Hartley, Claire Louise (2008) *Characterisation of factors that regulate homologous recombination and antigenic variation in Trypanosoma brucei.*
PhD thesis.

http://theses.gla.ac.uk/410/

Characterisation of factors that regulate homologous recombination and antigenic variation in *Trypanosoma brucei*

Claire Louise Hartley

Wellcome Centre for Molecular Parasitology
Glasgow Biomedical Research Centre
University of Glasgow

Supervisor: Dr Richard McCulloch

Submitted for the degree of Doctor of Philosophy
July 2008

Abstract

Trypanosoma brucei is an evolutionarily divergent eukaryotic parasite of mammals in sub-Saharan Africa and is transmitted by the tsetse fly vector. To evade the mammalian immune response, *T. brucei* utilises antigenic variation, which involves switches in the Variant Surface Glycoprotein (VSG) expressed on the cell surface. Such reactions can occur at very high rates (~10^{-3} switches/cell/generation) and occur primarily by the recombination of *VSG* genes, selected from an enormous silent archive, into specialised expression sites. It has been previously shown that such *VSG* switching is a form of homologous recombination, as mutation of *RAD51* and a related gene, *RAD51-3*, impairs the process.

BRCA2 has emerged as a significant regulatory factor during RAD51-catalysed recombination. In humans, BRCA2 contains eight BRC repeats, six of which have been shown to bind RAD51. Similar repeats are present in BRCA2 from other organisms, though normally in smaller numbers. This thesis describes a *T. brucei* BRCA2 homologue that appears exceptional in that it contains up to 12 BRC repeats. Furthermore, the sequence degeneracy that is observed between the BRC repeats in most organisms is absent in *T. brucei*, with all but the C-terminal proximal repeat being identical. It was hypothesised that this unusual BRCA2 organisation is due to the high levels of RAD51-directed recombination needed during antigenic variation.

To examine the function of the putative *T. brucei* BRCA2 homologue, mutants were generated and found to display impaired growth, sensitivity to induced DNA damage, impairment in the ability to form sub-nuclear RAD51 foci, a reduced ability to recombine DNA constructs into their genome and a reduction in frequency of *VSG* switching, all of which are consistent with roles for BRCA2 in DNA repair and recombination. Furthermore, genome instability in the mutants was observed through the loss of silent *VSG* gene copies and substantial reductions in the size of the mega-base chromosomes. Interestingly, other chromosome classes (the so-called mini- and intermediate-chromosomes) appear not to be susceptible to such instability.

A potentially novel function for BRCA2 was identified through DNA content analysis of the *T. brucei BRCA2* mutants. Mutation of *BRCA2* was shown to result in an accumulation of cells with aberrant DNA content that is most readily explained by an increased number of cells that undergo cytokinesis without having completed nuclear division, phenotypes that are not observed in other *T. brucei* recombination mutants, such as *RAD51*. This result

suggests that BRCA2 has a role in the regulation of cell division, with mutation causing impaired replication of *T. brucei* nuclear DNA, but without a cell cycle stall, leading to the accumulation of chromosomal aberrations.

In order to investigate the potential role of *T. brucei* BRCA2 in DNA replication and the unusual BRC repeat organisation phenotypes further, various truncations of BRCA2 were expressed in a mutant background. Cell lines expressing BRCA2 with only 1 BRC repeat displayed reduced efficiency in recombination, DNA repair and RAD51 foci formation, indicating that the large BRC repeat expansion in *T. brucei* BRCA2 plays a critical role in the proteins function. Expression of a BRCA2 variant encompassing only the region of the protein, C-terminal to the BRC repeats appeared able to function, at least partially, in regulating cell cycle progression. Moreover, this DNA replication role appears not to be provided by conserved DNA binding motifs present within the C terminus of BRCA2 since a fusion of *T. brucei* BRCA2 and the parasites homologue of the replication protein A 70 kDa subunit was impaired in cell division, but was proficient in repair of DNA damage. Taken together, these data infer that *T. brucei* BRCA2 possesses a function that is distinct from BRCA2's role as a regulator of RAD51, and acts in DNA replication or cell division.

In addition to the above research on BRCA2, I sought to examine the factors that interact with RAD51 in *T. brucei*. This work demonstrated that it is possible to add an epitope tag for tandem affinity purification (TAP) to the N-terminus of RAD51 in both the bloodstream and procyclic stages of *T. brucei* without disrupting its function. Preliminary data suggest that TAP is potentially a feasible way of examining RAD51 interacting factors.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would firstly like to thank Richard, not only for coming up with the initial ideas for this thesis, but also for his advice and support along the way.

I would also like to thank all members of WCMP, past and present, for helping to create a fun working atmosphere within the laboratory. Your ideas and suggestions have been really appreciated, as have regular coffee breaks and visits to the pub.

I would like to thank my parents, who continue to support me both emotionally and financially. Hopefully, a day will soon arrive when I can finally afford a basic standard of living!

Finally, I would like to thank Joe, who has been through all the rough times and all the great times with me. Thanks for always being there to support me and for putting up with my sometimes erratic mood swings!

Thanks also go to the Medical Research Council for funding and to the Roberts Fund for awarding me a travel scholarship.

Author's Declaration

I declare that this thesis and the results presented within it are entirely my own work except where otherwise stated. No part of it has been previously submitted for a degree at any university.

Claire Hartley

Definitions

A	adenine
aa	amino acid
ADP	adenosine diphosphate
APS	ammonium persulphate
ATP	adenosine triphosphate
BER	base excision repair
BES	bloodstream expression site
BIR	break-induced replication
BLE	bleomycin resistance protein gene
bp	base-pairs
BSA	bovine serum albumin
BSD	blasticidin–S-deaminase gene
C	cytosine
CBSS	Carters balanced salt solution 1 x: 0.023M HEPES, 0.12M NaCl, 5.41mM KCl, 0.55mM $CaCl_2$, 0.4mM $MgSO_4$, 5.6mM Na_2HPO4, 0.035M glucose, 0.04mM phenol red, pH adjusted to 7.4.
cDNA	complementary DNA
CIP	alkaline phosphatase, calf intestinal

DAPI	4, 6-diamidino-2-pheylindole
DEPC	diethyl pyrocarbonate
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DSB	double strand break
dsDNA	double-stranded DNA
EATRO	East African Trypanosomiasis Research Organisation
EDTA	ethylenediaminetetraacetic acid
ELC	expression linked copy
ESAG	expression site associated gene
ESB	expression site body
EtBr	ethidium bromide
FITC	fluorescein isothiocyanate
G	guanine
gDNA	genomic DNA
GPI	glycophospatidylinositol
GPI	glucose-6-phosphate isomerase.

HJ	Holliday junction
HR	homologous recombination
HYG	hygromycin phosphotransferase gene
IDL	insertion/deletion loop
ILTat	International Laboratory for research on animal diseases, Trypanozoon antigen type
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase-pairs
LB broth	Luria-Bertani broth
Mb	megabase-pairs
MES	metacyclic expression site
MITat	Moltino Institute – *Trypanozoan* antigen type
MMS	methyl methanesulphonate
MNE	MOPS/Sodium acetate/EDTA buffer 1 x: 0.024M MOPS, 5mM NaOAc, 1mM EDTA. pH adjusted to 7.0 with NaOH and stored in the dark.
mRNA	messenger RNA
MRN complex	Mre11, Rad50, Nbs1 complex (in mammals)
MRX complex	Mre11, Rad50, Xrs2 complex (in yeast)

NDS	solution for the manufacture of genomic plugs
	1 x: 0.5M EDTA, 0.5M TRIS base, 0.5M NaOH, 17mM lauroyl sarcosine. pH adjusted to 8.0 or 9.0 with NaOH.
NER	nucleotide excision repair
NHEJ	non-homologous end joining
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PSG	phosphate/ sodium chloride/ glucose buffer
	1 x: 0.06M Na_2HPO_4, 3.6mM NaH_2PO_4, 46mM NaCl, 55mM glucose, pH adjusted to 8.0.
PUR	puromycin-N-acetyltransferase gene
RNA	ribonucleic acid
RNAi	RNA interference
RPA	replication protein A
rRNA	ribosomal RNA
RT	reverse transcriptase
RT-PCR	reverse transcription polymerase chain reaction
SDS	sodium dodecyl sulphate
SDSA	synthesis-dependent strand annealing
SNP	single nucleotide polymorphism

SOB	bacterial media (per litre)
	20g bacto-tryptone, 5g bacto-yeast extract, 0.5g NaCl.
SOC	SOB + 20mM glucose
SRA	serum resistance-associated gene
SSA	single-strand annealing
SSC	sodium chloride / sodium citrate solution
	1 x: 0.15 M NaCl, 0.015 M $Na_3C_6H_5O_7.2H_2O$
ssDNA	single-stranded DNA
T	thymine
TAE	TRIS / acetate / EDTA buffer
	1 x: 0.04M TRIS base, 0.04M glacial acetic acid, 1mM EDTA
TBE	TRIS / borate / EDTA buffer
	1 x: 0.089M TRIS base, 0.089M ortho-boric acid, 2mM EDTA
TB1/10E	TRIS / borate /1/10 EDTA buffer
	1 x: 0.089M TRIS base, 0.089M ortho-boric acid, 0.2mM EDTA
TE	10 mM Tris.Cl, 1 mM EDTA
TEMED	N, N, N', N' - Tetramethylethylenediamine
TLF	trypanosome lytic factor
TREU	Trypanosomiasis Research Edinburgh University
UTR	untranslated region
UV	ultraviolet

VAT   variable antigen type

VSG   variant surface glycoprotein

X-Gal   5-bromo-4-chloro-3-indolyl-β-D-galactoside

ZM   Zimmerman postfusion medium

ZMG   Zimmerman postfusion medium supplemented with glucose
132 mM NaCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 0.5 mM
magnesium acetate, 0.09 mM calcium acetate [pH 7.0], 1% glucose

CHAPTER 1

Introduction

1.1 General introduction

Trypanosoma brucei is a protozoan parasite which belongs to the family *Trypanosomatidae*, of the class *Kinetoplastida* and the phylum *Euglenozoa*. The *Kinetoplastida* are highly divergent and present an unusual aspect of biology, such as extensive trans-splicing and RNA editing (Sogin *et al.*, 1986;Sogin *et al.*, 1989). Phylogenetic analysis of many eukaryotic protein sequences suggest that *Kinetoplastida* diverged from a common ancestor approximately 1.98 billion years ago (figure 1.1) (Hedges *et al.*, 2004). They are characterised by the possession of the kinetoplast, which is an unusual DNA network that forms an integral part of the mitochondrial system. The *Trypanosomatidae* include many vertebrate parasites (Maslov *et al.*, 2001), such as *Leishmania, Trypanosoma* and *Endotrypanum*, all of which are transmitted by insects. The *Trypanosoma* genus includes *Trypanosoma cruzi,* which causes Chagas disease in the New World, and the African Salivarian branch including *Trypanosoma congolense, Trypanosoma vivax* and *Trypanosoma brucei*, which causes Human African Trypanosomiasis (or Sleeping sickness) in humans and Nagana in cattle.

African trypanosomes are typically distinguished from other organisms belonging to the *Kinetoplastida* by their ability to survive extra-cellularly, due to having evolved antigenic variation processes for immune evasion. They undergo antigenic variation through the periodic switching of their major surface antigen, Variant Surface Glycoprotein (VSG). This process allows the parasite to establish chronic infections in the vertebrate hosts, which can often be fatal. Most research to date has been conducted on *T. brucei*, due to it being the only one to infect humans, and because *in vitro* growth is relatively easy in laboratory conditions. *T. congolense* and *T. vivax,* whilst not being infective to humans, are important livestock pathogens, with the former being the most widespread and the latter, the most pathogenic.

T. brucei is classified into three subspecies: *T. brucei brucei*, which does not have the ability to infect humans and *T. brucei gambiense* and *T. brucei rhodesiense*, which are human infective. *T. b. brucei* is believed to be incapable of infecting humans due to a high-density lipoprotein, called trypanosome lytic factor (TLF), that is found in human serum and lyses the parasites (Hajduk *et al.*, 1992;Smith and Hajduk, 1995;Smith *et al.*, 1995). *T. brucei gambiense* and *T. brucei rhodesiense*, conversely, are resistant to lysis *in vitro* by human serum and are therefore capable of infecting humans, a phenotype which has been attributed to the action of one or a few genes (Turner *et al.*, 2004). In *T. b. rhodesiense,* the product of the serum resistance associated gene (*SRA*) has been

determined as being responsible for this phenotype (Oli *et al.*, 2006). However, the extent
to which these subspecies are truly distinct has recently been called into question, with the
presence of an intermediate phenotype in *T. b. brucei* being uncovered, which has the
ability to develop resistance upon prolonged exposure to human serum (Turner *et al.*,
2004). This intermediate phenotype could well have arisen due to genetic exchange
occurring between the subspecies (section 1.1.2).
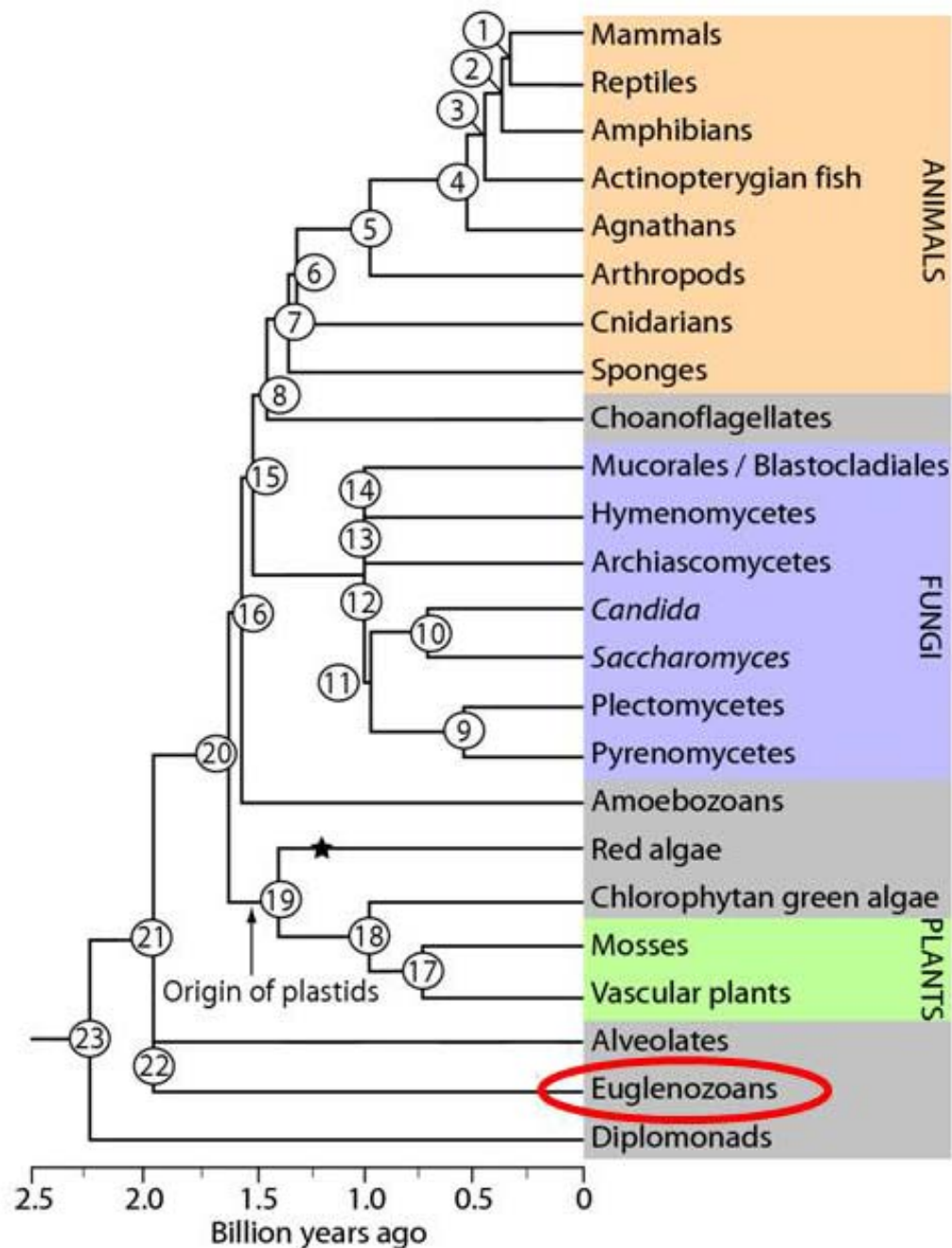
Figure 1.1 – A phylogenetic tree of a number of eukaryotes. A phylogenetic tree of a number
of eukaryotic organisms was constructed based upon a number of protein sequences and used to
calculate the divergence of various organisms from the tree. *T. brucei* belongs to the Euglenozoan
order, circled in red. Taken from Hedges *et al.*, 2004.

1.1.1 Symptoms, prevalence and treatment of Human African Trypanosomiasis

African trypanosomes are extracellular parasites, which proliferate in the lymphatic and vascular systems of their mammalian host. The early stage of the disease is characterised by fever, anaemia, lack of appetite and wasting caused by interstitial inflammation and necrosis within the capillaries of major organs (Vickerman, 1985). If the infection is allowed to progress, the parasites eventually cross the blood brain barrier. This late stage infection is characterised by motor and sensory disorders, sleep disturbances, followed by seizures and finally coma. If untreated, sleeping sickness is always fatal (Sternberg, 2004).

In 1986, the World Health Organisation estimated that 70 million people lived in areas where *T. brucei* disease transmission could take place, namely in sub-Saharan and equatorial Africa. In 1998, almost 40,000 cases were reported, but it was estimated that the actual number of cases was between 300,000 and 500,000. The seriousness of the disease was further highlighted by the fact that in certain areas sleeping sickness was considered to be a greater cause of mortality than HIV or AIDS. By 2005, surveillance had been reinforced and the number of new cases reported throughout the continent had substantially reduced. Currently, the estimated number of cases lies between 50,000 and 70,000 (http://www.who.int/mediacentre/factsheets/fs259/en/).

Only four drugs are registered for the treatment of sleeping sickness, and all are associated with major problems, including side effects and an increasing rate of treatment failure (Barrett *et al.*, 2003;Kennedy, 2004). Pentamidine is used in the treatment of the first stage of a *T. b. gambiense* infection and is a diamidine compound with antiprotozoal activity. Some of the observed side effects from pentamidine treatment include nephrotoxicity and pancreatic damage. A second drug, suramin, is used for first stage *T. b. rhodesiense* infections and is a complex derivative of urea with antiprotozoal activity. This drug has the ability to enter extracellular spaces but cannot cross the blood-brain barrier. Some of the observed adverse effects are heavy proteinuria, stomal ulceration, exfoliative dermatitis, severe diarrhoea, prolonged high fever and prostration. Melarsoprol is an organic arsenical compound that has the ability to enter the central nervous system, thereby making it a suitable drug for treating second stage cases of *T. b. gambiense* and *T. b. rhodesiense* infections. Fatalities have been known to occur with this treatment, but the most common side effects include headache, tremor, slurring of speech and convulsions. Finally, eflornithine is used for the treatment of both early and late stage *T. b. gambiense* infections. It is an ornithine derivative that acts by inhibiting the enzyme ornithine

decarboxylase, which is involved in polyamine synthesis in trypanosomes. The most common side effects include diarrhoea, anaemia, leukopenia, thrombocytopenia and convulsions.

1.1.2 The life cycle of T. brucei

T. brucei has a complex life cycle, proliferating in both the bloodstream of the mammalian host, and in the midgut and salivary glands of the tsetse fly (*Glossinidae* family). Each of these provides a contrasting environment and the parasite has therefore developed several distinct life cycle stages, which allow it to proliferate and transmit through each stage (figure 1.2).

The metacyclic form of the trypanosome is transmitted into the mammalian host when the tsetse fly takes its blood meal, injecting trypanosomes below the skin. Metacyclic form cells develop in the salivary glands of the tsetse fly and possesses an MVSG (Metacyclic Variant Surface Glycoprotein) coat (Tetley *et al.*, 1987), which serves to not only protect against the alternative complement system, but also to hide the invariant surface molecules from the hosts acquired immune system. In order for the trypanosome to be capable of transmitting from the fly into a vertebrate host, the cell cycle arrests, lying in the G0 phase. If the metacyclic form cells fail to transmit into a mammalian host, the cells die (Matthews and Gull, 1997;Shapiro *et al.*, 1984).

Once present in the mammalian bloodstream, the non-dividing metacyclic form trypanosomes differentiate into the long slender bloodstream form trypanosomes. These are distinguishable physiologically, and also by the fact that the MVSG coat is replaced with bloodstream form VSGs and the cells proliferate by rapid mitotic division. An infection is capable of establishing and being maintained in the mammalian host by the trypanosome population evading the host's immune system. This is achieved by antigenic variation, which occurs by periodic, spontaneous changes of the VSG being expressed, which contributes to peaks of parasitaemia that correspond to parasite populations expressing antigenically distinct VSGs (Capbern *et al.*, 1977). The fluctuations in parasitaemia result from immune reactions generated against VSGs causing destruction of a majority of the parasite population. However, a small proportion that switched their VSG to an antigenically distinct variant manages to escape. These continue to proliferate and establish another parasitemic peak, but are, in turn, eliminated by host immunity. Another contributing factor to this succession of parasitemic waves is the differentiation of

long slender form trypanosomes into non-dividing short stumpy forms, which occurs as the parasitaemia peaks (Gruszynski *et al.*, 2006).
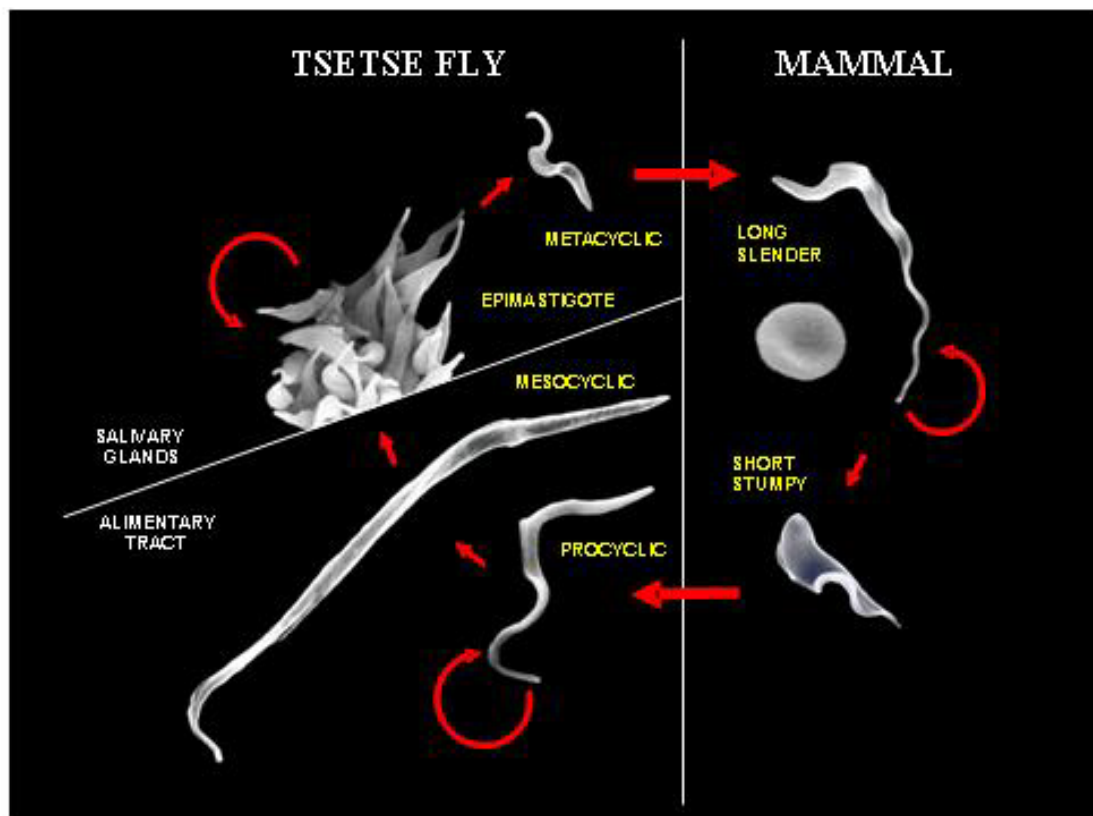
Figure 1.2 – The *T. brucei* life cycle. *T. brucei* life cycle stages are shown as scanning electron micrographs, shown to scale; an erythrocyte is shown next to the long slender bloodstream stage for comparison. The host organism and the name of the life cycle stage are indicated. Circular arrows represent replicative stages, whereas straight arrows represent differentiation and progression through the life cycle. Taken from Barry and McCulloch, 2001.

This differentiation is density dependent, which occurs due to the accumulated secretion of a low molecular weight factor, termed stumpy induction factor (SIT), from the long slender form trypanosomes. The SIT induces a growth arrest through a cyclic adenosine monophosphate (cAMP) signalling pathway (Vassella *et al.*, 1997). The short stumpy form trypanosomes, similar to the metacyclic forms, have a finite life if they are not transmitted to the tsetse fly when the fly feeds (Turner *et al.*, 1995). In addition, the short stumpy form trypanosomes are pre-adapted to life in the tsetse fly, with metabolic changes which allow them to switch from the glucose energy source in the bloodstream to the proline energy source found in the tsetse fly midgut (Hendriks *et al.*, 2000). However, despite this, the majority of trypanosomes do not survive long enough to differentiate to procyclic form cells (Van den Abbeele *et al.*, 1999).

Once ingested by the tsetse fly, short stumpy form trypanosomes differentiate into procyclic form cells within hours of the tsetse fly ingesting its feed (Hendriks *et al.*, 2000;Matthews *et al.*, 2004). This differentiation involves cell lengthening, re-positioning

of the kinetoplast, expression of procyclins on the cell surface and release from cell cycle arrest (Roditi *et al.*, 1998;Liniger *et al.*, 2004). The procyclin coat consists of different variants of procyclins, each of which are composed partly by internal amino acid repeat motifs, and expressed differentially throughout the tsetse infection (Acosta-Serrano *et al.*, 2001;Roditi and Liniger, 2002). EP procyclin contains internal Glu-Pro repeats, whilst GPEET procyclin contains Gly-Pro-Glu-Glu-Thr repeats. Whilst the exact function of the procyclin coat is unknown, it is thought to protect against trypanocidal factors in the tsetse midgut and prevent differentiation from tsetse specific factors (Roditi and Liniger, 2002).

Procyclic form trypanosomes continue to proliferate in the midgut of the tsetse fly before migrating to the anterior end of the midgut and differentiating into the very long mesocyclic form. Mesocyclic form trypanosomes enter the tsetse fly's foregut and proboscis whilst simultaneously replicating their DNA to become 4N. An asymmetrical division then occurs, which results in a small daughter cell that differentiates into the epimastigote form and a larger daughter cell, which is presumed to be unable to differentiate and therefore fails to survive (Van den Abbeele *et al.*, 1999).

Epimastigote form trypanosomes are proliferative cells and migrate to the salivary glands, where they become attached. It is here that genetic exchange between different *T. brucei* strains is thought to take place (Tait and Turner, 1990;Gibson *et al.*, 1995), contributing to genetic diversification of the parasite (Schweizer *et al.*, 1988). Finally, differentiation of the epimastigote form trypanosomes results in the mammalian infective metacyclic form trypanosomes, completing the life cycle.

It is important to note that some laboratory strains have lost the ability to be transmitted through the fly and have been termed monomorphic as only the long slender bloodstream form is present in the vertebrate stage. Other strains, which are capable of completing the life cycle, have subsequently been termed pleomorphic (Matthews and Gull, 1994;Wijers and Willet, 1960).

1.1.3 The genome of T. brucei

The *T. brucei* genome consists of 11 diploid megabase chromosomes, a set of intermediate sized chromosomes and a large number of mini-chromosomes. The megabase chromosomes of the strain TREU 927/4 have been sequenced (Berriman *et al.*, 2005), and represent a haploid genomic content of 26 Mb, containing 9068 predicted genes that includes ~900 pseudogenes and ~1700 *T. brucei* specific genes. Megabase chromosomes vary in size from 0.9 to over 6 Mb and have been named I to XI in order of increasing size

(Melville *et al.*, 1998). Genomic content can vary by as much as 25 % between *T. brucei* strains, as can the sizes of individual chromosomes between allelic copies in the same strain (Melville *et al.*, 2000;El Sayed *et al.*, 2000). This large degree of fluctuation is thought to be largely due to telomeric and subtelomeric rearrangements that are associated with antigenic variation (Callejas *et al.*, 2006).

The intermediate chromosomes vary in number between strains (between 1and 7), are of uncertain ploidy and range in size from 200-900 kb (Wickstead *et al.*, 2004). The ~100 mini-chromosomes range from 50-150 kb and are composed of mainly repetitive, palindromic sequences, known as the 177 bp repeats, which are also present in the intermediate chromosomes (Wickstead *et al.*, 2004;El Sayed *et al.*, 2000;Wickstead *et al.*, 2003a). The 177 bp repeats have been shown to be present in the core of these smaller chromosomes in an inverted symmetry (Wickstead *et al.*, 2004). These repeats are less abundant in the intermediate chromosomes and vary in the length of non-repetitive subtelomeric sequences they possess (Wickstead *et al.*, 2004). It has been suggested that these 177 bp repeats play an important role, whereby they help in the maintenance of the mini-chromosomes and intermediate chromosomes through associating with replication bubbles (Weiden *et al.*, 1991). To date, only *VSG* and expression site associated genes have been found on the minichromosomes (Wickstead *et al.*, 2004;Rudenko *et al.*, 1998;Melville *et al.*, 2000), and none have been found to possess an active *VSG* expression site, unlike in the intermediate chromosomes. This therefore suggests that in order for these *VSGs* to be expressed, they must either be duplicated into an active expression site or be part of a telomere exchange with one (El Sayed *et al.*, 2000).

1.1.4 Transcription and translation

The genes of *T. brucei* are orientated unidirectionally over long distances and are thought to be transcribed polycistronically (Berriman *et al.*, 2005;El Sayed *et al.*, 2003;Johnson *et al.*, 1987). Mature mRNA is produced from the polycistronic transcript by a process of *trans*-splicing and polyadenylation. *Trans*-splicing involves the addition of a 39 nt capped RNA, termed the 'spliced leader' to the 5′ end of mRNA, whilst polyadenylation of mRNAs occurs at the 3' end (Ullu *et al.*, 1993;Matthews and Gull, 1994;Clayton, 2002). It appears that the *trans* splicing and polyadenylation are inextricably linked, since inhibiting either process prevents the other (Ullu *et al.*, 1993). Addition of the spliced leader RNA adds a cap to the mRNA, and this structure appears unique to kinetoplastids, consisting of 7-methylguanosine and 4 methylated nucleotides (Bangs *et al.*, 1992). Polyadenylation signals also appear unusual, due to being poorly defined and occurring at a fixed distance

upstream of the splice signal of the downstream gene in the polycistron (Matthews and Gull, 1994).

T. brucei transcription also appears unusual in the action of the polymerase enzymes involved. In most eukaryotes, transcription is mediated by 3 types of RNA polymerase: RNA pol I generates rRNA, RNA pol II produces mRNA, and RNA pol III yields tRNA (Rutter W Jr *et al.*, 1976;Cramer, 2002;Tamura *et al.*, 1996). In *T. brucei*, it is RNA pol II that appears to be primarily responsible for polycistronic transcription (Devaux *et al.*, 2006). Notable exceptions are found in VSG and procyclin expression, which are transcribed by RNA pol I (Navarro and Gull, 2001). This is the only example of RNA pol I transcription directing expression of protein-coding mRNAs, and is thought to occur in order to allow for the high of levels transcription needed for these abundant proteins (Gunzl *et al.*, 2003). In contrast, tubulin-encoding mRNAs, which are transcribed by RNA polymerase II, are found in relatively similar abundance because multiple copies of the gene are found in the genome (Kooter and Borst, 1984).

1.2 Phase and antigenic variation

Pathogenic organisms face many challenges in order to ensure their long term survival within a host. These include the crossing and colonisation of novel surfaces, such as endothelia, and the host's specific immune response. Often it is too late if the organism only activates a phenotypic change once it has sensed an alteration in the environment. Instead, many organisms have adopted a strategy which generates diversity in the population before the challenge arises. Such a strategy is the spontaneous mutation of a set of genes that have been termed contingency genes (Moxon *et al.*, 1994). These contingency genes undergo spontaneous mutation at rates that are higher than the background rate of housekeeping genes (10^{-6}) (Barry *et al.*, 2003) and provide the organism with a level of diversity within the population that should allow for selection of individuals that can respond to environmental changes and continue the infection. The functions of these contingency genes are diverse and include attachment to host surfaces, cell invasion and cell protection. Moreover, such contingency genes are found in a variety of pathogens, including viruses, fungi, bacteria and protozoans (Deitsch *et al.*, 1997).

Phase variation is a reversible process that allows two distinct states to be switched between using mechanisms such as promoter inversion mediating gene transcriptional switching, recombination-mediated genetic rearrangements and slipped strand mis-pairing. An example of the latter mechanism is found in *Haemophilus influenzae*. *H. influenzae* is

one of the causative agents of bacterial meningitis and escapes specific immune responses by the lipopolysaccharide (LPS) on its cell surface undergoing structural changes. Through the loss and gain of tandem repeats (CAAT) in the coding region of LPS biosynthesis genes, translation initiation codons are placed in and out of frame, resulting in the switching on and off of gene expression (Levinson and Gutman, 1987). Examples of recombination mediated genetic rearrangements are present in *Escherichia coli* and *Neisseria gonorrhoea*, which utilise site specific recombination in the *fimA* gene and recA-dependent homologous recombination involving the *pilin* genes, respectively (Kulasekara and Blomfield, 1999;Mehr and Seifert, 1998).

Antigenic variation is another important mechanism in prolonging pathogen survival, but differs from phase variation in that the process is not phenotypically reversible and, indeed, is more complex by virtue of a progressive switching between multiple states, rather than switching between two states. The purpose of antigenic variation is to evade the acquired immune system of the host and, as such, occurs by switches solely involving surface molecules (antigens).

Arguably, the best studied example of antigenic variation is in *T. brucei*, and this will be discussed at length in section 1.3. Well-studied examples of antigenic variation in protozoa are found in *Plasmodium* (Kraemer and Smith, 2006) and *Giardia* (Nash, 2002). In bacterial systems, antigenic variation has been documented in the spirochetes *Borrelia hermsii* (Dai *et al.*, 2006) and *B. burgdorferi* (Zhang *et al.*, 1997), as well as in *Anaplasma marginale* (Futse *et al.*, 2005) and *Neisseria gonorrhoeae* (Sechman *et al.*, 2005;Zhang *et al.*, 1992).

Plasmodium falciparum, the causative agent of malaria, infects erythrocytes in the host bloodstream, causing their surface morphology to alter and subsequently be targeted for destruction in the spleen. The parasite counteracts this by adhering to endothelia of the host blood vessels by expressing PfEMP1 (*P. falciparum* erythrocyte membrane protein 1) molecules on the red blood cell surface. PfEMP1 molecules are encoded by the *var* genes, and the parasite avoids destruction by acquired immunity against PfEMP1 through transcriptional switching between a repertoire comprising 50 -150 *var* genes copies (Kyes *et al.*, 2007;Deitsch *et al.*, 1997). *Giardia lamblia*, the causative agent of giardiasis, also utilises transcriptional control mechanisms to switch between a repertoire of 150 genes, which encode the variant surface protein (VSP) (Mowatt *et al.*, 1991;Kulakova *et al.*, 2006).

1.3 Antigenic variation in T. brucei

Trypanosoma brucei was the first organism discovered to undergo antigenic variation as a means of escaping the immune defence system (Vickerman, 1978;Borst, 1986;Borst and Greaves, 1987) and remains one of the best-studied examples of this process (reviewed in Barry and McCulloch 2001; McCulloch 2004; Pays 2006; Taylor and Rudenko 2006 and http://www.VSGdb.org/).

Infections with *T. brucei* are characterised by successive waves of parasitaemia, in which the Variant Surface Glycoproteins (VSG) expressed during one wave are different from those of the preceding ones (Pays, 1985;Roth *et al.*, 1991;Borst and Rudenko, 1994;Cross, 1990). This typical pattern of infection is displayed in figure 1.3 and demonstrates how the infection persists until the host is treated or dies of the disease (Molyneux, 1983).
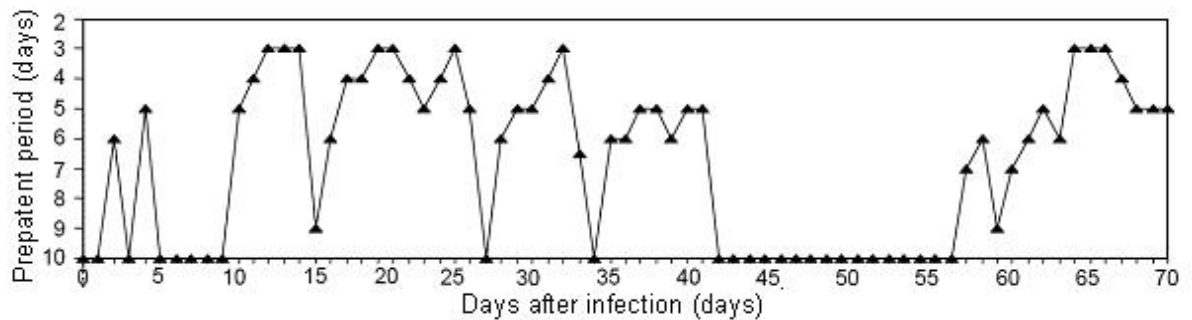
Figure 1.3 – Parasitic wave of a *T. brucei* infection in a cow. Chronic infection of a cow with *T. brucei* strain ILTat 1.2. The timeline of the 70-day infection is shown on the x-axis, whereas parasite density in the cow is shown as prepatent period on the y-axis; the number of days for parasitaemia to reach a certain level following inoculation of an immunosuppressed mouse with cattle blood. Figure taken from Morrison *et al.*, 2005.

Antigenic variation in *T. brucei* is mediated by VSGs, which generate thick glycoprotein coats that completely envelope the bloodstream stage parasites. The VSG coat serves to hide invariant antigens on the cell surface and also helps to protect against innate immune responses, such as phagocytosis (Cross, 1975;Turner *et al.*, 1988). Generally, a single *T. brucei* cell expresses a single VSG at a time. Throughout an infection subpopulations of trypanosomes expressing antigenically different VSGs arise due to reactions that cause switches in the expressed VSG. As a result of such switching, these subpopulations are able to escape the antibody mediated response to the parental populations expressing the preceding VSG. In turn, the switched subpopulations are cleared by antibodies against the VSG, but further subpopulations expressing novel VSGs are continually generated, allowing the infection to continue. The extent of the antigenic variation is determined by the large number of *VSG* genes in the genome that are devoted to antigenic variation, and

the mechanism involved in the switching process (see below). Despite the potential for antigenic variation to generate thousands of distinct VSGs, hosts that have been continuously exposed to infection eventually acquire a degree of immunity to re-infection (Browning and Gulbransen, 1936). However, repeated immunological responses can eventually lead to prolonged immuno-depression, which in turn is detrimental to the host.

1.3.1 VSG in T. brucei

Variant Surface Glycoprotein molecules cover the surface of bloodstream form *T. brucei* cells in a densely packed monolayer (figure 1.4). It is this packaging that serves to protect the invariant surface molecules from the immune system (Borst and Fairlamb, 1998;Overath *et al.*, 1994) and also protects against innate responses (Turner *et al.*, 1988). VSG-specific antibody responses are raised against a small part of the VSG molecule (Cross, 1990), the hypervariable N-terminal domain (Berriman *et al.*, 2005).

The surface of each *T. brucei* cell possesses approximately 5.5×10^6 VSG homodimers (Cross, 1975;Auffret and Turner, 1981), which are attached to the plasma membrane by a glycosylphosphatidylinositol (GPI) anchor (Ferguson *et al.*, 1988). The VSGs are continuously endocytosed and recycled at a high rate via the flagellar pocket (Overath and Engstler 2004). The VSG homodimers typically consist of 400 to 500 amino acid residues and are composed of an elongated N terminal domain consisting of 350-400 residues and one or two smaller C terminal domains consisting of 40-80 residues each (Pays *et al.*, 2007). Whilst the sequences of VSGs display high levels of diversity, the three dimensional structures appear to be well conserved (Blum *et al.*, 1993;Chattopadhyay *et al.*, 2005). Greater levels of diversity are found within the N terminus, which adopts an alpha helical coiled-coil structure containing exposed surface loops. Conversely, greater levels of conservation are observed at the carboxyl terminus, where the VSGs are anchored at the parasite membrane by GPI linkage to ethanolamine (Ferguson *et al.*, 1988). Each VSG is defined as a combination of an N terminal type and a C terminal type, based on the distribution of cysteine residues in the molecule (Carrington *et al.*, 1991).
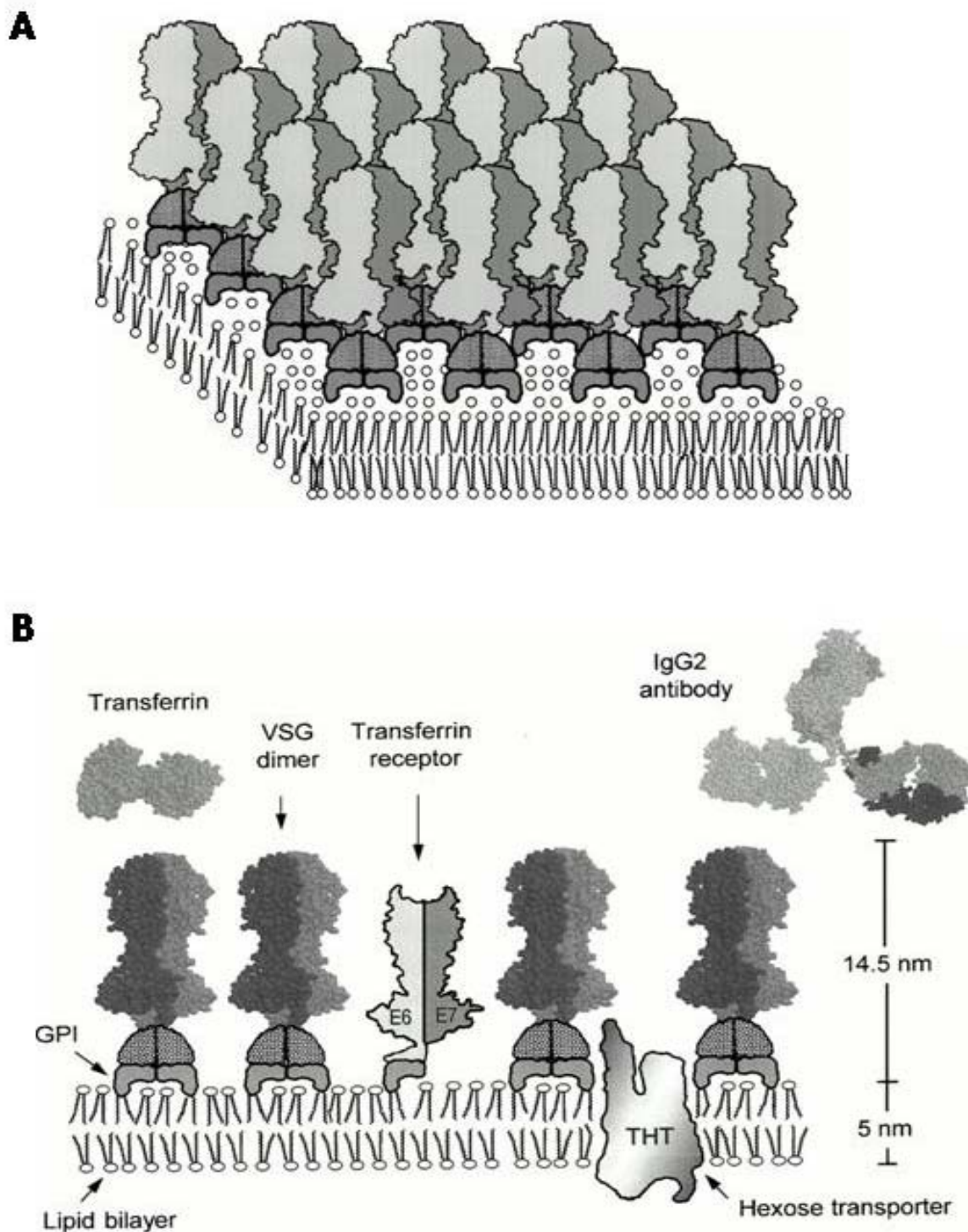
Figure 1.4 – The cell surface of bloodstream form _T. brucei_. (A) A three dimensional depiction of the tightly packed VSG dimers present on the cell surface of bloodstream form _T. brucei_. (B) A schematic representation of the cell surface of bloodstream form _T. brucei_. VSG dimers (attached to the GPI anchor via the C-terminal domain), a transferrin receptor and a hexose transporter are shown associated with the plasma membrane. A transferrin and immunoglobulin G (IgG2) molecule are also shown for size comparison. Taken from Borst and Fairlamb, 1998.

The *T. brucei* genome contains a huge repertoire of silent *VSG* genes, which facilitate the process of antigenic variation. It was previously thought that the genome contained approximately 1000 *VSGs* (Van der Ploeg *et al.*, 1982). However, with the aid of the genome sequencing project looking at strain TREU927, this has recently been updated, and the number of *VSGs* on the megabase chromosomes has now been estimated at approximately 1600, with a further ~200 present on minichromosomes, plus ~20 at the VSG transcription loci known as bloodstream expression sites (see below) (Berriman *et al.*, 2005;Wickstead *et al.*, 2004;Marcello and Barry, 2007b;Marcello *et al.*, 2007). 940 of the *VSGs* in the main contigs of chromosomes have been analysed and, remarkably, the majority have been shown to be pseudogenes or gene fragments located at the subtelomeres of chromosomes (figure 1.5) (Barry *et al.*, 2005;Marcello and Barry, 2007b): only 5 % have been shown to be fully intact, encoding all known features of functional VSGs; 9 % of the VSGs were described as 'atypical', meaning they are predicted not to be accurately folded or modified; 62 % were full length pseudogenes, containing frame shifts or stop codons; and 19 % were gene fragments (Barry *et al.*, 2005).

It has recently emerged that the archives of surface antigens for protozoan pathogens are commonly found at subtelomeric locations. The reason for this is thought to be due to their proneness for ectopic recombination, thereby enabling the expansion of contingency gene families and promoting antigenic variation (Barry *et al.*, 2003).

The annotation of the VSG repertoire raises an important question: are the location and functional degeneration of the VSGs, unusual for organisms employing antigenic variation of surface antigens? The answer is almost certainly no. Pseudogenes are known to contribute to immune evasion, by providing substrates for gene conversions (see below), in bacterial systems such as in the spirochetes *Borrelia hermsii* (Dai *et al.*, 2006), and *B. burgdorferi* (Zhang *et al.*, 1997;Craig and Scherf, 2003). Locating most of the *VSG* genes at the subtelomeres of chromosomes is also not an unusual phenomenon (Barry *et al.*, 2003). Indeed, this strategy is employed in other pathogens such as *Pneumocystis carinii* (Keely *et al.*, 2005) and *P. falciparum* (Scherf *et al.*, 2001), with *Giardia lamblia* (Adam, 2000;Arkhipova and Morrison, 2001) providing a notable exception whereby the *vsp* genes are located interstitially. Subtelomeric locations are thought to contribute to antigenic variation due to the fact that they are prone to recombine ectopically during mitotic, and perhaps meiotic, recombination, promoting not only gene conversion reactions but also the expansion and sequence diversification of the gene families involved in antigenic variation (Barry *et al.*, 2003;Freitas-Junior *et al.*, 2000).

Figure 1.5 – Schematic representation of the silent *VSG* repertoire on the megabase chromosomes of *T. brucei* strain TREU 927/4. Chromosomes are represented by horizontal lines, with the chromosome number in a grey box to the left of each chromosome. Arrays of *VSG*s are depicted by black blocks; the orientation of sets of *VSG*s is shown by the position of the box above or below the line. The provisional number of *VSG*s in each array is shown. Breaks in contiguation are represented by oblique lines. Figure taken from Barry *et al.*, 2005.

Metacyclic form trypanosomes also encode a VSG coat, known as metacyclic VSG (MVSG). MVSG proteins are structurally indistinguishable from VSG proteins expressed in the bloodstream stage, but appear to be a specific type, of which 27 have been identified by immunological methods (Turner *et al.*, 1988). MVSGs are expressed for several days following infection of the vertebrate host, before the parasite replaces them with the bloodstream specific VSGs (Ginger *et al.*, 2002).

1.3.2 The expression sites of T. brucei

In order for *VSGs* to be expressed, they need to be localised in special polycistronic transcription units known as expression sites (ES) (Pays *et al.*, 2001). Up to 20 of these transcription sites are used in the bloodstream stage of the parasite. These are known as bloodstream expression sites (BESs) (figure 1.6) and are located at the telomeres of megabase and intermediate chromosomes (Becker *et al.*, 2004). In addition to the BESs, there is a set of expression sites reserved for *VSG* expression in the metacyclic stage of the parasite, known as metacyclic ESs (MESs) (figure 1.6) (Graham *et al.*, 1999;Bringaud *et al.*, 2001). MESs are also telomerically located.

Figure 1.6 – The expression sites of *T. brucei*. (A) The BESs shown are AnTat 1.3A, ETat 1.2CR, VSG10.1, 221ES, VO2 ES and Bn-2 ES. The *VSG* is indicated by a white box, the 50-bp and 70-bp repeat arrays (not to scale) by striped boxes, and the promoters by flags. The *ESAG*s and pseudo-*ESAG*s (indicated by ψ) are represented by numbered grey and black boxes. Retrotransposon Hot Spot genes (RHS) are annotated R, and Serum Resistance Associated (SRA) genes SRA. Taken from Berriman *et al.*, 2002. (B) Two MESs are shown. The ES promoters are indicated by a white flag, the 70 bp repeats by striped boxes, the ESAGs and pseudo-*ESAG*s (indicated by ψ) by white boxes, arrows indicate direction of transcription and black boxes represent the *VSGs*. Taken from Rudenko, 2000.

The BESs can vary in size from 40 kb to 100 kb, but all retain a conserved structure, with the *VSG* gene located closest to the telomere (Becker *et al.*, 2004;Berriman *et al.*, 2002). Upstream of the *VSG* lies a set of short repetitive sequences known as the 70 bp repeats (Liu *et al.*, 1983). These 70 bp repeats can span a distance of up to 20 kb in the BES, and are also found, albeit in shorter arrays, upstream of most silent *VSG*s (>90 %) within the genome (Barry *et al.*, 2005;McCulloch *et al.*, 1997). These repeats define the boundaries of many of the regions involved in *VSG* gene conversion reactions (section 1.4.2.3), leading to a hypothesis that the 70 bp repeats were involved in the initiation of *VSG* switching (Barry, 1997). However, deletion of the 70 bp repeats from the active expression sites did not affect duplicative transposition of *VSG* genes from silent expression sites (McCulloch *et al.*, 1997), showing that the reaction can occur in their absence.

Further upstream of the 70 bp repeats lies between 8 and 10 expression site associated genes (*ESAGs*) (Pays *et al.*, 1989). *ESAGs* 6 and 7 are the only *ESAGs* to date that have been identified in each BES, and have been found to encode the two subunits of a transferrin receptor, which allows the parasites to uptake iron from the host (Schell *et al.*, 1991;Ligtenberg *et al.*, 1994;Steverding *et al.*, 1994;Berriman *et al.*, 2002). Allelic variants of *ESAGs* 6 and 7 are found throughout the BESs and are thought to provide different affinities for transferrin from different hosts. It has been postulated that through the parasites use of different BESs, it can adapt to different hosts and therefore establish host specificity (Bitter *et al.*, 1998). However, this hypothesis has been questioned recently in a series of *in vitro* experiments (Salmon *et al.*, 2005). *ESAG 4* has been identified as encoding an adenylate cyclase (Paindavoine *et al.*, 1992) and the Human Serum Resistance gene (*SRA*) has also been identified within some BES, and therefore constitutes an *ESAG* (Xong *et al.*, 1998). The products of most of the other *ESAGs* still remain uncharacterized and none have been discovered to be directly involved in antigenic variation (Borst and Rudenko, 1994;Cross, 1996).

Transcription of the active BES is carried out by RNA polymerase I (Gunzl *et al.*, 2003). Directly upstream of the BES promoter lies a large array of repetitive sequences known as the 50 bp repeats, which can span up to 50 kb (Zomerdijk *et al.*, 1990;Zomerdijk *et al.*, 1991). This repeat region is thought to function as a barrier between the upstream sequences and the BES transcriptional unit (Sheader *et al.*, 2003).

MESs, like BESs, are also found in subtelomeric locations and are transcribed by RNA polymerase I (Graham *et al.*, 1999;Barry and McCulloch, 2001;Berriman *et al.*, 2002).

Where they differ is in their much simpler structure, which consists of a single *VSG* gene, located adjacent to the telomere, followed upstream, by a stretch of 70 bp repeats and a promoter (figure 1.6). This means that the MESs also differ from BESs and from the majority of kinetoplastid genes, by the fact that they are transcribed monocistronically (Alarcon *et al.*, 1994;Nagoshi *et al.*, 1995).

In bloodstream form *T. brucei*, the parasite ensures that only one *VSG* is expressed at a time by mechanisms that ensure that only one BES is actively transcribed at a time. Many organisms utilise such mono allelic expression (Borst, 2002). In *T. brucei*, this may be explained by the discovery of a sub nuclear body, known as the expression site body (ESB) (Navarro and Gull, 2001). The ESB is distinct from the nucleolus, contains RNA polymerase I and the single actively transcribed BES. Exactly how the ESB exhibits control over *VSG* transcription is unknown, but it is thought that the factors for BES transcription are sequestered in the ESB, and are therefore unavailable to the silent BESs. This body has been linked to a role in antigenic variation, not only for its association with transcribing BESs, but also due to the fact that it has only been observed in bloodstream form cells (Navarro and Gull, 2001).

Another unusual biological feature has been discovered in kinetoplastids, as well as in two distantly related organisms; *Diplonema* and *Euglena,* and may help to explain the mono allelic expression of *VSGs* (Borst and van Leeuwen, 1997;van Leeuwen *et al.*, 1998;Dooijes *et al.*, 2000). This is β-D-glucosyl-hydroxymethyluracil, or base J, which is a modified version of uracil and replaces a subset (~0.2 %) of thymine residues within the *T. brucei* genome (Gommers-Ampt *et al.*, 1991;Gommers-Ampt *et al.*, 1993;van Leeuwen *et al.*, 1997). The function of base J has not been clearly defined, but it has been implicated to have a role in antigenic variation. Base J has been localised to repeated sequences within the genome, telomeric repeats and in *VSG*s and other sequences within the silent BESs. Its notable absence from active BESs led to the hypothesis of a role in BES silencing, in which J directly blocks transcription activation or elongation (van Leeuwen *et al.*, 1996). The ESB could therefore be involved in the removal of J, resulting in BES activation. J is known to be bound by two J-binding proteins (JBP1 and JBP2) (Cross *et al.*, 1999;Cross *et al.*, 2002;Dipaolo *et al.*, 2005). The latter of these is a member of the SWI2/SNF2 family and functions in chromatin remodelling, leading to an alternative hypothesis for J function, not involving antigenic variation. In this hypothesis, J acts as an epigenetic marker of heterochromatin (Borst and Ulbert, 2001;Pays *et al.*, 2004). More recently, JBP2 null mutants have been generated in *T. brucei* and not only contain five fold less base J, but are also incapable of synthesising J in newly generated telomeric arrays

(Kieft *et al.*, 2007). Further research is currently underway to finally determine J's biological role.

1.3.3 The mechanisms of VSG switching in T. brucei

Multiple mechanisms exist by which a trypanosome can switch from one *VSG* to another. The switching events can be broadly described as being transcriptional-based or recombinational-based events. A third class of switching has also been proposed in which multiple point mutations arise either during the generation of a copy of a *VSG* gene or within the *VSG* silent archive (Donelson, 1995).

As previously mentioned, prolonged syringe passaging between mammalian hosts can result in loss of the parasites ability to differentiate beyond the long slender bloodstream stage. The resulting 'monomorphic' trypanosome cell lines also have a greatly depressed rate of antigenic variation, the causal reason of which is unknown, and switch at an overall rate of 10^{-6} to 10^{-7} switches/cell/generation (Lamont *et al.*, 1986). 'Pleomorphic' trypanosome cell lines that can undergo differentiation from long slender to short stumpy forms, in contrast, can switch at much higher rates of 10^{-2} to 10^{-3} switches/cell/generation (Turner and Barry, 1989;Turner, 1997). It should be noted that all of the work in this thesis was performed on the monomorphic, low switching cell lines, in the strain Lister 427 (Cross, 1975).

1.3.3.1 Transcriptional (*in situ*) switching

Transcriptional, or *in situ,* switching occurs by activating transcription from a silent BES and silencing transcription from the active BES (*in situ*; figure 1.7). This mechanism is not considered to significantly contribute to *VSG* switching, at least in the pleomorphic cell lines, since it is only able to occur between the BESs and therefore only a small number of *VSGs* (Robinson *et al.*, 1999). In monomorphic cell lines, however, transcriptional switching is considered to predominate, as the recombinational switches have been proposed to be repressed (Barry, 1997).

The mechanisms by which transcriptional switching occurs still remain unknown. However, it is generally considered that DNA rearrangements are not required, except in rare cases where the active BES is deleted (Cross *et al.*, 1998). The possibility that DNA repair mechanisms are involved, however, cannot be discounted, since it has been shown that genome wide DNA damage can trigger transcriptional activation of silent BESs (Sheader *et al.*, 2004). Some recent work may support this possibility by implicating that

the DNA repair factors RAD51 and RAD51-3 may be involved (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005).

Another potential mechanism proposed to explain transcriptional switching is telomere silencing, which was first discovered in *Saccharomyces cerevisiae* (Gottschling *et al.*, 1990). This process represses transcription in telomere proximal regions of the genome and was proposed to contribute to VSG expression regulation (Horn and Cross, 1995;Rudenko *et al.*, 1995). However, more recently doubt has been cast on the importance of telomere silencing in transcriptional-based VSG switching, due to the fact that VSG promoters appear to lie too far away from the telomere for this process to have an effect. Additional support is provided by the deletion of genes that are likely to be involved in telomere silencing and the deletion of telomeric repeats, the results of which did not affect antigenic variation (Pays *et al.*, 2004;Horn and Barry, 2005;Alsford *et al.*, 2007;Glover *et al.*, 2007).

1.3.3.2 Recombinational switching

DNA recombination is the most common route for *VSG* switching to occur since the majority of the *VSG* repertoire is located in non-transcribed loci, and can only be expressed through recombination into the active BESs. The major routes of recombination in *VSG* switching are duplicative transposition reactions, which occur by gene conversion events whereby genetic information is transferred from a silent locus into the active BES, deleting the existing *VSG* gene (Robinson *et al.*, 1999).

Duplicative transposition is largely thought to occur through gene conversion events involving the replacement of the VSG at the active BES with a silent VSG from a tandem array in a megabase chromosome (array VSG GC; figure 1.7), from a silent BES (ES VSG GC; figure 1.7), or from the subtelomere of a minichromosome (MC VSG GC; figure 1.7). These gene conversion reactions generally occur using the 70 bp repeats as upstream homology (Liu *et al.*, 1983;Matthews *et al.*, 1990) and the 3' end of the VSG ORF as downstream homology (Michels et al 1983), although this can extend beyond the ORF to the 3' UTR (Michels *et al.*, 1983;Timmers *et al.*, 1987).

In monomorphic cell lines, the gene conversion reactions can utilise homology from much further upstream. These reactions occur between BESs (ES GC; figure 1.8) and can use homology at least up to 6 kb upstream of the *VSG*, beyond the 70 bp repeats (Lee and Van der Ploeg, 1987). That these reactions need not rely on the 70 bp repeats is demonstrated by no observed effect on antigenic variation by their deletion (McCulloch *et al.*, 1997).

Another type of duplicative transposition reaction is a telomere conversion event (telomere conversion; figure 1.8) (Shah *et al.*, 1987;de Lange *et al.*, 1983). This gene conversion reaction again uses homology from the 70 bp repeats upstream (and could, in theory initiate further upstream from another BES), but the 3' end extends down to the telomere. Clearly, therefore, this reaction is limited to switches between telomeric *VSGs*. It is unclear whether this is a distinct mechanism from the duplicative transposition events described above, though it has been proposed that this pathway may occur by break induced replication (BIR) (Dreesen and Cross, 2006).

Telomere reciprocal exchange is another mechanism which utilises recombination events to drive *VSG* switching. Unlike gene conversion events, however, this reaction involves a simple crossover event between telomeric *VSGs*, with both *VSGs* remaining intact at telomeric locations. Homology can be obtained from the 70 bp repeats or further upstream in the BES (Pays *et al.*, 1985;Shea *et al.*, 1986). This reaction is generally considered to occur less commonly than gene conversion events since it is limited to the telomeric *VSGs*.

Finally, mosaic gene formation has been described to occur in *VSG* switching (mosaic gene formation; figure 1.8) (Thon *et al.*, 1990;Barbet and Kamper, 1993). This mechanism occurs by using two or more segmental gene conversion events from silent *VSGs*, which allows pseudogenes to contribute to antigenic variation. Rather than using flanking homology regions as used in duplicative transposition reactions, it utilises short regions of homology within the *VSG* ORFs. This pathway was originally considered to occur relatively late in infections when the intact *VSGs* has been recognised by the immune system and contribute to only a small percentage of *VSG* switching events (Barbet and Kamper, 1993). However, in light of the genome sequencing project, where it has been uncovered that the majority of *VSGs* that exist in the genome are in fact non-functional (Berriman *et al.*, 2005;Marcello and Barry, 2007b;Marcello and Barry, 2007a), it appears that mosaic gene formation has more of an important role than was previously considered. Indeed, it has been proposed that segmented gene conversion is key to the success of antigenic variation (Barbet and Kamper, 1993;Marcello and Barry, 2007b), allowing long-term survival in a single host and for anti-VSG immunity in host herds.
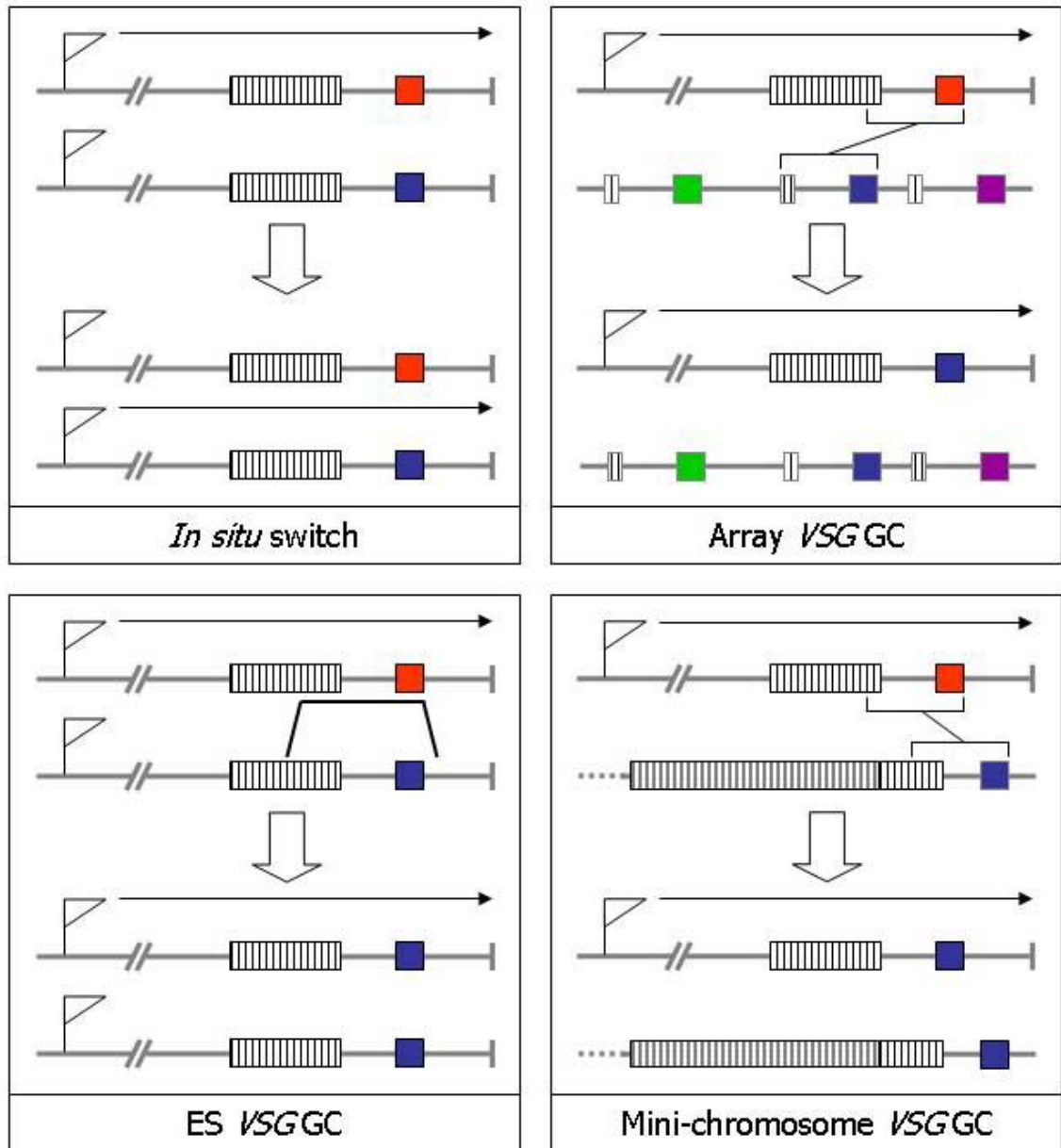
Figure 1.7 – VSG switching mechanisms used in *T. brucei*. A schematic representation of mechanisms of VSG switching. See text for full explanations. Horizontal grey lines represent chromosomal DNA, whereas vertical grey lines represent the end of a telomere. A flag depicts the BES promoter, and transcription from this promoter is shown by a horizontal black arrow. *VSG*s and *VSG* pseudogenes are represented by coloured squares and rectangles respectively. 70 bp repeat tracts are shown by black and white striped boxes, whilst the 177 bp repeats found in the mini-chromosomes are shown by grey and white striped boxes. The black lines show the extent of sequence copied into the expression site. GC – gene conversion. Adapted from C. Proudfoot, PhD thesis, 2005.

Figure 1.8 – VSG switching mechanisms used in *T. brucei* (cont.). A schematic representation of mechanisms of VSG switching. See text for full explanations. Horizontal grey lines represent chromosomal DNA, whereas vertical grey lines represent the end of a telomere. A flag depicts the BES promoter, and transcription from this promoter is shown by a horizontal black arrow. *VSG*s and *VSG* pseudogenes are represented by coloured squares and rectangles respectively. 70 bp repeat tracts are shown by black and white striped boxes, whilst the 177 bp repeats found in the mini-chromosomes are shown by grey and white striped boxes. The black lines show the extent of sequence copied into the expression site. GC – gene conversion. Adapted from C. Proudfoot, PhD thesis, 2005.

1.3.4 Use and timing of different switching mechanisms

The difference between monomorphic and pleomorphic cell lines in *VSG* switching appears not only to affect the rate of switching but also the mechanisms used. Early in an infection, monomorphic cell lines are thought to primarily utilise *in situ* switching mechanisms, as demonstrated by two separate studies (50 % and 59 %) (Liu *et al.*, 1985;Aitcheson *et al.*, 2005). As the infection progresses, and antibodies are generated against the BES VSG, recombination mechanisms are likely to play a role in activating the silent *VSGs*. Pleomorphic cell lines, however, appear to utilise *VSG* gene conversion events throughout an infection. It remains unclear exactly how frequently *in situ* switching mechanisms are used, but it appears that despite studies demonstrating an early transcriptional switching event, these generally occur infrequently (~ 9 %) (Robinson *et al.*, 1999;Morrison *et al.*, 2005).

A feature that is common to antigenic variation in a number of pathogens is the 'ordered expression' of variant antigens, meaning that specific surface molecules appear at somewhat predictable times in an infection (Barry, 1986;Capbern *et al.*, 1977). This ordered pattern of expression is based upon the probability of activation of each *VSG* and allows variants to arise gradually, rather than the co-expression of many variants which could overwhelm the host and therefore prove detrimental to the success of antigenic variation. In the hierarchy of activation in *T. brucei*, it appears that *VSGs* within telomeric locations are preferentially activated (Liu *et al.*, 1985;Morrison *et al.*, 2005), followed by intact *VSGs* located in arrays (Lee and Van der Ploeg, 1987;Timmers *et al.*, 1987) and finally by mosaic genes (Thon *et al.*, 1990). Recent research has also revealed that it is the activating *VSG* that determines which *VSG* is activated, not the previously active one (Morrison *et al.*, 2005).

Within mosaic genes, there also exists a suborder of variants. Early mosaic genes express similar *VSGs* and are thought to arise from recombination between close homologues. As the infection progresses, mosaic genes gradually diverge to distinct variants due to the assembly of genes encoding antigenically novel *VSGs* (Marcello and Barry, 2007b).

1.4 DNA double strand break repair

DNA double strand breaks (DSBs) arise frequently during DNA replication and can also be induced by ionising or UV radiation, by mutagenic chemicals and by free radicals (Kuzminov, 1995). Severe consequences can occur if DSBs remain unrepaired, including chromosomal fragmentation and translocation, which in multi-cellular organisms can lead to cancer and ultimately death (Khanna and Jackson, 2001).

The repair of these DSBs is mediated by two independent pathways; non-homologous end-joining (NHEJ) and homologous recombination (HR). These pathways differ in their requirement for a homologous DNA template and the fidelity of the repair reaction, but both appear to be conserved in eukaryotes. NHEJ occurs without a template and involves the re-ligation of the broken strands, which can lead to sequence changes. The primary proteins required for this process are Ku70/Ku80, the catalytic subunit of the DNA-dependent protein kinase and a DNA ligase heterodimer, composed of DNA ligase IV and XRCC4 (Chu, 1997). Homologous recombination, or the exchange of strands between homologous DNA molecules, not only repairs DNA damage but also ensures chromosome segregation and accurate genome duplication. The key protein that catalyses this reaction in bacteria is the RecA recombinase, which recognises homology between DNA molecules, pair homologous strands and mediates exchange (West *et al.*, 1981). In eukaryotes, the orthologue of RecA is named Rad51, while it is called RADA in archaea.

NHEJ appears to be the DNA repair mechanism that is favoured in mammalian cells whilst lower eukaryotes, such as budding yeast, appear to favour HR (Liang *et al.*, 1998). However, both mechanisms exist in most eukaryotes, and many prokaryotes, and it was once postulated that they competed against each other for the broken DNA ends at a DSB (Van Dyck *et al.*, 1999). It now appears that this is not the case and, in fact a number of determinants effect which pathway is chosen. One of these determinants is the cell cycle stage at which the DNA damage occurs. For instance, it has been established in chicken cells that if DNA damage occurs at G1-early S phase, the predominating pathway appears to be NHEJ, whilst at late S-G2 phase, HR appears to be favoured (Takata *et al.*, 1998). The position of a DSB along the chromosome is also considered to have a determining role on the repair mechanism, with NHEJ being favoured when the break it situated proximal to the telomere (Ricchetti *et al.*, 2003). Finally, the DNA substrate has also been shown to have an influence, with HR preferring to act on long ssDNA that is generated at DSBs (Ristic *et al.*, 2003).
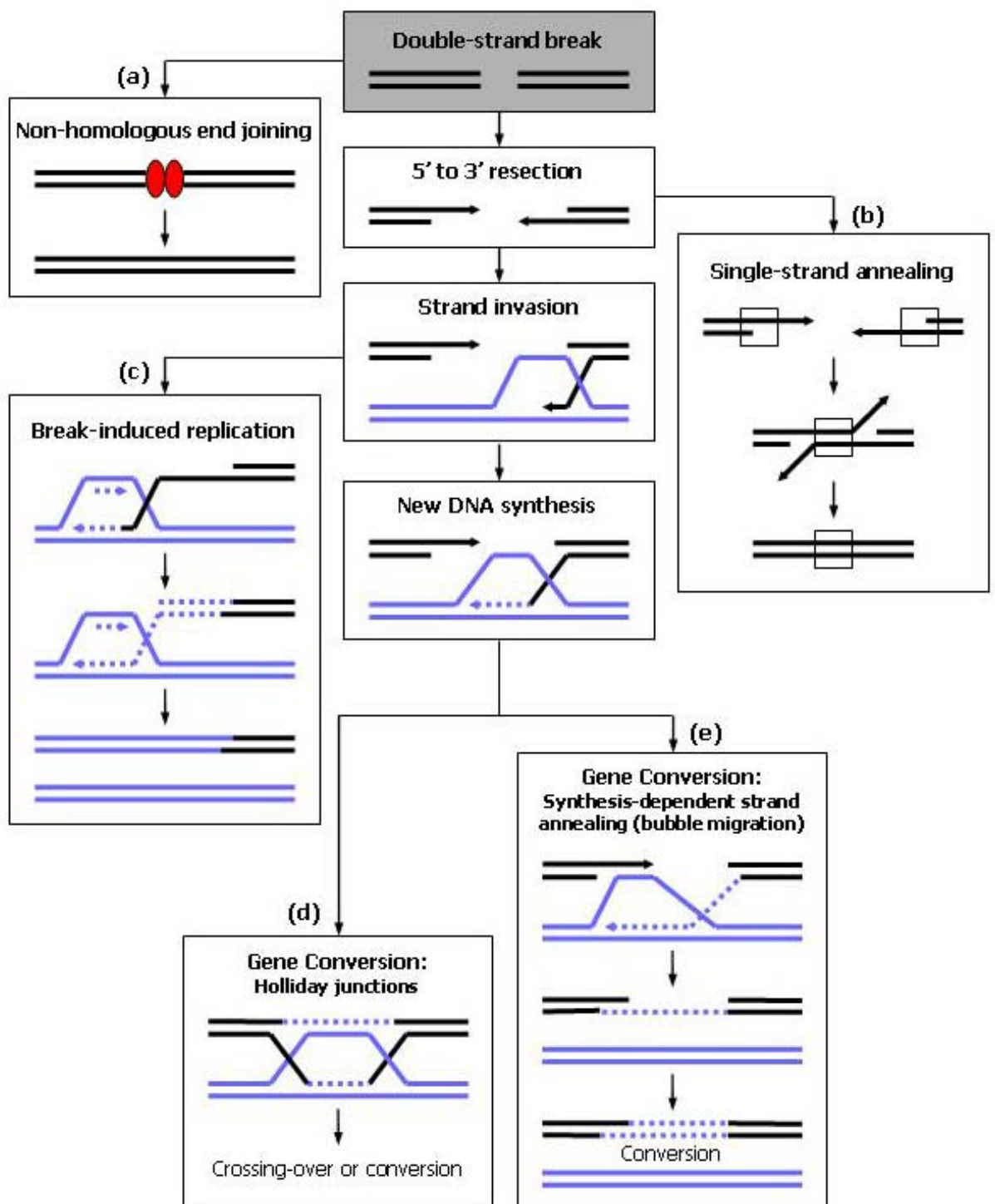
Figure 1.9 – Pathways of eukaryotic DNA double strand break repair. A schematic representation of DSB repair mechanisms in eukaryotic cells, including NHEJ and HR pathways. DNA containing a DSB is represented by black lines, intact duplex DNA by blue lines, newly synthesised DNA by dashed lines, and NHEJ machinery by red ovals. Taken from J.S. Bell, PhD thesis 2002.

Other additional roles exist for both NHEJ and HR, beyond the general repair of DSBs. In V(D)J recombination and meiosis, DSBs are caused deliberately in order to generate diversity (Xu *et al.*, 2005), and the repair of these DSBs is mediated by NHEJ and HR, respectively. HR is also involved in the repair of stalled replication forks (Michel *et al.*, 2004;Krogh and Symington, 2004) which can arise for many reasons, including DNA damage in the replication substrates and blockage of the replication machinery. Recombination proteins can target blocked replication forks and have been found to actually prevent the occurrence of DNA damage rather than to repair damage (Michel, 2000). HR proteins can also help to reverse stalled replication forks, allowing them to be reset without any DNA damage (Michel *et al.*, 2001;Seigneur *et al.*, 1998). In addition, HR has also been shown to rescue telomere length in yeast cells lacking telomerase (Lundblad and Blackburn, 1993;Le *et al.*, 1999).

1.4.1 Non-homologous end joining

Non-homologous end joining (NHEJ) involves the repair of a DSB by a re-ligation of the DNA ends (figure 1.9 – a). Very little (2-4 bp) or no sequence homology is required for this reaction to occur, though sequence changes often occur at the DSB site. The main components of the NHEJ machinery are the DNA dependent protein kinase catalytic subunit (DNA-PKcs), the Ku heterodimer (composed of Ku70 and Ku80), and the DNA ligase IV – XRCC4 complex (figure 1.10).

The MRX complex (composed of the proteins Mre11/Rad50/Xrs2 in yeast) also has a role in NHEJ, serving to act in the early stages by removing any proteins that may already be bound to the DNA ends and bridging the ends together (Stracker *et al.*, 2004;Connelly and Leach, 2002). This complex has been shown to provide an essential role in NHEJ in yeast, with deletion of each protein causing NHEJ to reduce by more than 70 fold (Moore and Haber, 1996). The severity of this may be due to the absence of DNA-PKcs in yeast.

Following on from MRX, the Ku heterodimer binds to the DNA ends and translocates along the DNA in an ATP dependent manner, serving to provide end protection (Mimori and Hardin, 1986) and stabilising the binding of the DNA-PKcs (Smith and Jackson, 1999). Ku acts as the DNA binding component of the DNA-dependent kinase multimer. However, quite how DNA-PK operates is currently unknown, though it is known to belong to the PIKK family, which includes ATM and ATR, and is known to phosphorylate p53, Ku, XRCC4 and itself (Smith and Jackson, 1999). It has been postulated that this phosphorylation is not essential, however, due to a DNA-PK being absent from yeast

(Critchlow and Jackson, 1998;Featherstone and Jackson, 1999). Finally, the DNA ligase IV – XRCC4 complex is recruited to the ends forming a tetrameric structure in which ligation of the ends occurs (Sibanda *et al.*, 2001).

NHEJ also appears to be conserved in at least some bacteria, where a homodimeric Ku homologue is found, plus NHEJ-specific ligases. Why a gene duplication has occurred to result in the 2 Ku copies found in eukaryotes is unknown (Della *et al.*, 2004;Wilson *et al.*, 2003).
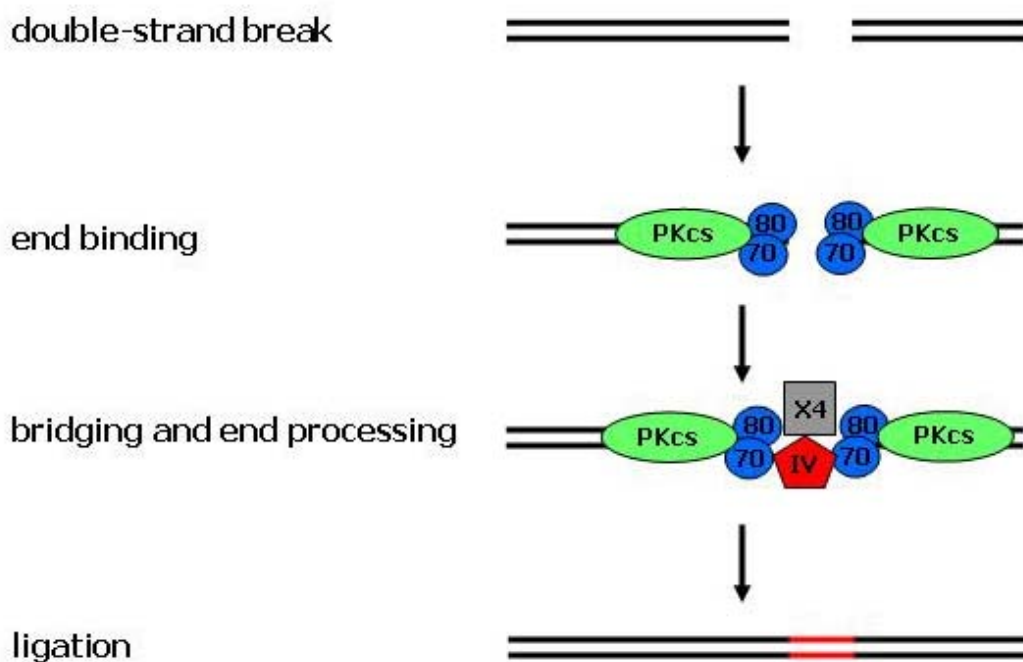
Figure 1.10 – Non-homologous end joining. A schematic representation of the process of non-homologous end joining (NHEJ). Following a DSB the Ku heterodimer binds to the DNA ends and recruits the catalytic subunit of the DNA protein kinase. The DNA ligase IV – XRCC4 complex is finally recruited to complete the ligation reaction. The Ku70/Ku80 heterodimer is represented by 2 blue circles whilst the DNA protein kinase catalytic subunit is depicted by a green oval. DNA ligase IV is represented by a red pentangle and XRCC4 by a grey square. PKcs - DNA protein kinase catalytic subunit; 70 – Ku70; 80 – Ku80; IV – DNA ligase IV; X4 – XRCC4.

1.4.2 Homologous recombination

Homologous recombination (HR) involves the accurate repair of a DSB utilising homologous sequence in an unbroken DNA molecule as a template. The process is conserved from bacteria to humans (Cromie *et al.*, 2001) and is more complex than NHEJ, involving a greater number of proteins (figure 1.11). In addition, HR can be divided into a number of different mechanisms, including single strand annealing, break induced replication and gene conversion events. Despite these processes appearing quite distinct

(figure 1.9, b – e), they all share the use of a homologous template. Moreover, these processes follow essentially the same catalytic steps of pre-synapsis, synapsis and post-synapsis (Hamatake *et al.*, 1989).

Pre-synapsis is the first stage and involves processing the ends of the DSB by resecting the 5' ends to provide 3' ssDNA substrates for HR. Following this is synapsis, in which the 3' overhangs invade homologous DNA during strand exchange. Finally, the reaction terminates with post-synapsis, which can be mediated through different mechanisms depending on whether both strands of the DSB have invaded duplex DNA. If both ends of the DSB invaded the duplex, a four stranded branched DNA structure known as a Holliday junction can form (Holliday, 1964), which requires specific enzymes to be resolved. Alternatively, the invading DNA strand can re-anneal with the broken DNA in a process known as synthesis dependent strand annealing (SDSA) (Nassif *et al.*, 1994). Both of these reactions can result in gene conversion events and are the main form of HR found in eukaryotes (Chen *et al.*, 2007). Other, perhaps more minor, HR reactions also occur, however. If only one end of a DSB invades the duplex DNA, a process known as break induced replication (BIR) occurs (Paques and Haber, 1999). Finally, single strand annealing (SSA) is a process whereby the broken DNA ends do not invade a DNA duplex at all, but rather homology is found in the flanking regions surrounding the DSB (Paques and Haber, 1999).

HR in eukaryotes utilises a large number of proteins known as the Rad52 epistasis group, which was originally identified in *S. cerevisiae*. The proteins included in this group are Rad50, Rad51, Rad52, Rad54, Rad55, Rad57, Rad59, Mre11 and Xrs2 (Symington, 2002). As already stated, Rad51 is the key member of this group and is the eukaryotic homologue of bacterial RecA. This protein forms a helical nucleoprotein complex on ssDNA, and this structure facilitates DNA strand exchange to occur when it interacts with homologous dsDNA. The other proteins appear to promote Rad51 activity, either operating upstream of the formation of the Rad51 filament, aiding formation of or stabilising the filament or acting in the strand exchange step.

1.4.2.1 Single strand annealing

Single strand annealing (figure 1.9 – b) occurs on repetitive DNA sequences and is a Rad51-independent pathway, which usually involves some loss of genetic material (Paques and Haber, 1999). Following a DSB, the 5' ends are resected, exposing complementary regions within the 3'strands flanking the break site. These complementary sequences are able to anneal to each other to repair the break, thereby eliminating the need for a strand

invasion step. Although this reaction occurs without the need for Rad51, Rad54, Rad55 or Rad57, it does require Rad52 and Rad59 (Ivanov *et al.*, 1996;Sugawara *et al.*, 2000). Following the annealing of sequences, the 3' non-homologous ends are excised and DNA synthesis and ligation completes the repair.

1.4.2.2 Break induced replication

Break induced replication (BIR) involves the invasion of only one end of the DSB (figure 1.9 – c) and can occur by both Rad51-dependent and Rad51-independent mechanisms (McEachern and Haber, 2006). BIR reactions often remain undetected in WT cells due to the high amount of gene conversion reactions (see below) that occur (Davis and Symington, 2004). Once the 5' end of a DSB is resected, the 3' ssDNA overhang invades a homologous chromosome. Following this invasion, a replication fork is established and the chromosome is copied for up to 100 kb, normally up to the chromosome end. Rad51-dependent BIR is much more efficient than Rad51-independent BIR and involves the same co-factors that are required for gene conversion reaction (see section 1.4.2.3) (Malkova *et al.*, 2005;Davis and Symington, 2004). Rad51-independent BIR requires shorter lengths of homologous substrate than Rad51 dependent BIR and involves the proteins Rad52, Rad50 and Rad59 (Ira and Haber, 2002;Bosco and Haber, 1998;Signon *et al.*, 2001).

1.4.2.3 Gene conversion

Gene conversion events (figure 1.9, d – e) allow the transfer of genetic material from one DNA molecule to its homologue in a uni-directional manner. This occurs most often between two alleles of a gene, but also occurs between homologous sequences on different chromosomes, and is the most common mechanism of HR in DSB repair (Chen *et al.*, 2007).

Following the introduction of a DSB, both 5' ends are resected in a 5' to 3' manner by exonucleases, leaving 3' ssDNA overhangs that can be thousands of bp long (figure 1.11) (White and Haber, 1990;Sun *et al.*, 1991). The MRX complex in yeast or its mammalian homologue, the MRN complex (Mre11/Rad51/Nbs1), was thought to be responsible for this resection (Trujillo *et al.*, 1998). However, the nuclease activity of Mre11 acts in a 3' to 5' polarity, leading to the suggestion that its role is not one of resection, but rather that of tethering and cleaning up the DNA ends through a conformational change (Krogh and Symington, 2004;de Jager *et al.*, 2001;Moreno-Herrero *et al.*, 2005). The resection of the DSB ends therefore appears to be created by other, perhaps redundant, nucleases.

In *S. cerevisiae,* the 3' ssDNA tails that result from DSB resection are subsequently coated with Replication Protein A (RPA) (figure 1.11), a homologue of the bacterial ssDNA binding protein, SSB (Sugiyama *et al.*, 1997). This protein serves to protect the DNA from nucleases and also removes any secondary structures (Sugiyama *et al.*, 1997).

The single stranded tails then become bound by Rad51 in eukaryotes (figure 1.11) (RecA in bacteria and RadA in archaea) (Brendel *et al.*, 1997;Seitz *et al.*, 1998), forming a characteristic nucleoprotein filament on the ssDNA that has a key role in recombination (Shinohara *et al.*, 1992). However, in order for Rad51 to be able to bind to ssDNA efficiently, the RPA protein needs to be removed (Sung, 1997a;Sung, 1997b). In both yeast and mammals Rad52 facilitates the removal of RPA (Sung, 1997a;Benson *et al.*, 1998), but Rad52 is notably absent from *D. melanogaster*, *C. elegans* and *T. brucei*. Evidence exists that the breast cancer susceptibility protein, Brca2 (see below) can act preferentially at the interface between dsDNA and ssDNA, resulting in the displacement of RPA from the overhang (Yang *et al.*, 2002;Martin *et al.*, 2005). This mechanism appears to assist the loading of Rad51 onto ssDNA and could provide an explanation of how this occurs in the absence of Rad52. Brca2 appears to be conserved in most eukaryotic organisms with the notable exception of *S. cerevisiae* (Kowalczykowski, 2002). Why mammals utilise both Rad52 and Brca2 is unclear.

Most eukaryotes also express multiple Rad51-related proteins (often called Rad51 paralogues) that also aid Rad51 function. Rad55 and Rad57, which form a heterodimer, are examples in *S. cerevisiae* (Sung, 1997b). Rad55-57 helps the formation of the Rad51 nucleoprotein filament in a mechanism that is distinct to that of Rad52 (Gasior *et al.*, 1998). Notably, an absence of either of these proteins can be compensated for *in vivo* by an over-expression of either Rad51 or Rad52 (Johnson and Symington, 1995;Hays *et al.*, 1995). The situation is more complicated in mammals due to the existence of five Rad51-related proteins (Rad51B, Rad51C, Rad51D, XRCC2 and XRCC3) in addition to Rad52. These paralogues have been shown to form two distinct protein complexes *in vivo*; Rad51B-Rad51C-Rad51D-XRCC2 (designated the BCDX2 complex) and Rad51C-XRCC3 (Masson *et al.*, 2001b;Liu *et al.*, 2004). Their precise roles remain unclear, though, it has been demonstrated that the BCDX2 complex can bind ssDNA, gaps in dsDNA and nicks in duplex DNA (Masson *et al.*, 2001b), whilst the Rad51C-XRCC3 complex has been shown to have a DNA binding activity that might be important in resolving Holliday junctions (Liu *et al.*, 2004). Each of the mammalian Rad51 paralogues appears to possess an important role, since their disruption causes lethality in mice and an impaired ability to undergo recombination and repair in hamster and human cell lines (Shu

et al., 1999;Deans *et al.*, 2000;Pittman and Schimenti, 2000;Johnson *et al.*, 1999;Pierce *et al.*, 1999;French *et al.*, 2002;Godthelp *et al.*, 2002). The repertoire of RAD51 paralogues in *T. brucei* appears to be greater than that of *S. cerevisiae* and more similar to that of mammals, with four Rad51 paralogues being uncovered recently (Proudfoot and McCulloch, 2005).

Once the mature Rad51 nucleoprotein filament has formed onto the ssDNA tails, the protein uses the sequence of the DNA to scan the genome for homologous sequences and then catalyses invasion of the DNA into the duplex in a process known as 'strand invasion' (figure 1.11). This leads to the formation of a displacement loop (D-loop), a bubble of unwound DNA in which the complementary strand has been displaced from the intact duplex. Strand invasion is aided by Rad54, a dsDNA-dependent ATPase which is also a member of the Swi2/Snf2 family of chromatin remodelling proteins (Lisby and Rothstein, 2004;Eisen *et al.*, 1995;Emery *et al.*, 1991). Rad54 interacts with Rad51 to aid chromatin remodelling, using the ATP hydrolysis function of Rad54 to supercoil and separate the strands of the homologous DNA (Alexiadis and Kadonaga, 2002;Van Komen *et al.*, 2000;Sigurdsson *et al.*, 2002).

Following the formation of the D-loop, leading strand and lagging strand DNA synthesis occurs, using the 3' end as a primer for new DNA synthesis and the donor strand as a template (Paques and Haber, 1999;Holmes and Haber, 1999). In a process known as second end capture, the second 3' end of broken DNA also invades the D-loop, followed subsequently by DNA synthesis and ligation of the nicks, leading to the formation of a structure with two Holliday Junctions (HJs). Cleavage of the two HJs then occurs, yielding either a non-crossover (gene conversion) or a crossover product. The HJ resolvase that catalyses this step in eukaryotes is as yet unknown, but it has been postulated that in human cells it could be the Rad51C-XRCC3 heterodimer, which is capable of resolving HJs *in vitro* (Liu *et al.*, 2004;Liu *et al.*, 2007). The Mus81-Eme1 endonuclease also appears to be a contender, due to its ability to act on recombination intermediates (Boddy *et al.*, 2001;Chen *et al.*, 2001).

The above model predicts that the resolution of HJs occurs in two orientations, leading to the generation of an equal amount of crossover and non-crossover outcomes. However, mitotic recombination events result in an extremely low occurrence of crossover events (<8%) (Esposito, 1978;Haber and Hearn, 1985;Kupiec and Petes, 1988). To account for this low occurrence of crossovers, the synthesis-dependent strand annealing (SDSA) model was proposed (Nassif *et al.*, 1994;Paques and Haber, 1999). In this model, following

strand invasion and D‑loop extension, the newly synthesized strand is displaced from the template and anneals to the other 3'ssDNA tail, allowing DNA synthesis to occur on the receiving strand, followed by ligation of nicks (Haber *et al.*, 2004;Ira *et al.*, 2006). This model generally yields only non-crossover products, but can allow for cross-over events (Ferguson and Holloman, 1996).
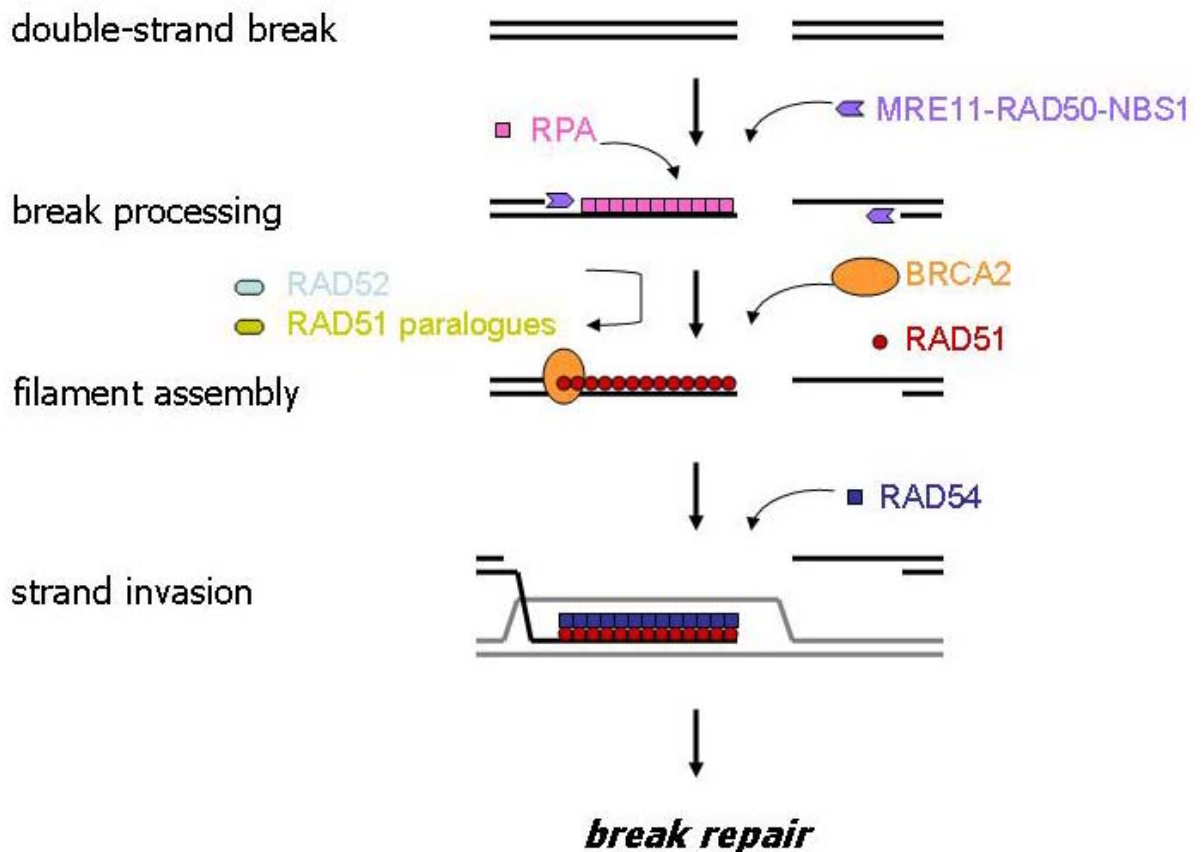
Figure 1.11 – Proteins involved in the early stages of eukaryotic homologous recombination. The black lines represent duplex DNA that has suffered a DSB and grey lines represent intact duplex used as a template for the repair of the damaged strand. The 5′ ends of the DSB are resected with the aid of the MRN complex (MRX in yeast) and other nucleases to form 3′ ssDNA tails. The ssDNA tails becomes coated with RPA to eliminate any secondary structure. The loading of the RAD51 nucleoprotein filament is aided by RAD52, the RAD51 paralogues and BRCA2, which also act to remove RPA. The tails then actively 'scan' the genome for homologous sequences in a 'strand invasion' process that is aided by RAD54. Following this one tail invades the homologous DNA duplex forming a displacement (D)-loop, which is then extended by DNA synthesis. See text for further details.

1.4.3 Mismatch repair

The mismatch repair (MMR) system has a vital role in maintaining genomic integrity and serves to recognise and repair any base mismatches that may arise during replication or due to mutagenesis by alkylating agents, such as MNNG (N-methyl-N'-nitro-N-nitrosoguanidine), or by cisplatin (Jiricny, 2006). Just as importantly, it also acts to prevent HR between non-identical sequences, thereby reducing the levels of HR and ensuring that exchanges only occur between homologous sequences (Datta *et al.*, 1996;Elliott and Jasin, 2001).

The basic processes of MMR appear to be conserved from bacteria to higher eukaryotes, involving a number of proteins. In bacteria, mismatched bases are recognised by the MutS protein (MSH2/MSH6 or MSH2/MSH3 heterodimers in eukaryotes) (Allen *et al.,* 1997). This is followed by the recruitment of MutL (MLH1/PMS1, MLH1/PMS2 and MLH1/MLH3 heterodimers in eukaryotes) (Galio *et al.*, 1999), which is required for the activation of MutH. MutH is an endonuclease, which nicks the DNA in preparation for the removal of the mis-paired bases (Au *et al.*, 1992), though no eukaryotic homologue has been found. In bacteria, excision of the mis-paired DNA strand requires the UvrD helicase (Dao and Modrich, 1998) and a ssDNA exonuclease before the DNA can be re-synthesised by DNA polymerase III and DNA ligase (Kunkel and Erie, 2005). Eukaryotes appear somewhat different, in that mismatch repair can be reconstituted *in vitro* without a helicase, and requires ExoI, RPA, DNA polymerase δ, DNA ligase I and the non-histone chromatin factor HMGB1 (Jiricny, 2006).
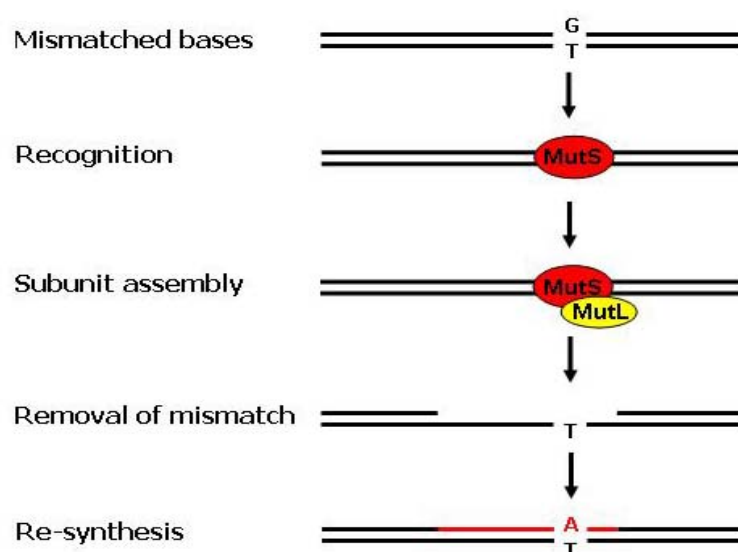
Figure 1.12 – Mismatch repair system in bacteria. The mismatch is recognized by MutS which, together with MutL, initiates MutH, an endonuclease which nicks the DNA. This is followed by the exonuclease degradation of DNA until the mismatched base is removed. This is subsequently filled in by DNA polymerase, which inserts the correct nucleotide. Figure adapted from Sancar, 1999.

1.5 BRCA2

Breast cancer is one of the most common causes of cancer related deaths in women and approximately 5-10 % of individuals who develop the disease are genetically predisposed to it (Lynch *et al.*, 1984). Linkage analysis identified the first breast cancer susceptibility gene, *BRCA1,* in 1991 (Hall *et al.*, 1990). A failure to assign all cases of breast cancer to a mutation in this gene led to the search for a second breast cancer susceptibility gene and in 1995, *BRCA2* was identified (Wooster *et al.*, 1995). Since then it has been established that women possessing an abnormal copy of *BRCA1* or *BRCA2* have up to an 85 % risk of developing breast cancer by the age of 70 (www.breastcancer.org). The products of these genes both participate in gene conversion events and therefore contribute a critical role in the maintenance of genome stability (Moynahan *et al.*, 1999;Moynahan *et al.*, 2001;Xia *et al.*, 2001). However, they are remarkably different proteins, most notably in terms of their size and the proteins with which they interact (see figure 1.13). For example, BRCA1 interacts with the MRN complex (Scully *et al.*, 1997), whilst BRCA2 interacts with Rad51 (Sharan *et al.*, 1997;Chen *et al.*, 1998b;Marmorstein *et al.*, 1998) and BRCA1 (Chen *et al.*, 1998a). Since the work in this thesis concentrates primarily on BRCA2, the following sections will focus solely on this protein.
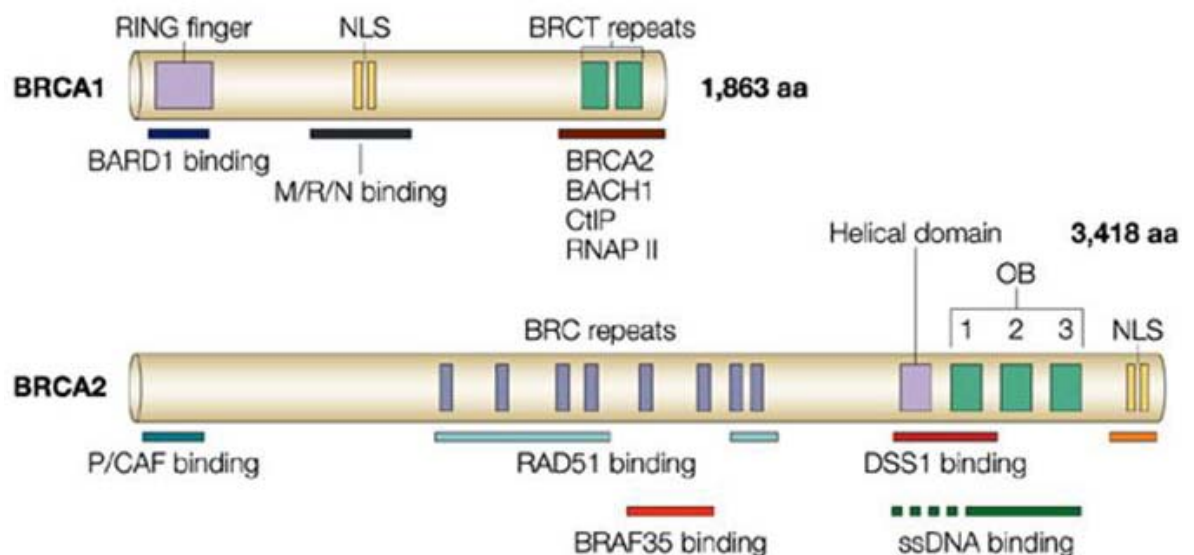
Figure 1.13 – The BRCA1 and BRCA2 proteins displaying the functional domains and interacting proteins. Both BRCA1 and BRCA2 are large polypeptides (1863 and 3418 amino acids respectively) which interact with each other and several other proteins. BRCA2 interacts with the histone acetylase P/CAF, BRAF35, RAD51 and DSS1. The sites of the eight BRC repeats (six of which interact with RAD51), the oligonucleotide binding (OB) domains and the nuclear localisation signal (NLS) sequences are indicated. Figure taken from West, 2003.

1.5.1 The structure of BRCA2

The human BRCA2 protein is a large polypeptide consisting of 3418 amino acids. Following a wealth of research into the protein in the last 13 years, BRCA2 is now known to directly regulate the recombinase Rad51 and form an essential part of the HR pathway (Jasin, 2002). Not only this, but it has also been shown to interact with several other proteins involved in the process of DNA repair and clues to the exact function of the protein are now being uncovered.

1.5.1.1 Rad51 binding occurs at the BRC repeats

Initial inspections of the gene revealed no obvious similarities to any other genes within available published sequences (Wooster *et al.*, 1995). However, the BRC repeat domain was identified and located to exon 11 following the failure to detect any similarity between BRCA2 and BRCA1 (Bork *et al.*, 1996). The BRC repeat domain contains a series of eight degenerate motifs which are approximately 30 amino acids long and are interspersed along a 1200 amino acid central region (see figure 1.13).

A major breakthrough in understanding the function of BRCA2 came when it was discovered that BRCA2 interacted with Rad51. This was evidenced by yeast two hybrid analyses and co-immunoprecipitation (Sharan *et al.*, 1997;Chen *et al.*, 1998b;Marmorstein *et al.*, 1998). In addition to this, it was observed that BRCA2 and Rad51 co-localise to DNA damaged induced foci, and that Rad51 foci fail to form in the absence of BRCA2 (Yuan *et al.*, 1999;Tarsounas *et al.*, 2003). Further research identified the region of this interaction as the BRC repeat domain (Wong *et al.*, 1997;Chen *et al.*, 1998b). A notable degree of sequence divergence exists across the 8 BRC repeats within BRCA2 in *H. sapiens*, with BRC1, BRC3, BRC4, BRC7 and BRC8 exhibiting the highest levels of similarity. It might be considered that this level of diversity across the BRC repeats could represent an example of evolutionary tuning, with the aim of producing binding sites with a range of affinities for Rad51. Indeed, this theory is supported by the demonstration that although all BRC repeats have the ability to bind Rad51 *in vitro*, some bind with a stronger affinity than others (Wong *et al.*, 1997;Chen *et al.*, 1998b): BRC3 and BRC4 have been shown to display a very strong interaction with Rad51, whilst BRC5 and BRC6 display a very weak interaction (Wong *et al.*, 1997;Chen *et al.*, 1998b), and are the most diverged repeats. This has led to the current thinking, that *in vivo*, only 6 out of the 8 BRC repeats in human BRCA2 bind Rad51. It has since been postulated that BRC5 and BRC6 could represent binding sites for other recombination factors (Pellegrini and Venkitaraman, 2004).

The large number of BRC repeats present within BRCA2 was originally considered to be necessary in order to deliver a sufficient number of Rad51 molecules onto the ssDNA at the site of the DSB, thereby allowing the nucleoprotein filament to be stabilised (Pellegrini and Venkitaraman, 2004). It was postulated that the BRC repeats with the weakest affinity would be the first to release Rad51 onto the damaged DNA, whilst those with the strongest affinity would be the last to release Rad51 (Pellegrini and Venkitaraman, 2004). However, recent investigations have shown that functional BRCA2 orthologues exist in many other eukaryotes, including the smut fungus *Ustilago maydis* and the nematode *Caenorhabditis elegans*, both of which have been shown to possess only 1 BRC repeat (Kojic *et al.*, 2002;Martin *et al.*, 2005). This therefore leaves the possibility that the additional repeats in higher eukaryotes could fulfil a different role or roles, or it could simply be indicative of the more complex biological systems.

Experiments with synthetic peptides corresponding to the BRC repeats have demonstrated that the interaction between Rad51 and BRC3 or BRC4 actually inhibited the DNA binding properties of Rad51. Not only was the formation of the Rad51 nucleoprotein filament prevented, but it was actually disrupted when the BRC peptides were present in molar excess of Rad51 (Davies *et al.*, 2001;Tarsounas *et al.*, 2004;Davies and Pellegrini, 2007;Esashi *et al.*, 2007). In light of this research, it appeared that the role of BRCA2 might be to provide a negative control mechanism over Rad51, possibly by keeping Rad51 inactive until it is needed for repair (Tarsounas *et al.*, 2004). Indeed, this hypothesis appeared to be supported by the *in vivo* overexpression of the BRC4 motif reducing the ability of cells to form Rad51 foci at sites of DNA damage (Chen *et al.*, 1999a).

Further insight into the mechanism of the interaction between the BRC repeats and Rad51 was provided when the structure of BRC4 bound to the core of Rad51 was solved by X-ray crystallography (Pellegrini *et al.*, 2002). This study revealed that the residues F-TASGK, which are conserved within the BRC repeats, are critical in mediating hydrophobic interactions with Rad51. Further evidence highlighting the critical nature of this motif arose from cancer-predisposing mutations being found within these residues (http://research.nhgri.nih.gov/bic/) (Bork *et al.*, 1996;Bignell *et al.*, 1997;Davies *et al.*, 2001). It could therefore be postulated that the differences in affinity for Rad51 arise not due to differences in the conserved residues but due to divergence in the flanking non-conserved residues (Pellegrini *et al.*, 2002). The crystal structure of BRC4 bound to Rad51 also provided evidence for the mechanism of the interaction. Through the comparison of this structure to the crystallographic RecA filament, it was revealed that BRCA2 appears to interact with Rad51 by mimicking the structure of the interaction domain between adjacent

Rad51 monomers. It also became apparent that the interaction between a BRC repeat and a Rad51 monomer prevents the ability of Rad51 monomers to interact with each other and thereby blocking filament formation (Pellegrini *et al.*, 2002). This would therefore allow the BRC repeats to maintain the Rad51 molecules in a monomeric state, which would only be able to form a nucleoprotein filament once deposited onto ssDNA. Indeed, this hypothesis has recently been shown to be accurate with the demonstration that only Rad51 monomers can bind to the BRC repeats and not Rad51 filaments (Davies and Pellegrini, 2007;Esashi *et al.*, 2007).

A feature common to all orthologues of BRCA2 that have been identified to date, is the presence of at least one BRC repeat. This observation, along with the conservation of the critical residues in the BRC repeat sequence between several different species appears to suggest that these domains are essential to the function of BRCA2 (Bignell *et al.*, 1997). Indeed, mutations located within the BRC repeats have been shown to associate with familial ovarian cancer (Gayther *et al.*, 1997), providing further evidence of their importance to the functioning of BRCA2. Generally, it appears that simpler organisms possess a smaller number of BRC repeats, whilst more complex, multicellualr organisms such as higher eukaryotes possess a larger number of repeats (Lo *et al.*, 2003). However, exceptions to this rule do exist, most notably in the *Trypanosoma brucei* homologue which is predicted to contain 15 BRC repeats, 14 of which are identical and are separated by exactly 20 amino acids (see chapter 3 for further details), and in *C. elegans* BRCA2, which has only one BRC repeat (Martin *et al.*, 2005).

1.5.1.2 Rad51 also binds to the C-terminus of BRCA2

In addition to Rad51 binding to the BRC repeats, it has been uncovered that there is also an unrelated Rad51 binding site located at the carboxyl terminus, situated within exon 27 of human BRCA2 (Mizuta *et al.*, 1997;Sharan *et al.*, 1997). This has been named the TR2 region by Esashi *et al.*, (2005) and has not only been demonstrated to bind Rad51, but has also been found to be phosphorylated in HeLa extracts (Esashi *et al.*, 2005) on serine 3291 (S3291), a reaction that is mediated by cyclin-dependent kinases (CDK's). This phosphorylation has been shown to have a direct effect on Rad51 binding, with interactions only occurring when S3291 is de-phosphorylated. The S3291 phosphorylation status fluctuates throughout the cell cycle, with low levels of phosphorylation observed in S-phase and high levels as cells enter mitosis. In addition to this, the induction of DNA damage has been shown to cause rapid de-phosphorylation of this site, thereby allowing Rad51 to bind.

These results appear to indicate that the phosphorylation status of S3291 may provide a 'molecular switch', through which HR can be down-regulated as cells approach mitosis and up-regulated in response to DNA damage. Indeed, this hypothesis is supported by an increase in tumour susceptibility when either exon 27 is deleted or mutations are located within the Cdk target site (McAllister *et al.*, 2002;Donoho *et al.*, 2003).

More recently, it has been uncovered that the TR2 region is only capable of binding multimeric forms of Rad51, such as filaments or rings, and cannot support the binding of Rad51 monomers (Esashi *et al.*, 2007;Davies and Pellegrini, 2007). In addition, it has been demonstrated that the TR2 region can also act to protect Rad51 filaments from the disruption caused by the BRC repeats. The importance of an alternative Rad51 binding site and mode is underlined by the apparent conservation of this activity in other orthologues of BRCA2. For example, the BRCA2 homologue in *C. elegans* (CeBRC-2) has been revealed to bind Rad51 in a similar manner to the TR2 region through a site located at the N-terminal domain (Petalcorin *et al.*, 2007). Furthermore, the *Ustilago maydis* homologue (Brh2) has also been demonstrated to contain a similar site (CRE), located at the C terminal domain (Zhou *et al.*, 2007).

Taken together, these findings have contributed to the current thinking of how BRCA2 operates within the HR pathway in that the BRC repeats and the TR2 region provide functionally distinct mechanisms (see section 1.5.1.4).

1.5.1.3 DNA binding domains

Apart from the Rad51 binding domains, the BRCA2 protein contains several other functional motifs lying downstream of the BRC repeats (figure 1.13). It was discovered in 1999 that the final third portion of the human BRCA2 protein associates with a highly acidic 70 amino acid polypeptide, DSS1 (Marston *et al.*, 1999), a protein which is mutated in split hand/split foot syndrome (Crackower *et al.*, 1996). DSS1 has been shown to provide a critical role for DNA damage induced Rad51 focus formation and for the maintenance of genomic stability (Gudmundsdottir *et al.*, 2004). Furthermore, it has been speculated that DSS1 is required for the BRCA2-Rad51 complex to associate with the sites of DNA damage, possibly due to its acidic nature mimicking ssDNA and thereby facilitating the accessibility of BRCA2 onto damaged DNA sites (Gudmundsdottir *et al.*, 2004). Co-expression of DSS1 along with the C terminal region of *H. sapiens* BRCA2 has allowed the X-ray crystallographic structure of this part of the breast cancer protein to be determined (Yang *et al.*, 2002). Structures were determined both with and without ssDNA, revealing five distinct domains in the C terminal region of BRCA2 (figure 1.13).

The first domain is described as the helical domain due to its 190 amino acids consisting mainly of α-helices. Lying further downstream are three oligonucleotide/oligosaccharide binding folds (OB1, OB2 and OB3), which are structurally similar to the OB folds present in the prokaryotic ssDNA-binding-protein (SSB) and the eukaryotic replication-protein-A (RPA), both of which also bind ssDNA. Indeed, the entire C terminal region of BRCA2 was shown to bind to ssDNA with high affinity. Located within OB2 exists a 130 amino acid insertion, which has been named the tower domain due to its tower-like structure protruding from the OB fold. DSS1 was found to associate with residues spanning the α-helical domain, OB1, the tower domain and OB2, whilst the tower domain is implicated in providing an additional role in dsDNA binding (Yang *et al.*, 2002).

In light of this structure, it was proposed that BRCA2 could be responsible for targeting Rad51 to the ssDNA/dsDNA junction at the sites of processed DSBs. Indeed, the affinity of the C terminal domain of BRCA2 for binding to ssDNA/dsDNA junctions could serve to displace the RPA molecules and thereby allow the formation of the Rad51 nucleoprotein filament (Martin *et al.*, 2005;Powell *et al.*, 2002;Wilson and Elledge, 2002).

1.5.1.4 The function of BRCA2 in HR

Despite the purification of the entire BRCA2 protein from any organism, so far proving unsuccessful, a number of experiments with smaller subunits of the protein have yielded a wealth of information. Not only this, but extensive research into organisms with BRCA2 deletions or tumour cells with mutations of BRCA2 have provided a clearer view of how the protein functions within the HR pathway.

BRCA2 has been demonstrated to be essential for genomic stability in eukaryotes, with *BRCA2* deficient murine and human cells both exhibiting an accumulation of chromosome breaks and radial chromosomes (Sharan *et al.*, 1997;Patel *et al.*, 1998;Yu *et al.*, 2000;Moynahan *et al.*, 2001;Tutt *et al.*, 2001). These phenotypes were presumed to result from the failed DNA repair of DSBs, since BRCA2 has been shown to co-localise with Rad51 in nuclear foci following damage (Tarsounas *et al.*, 2004). Indeed, this hypothesis was supported by the lack of Rad51 foci in the *BRCA2* deficient pancreatic cancer cell line, CAPAN-1 (Yuan *et al.*, 1999). This cell line possesses only one *BRCA2* allele, which encodes a truncated protein consisting of only 6 BRC repeats and no DNA binding domain. In addition to its impaired ability to form Rad51 foci, this cell line also displays hypersensitivity to DNA damaging agents (Goggins *et al.*, 1996).

Similar phenotypes have also been exhibited in lower eukaryotes deficient in *BRCA2*. For example, *Ustilago maydis* cells deficient in *Brh2* display defects in DNA repair, recombination and meiosis (Kojic *et al.*, 2002). Similarly, the absence of *Cebrc-2* in *C. elegans* caused defective DSB repair, induced by both ionising radiation and meiosis. These phenotypes were shown to result from an in ability to target Rad51 to the sites of the DSBs (Martin *et al.*, 2005). Taken together, these findings confirm a role for BRCA2 in DSB repair by directly regulating the recombinase Rad51 in the HR pathway.

It has been known for some time that the BRC repeats were responsible for interacting with Rad51(Wong *et al.*, 1997;Chen *et al.*, 1998b) and originally it was presumed that these allowed Rad51 to be recruited to the sites of damaged DNA. However, this model has since been forced to be updated, not only due to evidence implicating the BRC repeats in a negative role over Rad51 (Davies *et al.*, 2001;Chen *et al.*, 1999a;Tarsounas *et al.*, 2004;Davies and Pellegrini, 2007;Esashi *et al.*, 2007), but also due to the discovery of an additional, unrelated Rad51 binding domain (TR2) located at the C terminus of human BRCA2 (Esashi *et al.*, 2005). Indeed, with the recent discovery that the BRC repeats only support the binding of Rad51 monomers (see above), and the TR2 region supports the binding of Rad51 filaments (Esashi *et al.*, 2007;Davies and Pellegrini, 2007), a new model has been proposed (Lord and Ashworth, 2007;Petalcorin *et al.*, 2007) (figure 1.14). This model proposes that in a normal cell, BRCA2 holds Rad51 in an inactive monomeric form at the BRC repeats. However, when DNA damage is induced, not only does BRCA2 localise to the sites of damage, but de-phosphorylation of S3291 also occurs. This modifying event activates the TR2 region, which supports the formation of the Rad51 nucleoprotein filament. This, in turn, allows HR to progress and the DSB to be repaired. Finally, HR is halted when a cyclin dependent kinase phosphorylates S3291, causing the inactivation of TR2, which thereby allows the BRC repeats to disrupt the Rad51 nucleoprotein filament.
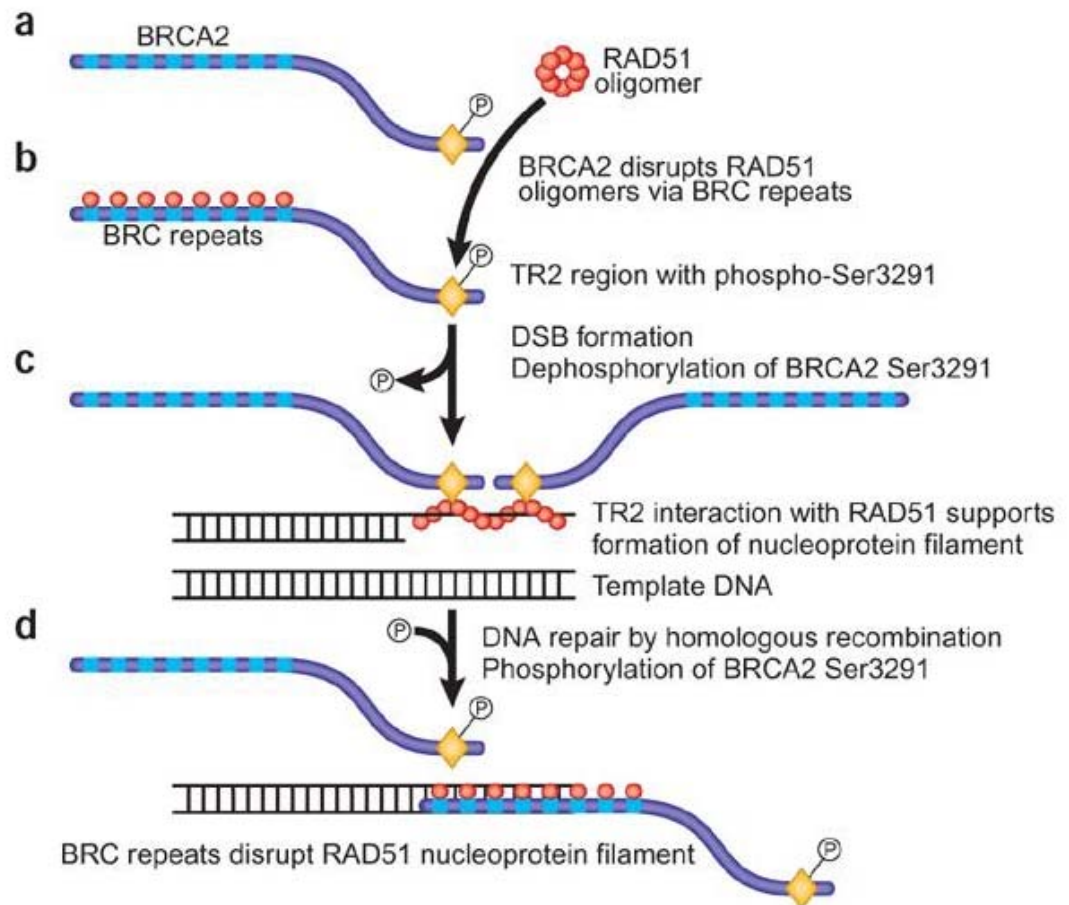
Figure 1.14 –BRCA2's role in the HR pathway and its interaction with Rad51. (a) Without DNA damage Rad51 oligomers are disrupted by the BRC repeats of BRCA2 (b) and remain bound in a monomeric form. (c) When DNA damage is detected, the site S3291 becomes dephosphorylated, activating the TR2 region. The TR2 region supports the binding of the Rad51 nucleoprotein filament, allowing the progression of homologous recombination. (d) When DSBs are repaired the site S3291 becomes phosphorylated by CDKs, inactivating the TR2 regions and allowing the BRC repeats to disrupt the nucleoprotein filament. Figure taken from Lord and Ashworth, 2007.

1.5.1.5 BRCA2 interacting proteins

Apart from Rad51 and DSS1, which have already been mentioned, BRCA2 is also known to associate with a number of other proteins (figure 1.13). In 1998, it was discovered that BRCA2 and BRCA1 co-exist in a complex and co-localise to sub-nuclear foci in somatic cells (Chen *et al.*, 1998a). In the same year, it also became apparent that BRCA2 interacts with p53 (Marmorstein *et al.*, 1998) and a transcriptional co-activator protein, P/CAF, which possesses histone acetyltransferase activity. The interaction with P/CAF was demonstrated both *in vitro* and *in vivo* and was found to be mediated by the *H. sapiens* BRCA2 residues 290-453 (Fuks *et al.*, 1998).

Three years later, Marmorstein *et al.*, (2001) identified a 2 MDa BRCA2-containing complex within which they were able to identify a structural DNA binding component, which they named BRCA2-Associated factor 35 (BRAF35) (Marmorstein *et al.*, 2001). In 2003, the N-terminal region of BRCA2 was discovered to co-immunoprecipitate with RPA, both *in vitro* and *in vivo* (Wong *et al.*, 2003). Not only this, but exon 3 of BRCA2 was found to associate with EMSY, a protein which is amplified in breast and ovarian cancer (Hughes-Davies *et al.*, 2003).

BRCA2 has also been discovered to interact with the mitotic Polo-like kinase (Plk1) in cell extracts (Lin *et al.*, 2003). This interacting domain has been localised to the region spanning the BRC repeats but, perhaps more interestingly, it has been discovered that the regions within the BRC repeats and not the BRC repeats themselves are phosphorylated by Plk1 as cells approach mitosis (Lee *et al.*, 2004). This phosphorylation event has also been discovered to cause the dissociation of P/CAF from BRCA2.

BRCA2 has also been found to function within the Fanconi Anemia (FA) pathway (see section 1.5.1.6). This finding was confirmed through co-immunoprecipitation and yeast-two-hybrid studies which revealed that BRCA2 interacts with FANCD2, a component of the FA pathway (Hussain *et al.*, 2004) and FANCG (Hussain *et al.*, 2003). Furthermore, BRCA2 function has also been demonstrated to require another binding partner, PALB2 (Xia *et al.*, 2006a), which has been recently identified as FANCN, another component of the FA pathway (Xia *et al.*, 2007;Reid *et al.*, 2007).

An interaction between BRCA2 and the meiosis-specific homologue of RAD51, DMC1, in *A. thaliana* has previously been described (Siaud *et al.*, 2004) and mapped to BRC2 (Dray *et al.*, 2006). More recently, this interaction has also been confirmed between the human proteins, highlighting the importance of BRCA2 in meiosis (Thorslund *et al.*, 2007). Through a series of yeast two-hybrid analysis, protein interaction assays and peptide arrays, this interaction has been mapped to residues 2386-2411 in human BRCA2 and has been named the PhePP motif (Thorslund and West, 2007). Unlike *A. thaliana*, this motif is unrelated to the BRC repeats, and only supports interactions between BRCA2 and DMC1, not Rad51. The PhePP motif appears to be highly conserved throughout vertebrates, but appears to have diverged in *U. maydis* and *C. elegans*. Recently, it has been discovered that the PhePP motif in *C. elegans* interacts with Rad51 (Petalcorin *et al.*, 2007). This fundamental difference between these eukaryotes could be explained by the lack of DMC1 in *C. elegans*. Instead, *C. elegans* expresses Rad51 at a high level in order to promote meiosis (Takanami *et al.*, 2000). A potential further difference between vertebrates and *C.*

elegans, is shown by the fact that the TR2 region of BRCA2 was shown to be dispensable for meiosis in vertebrates (Thorslund and West, 2007). Taken together, it is possible that these variations suggest considerable flexibility in the functions adopted by BRCA2 beyond RAD51 interaction.

1.5.1.6 BRCA2 is also a member of the Fanconi Anaemia pathway

Fanconi anaemia (FA) is an autosomal recessive disorder characterised by bone marrow failure, compromised genome stability, and a predisposition to cancer (D'Andrea and Grompe, 2003). The FA pathway has been implicated to function in DNA repair mediated through the HR pathway, and in the protection of stalled replication forks (Niedzwiedz *et al.*, 2004;Taniguchi and D'Andrea, 2006;Takata *et al.*, 2006;Yang *et al.*, 2005). The role in DNA repair was initially hypothesised due to observations that cells derived from FA patients exhibited sensitivities to cross linking agents (Grompe and D'Andrea, 2001;Joenje and Patel, 2001). There are currently 13 proteins proposed to function in the FA pathway, which include FANCA, FANCB, FANCC, FANCD1, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCJ, FANCL, FANCM and FANCN, but their precise roles still remain unclear (Joenje and Patel, 2001;Smogorzewska *et al.*, 2007;Levitus *et al.*, 2006;Meetei *et al.*, 2005).

The proteins of the FA pathway have been discovered to function closely with the breast cancer susceptibility proteins, BRCA1 and BRCA2. For example, BRCA1 is known to be required for the efficient foci formation of FANCD2 (Vandenberg *et al.*, 2003;Garcia-Higuera *et al.*, 2001), whilst cell lines defective in FANCD1 have been shown to carry bi-allelic mutations in BRCA2 (Howlett *et al.*, 2002). This latter observation, along with the discovery that BRCA2 can complement for the defect in FANCD1 cell lines, led to the conclusion that BRCA2 and FANCD1 are in fact one in the same (Howlett *et al.*, 2002). Further evidence displaying the relationship between the BRCA proteins and HR comes from the discovery that BRCA2 interacts with FANCD2 and FANCN (Hussain *et al.*, 2004;Reid *et al.*, 2007;Xia *et al.*, 2007). More recently it has been implicated that BRCA2 also forms a complex with FANCD2, FANCG and XRCC3 (Wilson *et al.*, 2008).

Of the remaining proteins, it is known that FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL and FANCM interact to form a multi-subunit nuclear complex, along with the FANCA associated polypeptides (FAAP), FAAP24 and FAAP100 (Meetei *et al.*, 2005;Garcia-Higuera *et al.*, 2001;Ciccia *et al.*, 2007;Ling *et al.*, 2007;Mathew, 2006). This complex is known to have multiple roles involving the activation, re-localisation, and monoubiquitylation of FANCD2 and FANCI (D'Andrea and Grompe, 2003;Takata *et al.*,

2006;Smogorzewska *et al.*, 2007). An example of this is found following exposure to DNA damage, where the FA core complex monoubiquitylates FANCD2, causing it to re-locate to sub-nuclear foci with BRCA1 and Rad51 (Taniguchi and D'Andrea, 2006;Garcia-Higuera *et al.*, 2001). Here, FANCD2 is thought to promote the loading of BRCA2 (FANCD1) into chromatin complexes, facilitates the assembly of Rad51 foci and thereby promoting HR (Wang *et al.*, 2004;Hussain *et al.*, 2004).

1.6 Rad51

Rad51, the eukaryotic homologue of bacterial RecA, is a relatively small protein (38kDa) that possesses a core domain structure containing Walker A and B motifs, which are responsible for ATP binding and hydrolysis (Walker *et al.*, 1982). This domain is homologous to the catalytic domain of *E. coli* RecA and also exhibits a high level of similarity to the Walker motifs present in DMC1 (Bishop *et al.*, 1992;Sauvageau *et al.*, 2004). Recombinant forms of Rad51 have been shown to form oligomeric ring structures of 7 or 8 monomers in the absence of DNA (Shin *et al.*, 2003;Kinebuchi *et al.*, 2004). However, when associated with DNA, Rad51 becomes functionally active and forms a highly ordered right-handed helical nucleoprotein filament which possesses homologous pairing and strand exchange activities (Benson *et al.*, 1994;Conway *et al.*, 2004;Sehorn *et al.*, 2004). Although recombinant forms of Rad51 and RecA can perform these functions alone *in vitro*, *in vivo* many other factors are required (Sung *et al.*, 2003;Symington, 2002) (see section 1.4.2.3).

When bacteria are exposed to DNA damage, RecA induction is observed to increase more than 15 fold in a 'SOS response' (Little and Mount, 1982;Walker, 1984). Mammalian cells do not exhibit such an induction (Tarsounas *et al.*, 2004), which may be surprising since the activity of human Rad51 in strand exchange assays has been shown to be comparatively much lower than that of RecA (Baumann *et al.*, 1996). Indeed, other eukaryotes, such as *S. cerevisiae* and *T. brucei's* Kinetoplastid relatives, *Leishmania major* and *T. cruzi*, do up-regulate Rad51 levels after damage (Shinohara *et al.*, 1992; McKean *et al.*, 2001; Regis-da-Silva *et al.*, 2006). Beyond up-regulation of expression, another response to DNA damage appears to be ubiquitous. When DNA damage is detected, Rad51 and other repair proteins that are normally diffused throughout the nucleus of eukaryotes (Haaf *et al.*, 1995;Scully *et al.*, 1997) are rapidly relocated and concentrated into sub-nuclear complexes that are microscopically detected as foci. This creates an overall effect that increases the local concentration of repair enzymes as the cell prepares for repair (Tarsounas *et al.*, 2004). In addition to this re-localisation occurring with the

onset of DNA damage, Rad51 foci formation is also observed to occur in undamaged S-phase cells, where they are proposed to repair broken replication forks (Tashiro *et al.*, 1996). Similar focal localisations of bacterial RecA have recently been described (Renzette *et al.*, 2005).

Of the five Rad51 paralogues which exist in vertebrates, Rad51 has been shown to interact with just Rad51C and XRCC3 through yeast two hybrid analysis (Masson *et al.*, 2001b;Masson *et al.*, 2001a;Schild *et al.*, 2000). In addition to this however, several other proteins have been shown to associate with Rad51, including Rad52 (Kurumizaka *et al.*, 1999;McIlwraith *et al.*, 2000), Rad54 (Mazin *et al.*, 2003), RPA (McIlwraith *et al.*, 2000), BRCA1 (Scully *et al.*, 1997), BRCA2 (Wong *et al.*, 1997;Chen *et al.*, 1998b), c-Abl (Chen *et al.*, 1999b), p53 (Sturzbecher *et al.*, 1996;Buchhop *et al.*, 1997;Linke *et al.*, 2003), UBL1 (Li *et al.*, 2000), Pir51 (Kovalenko *et al.*, 1997) and UBE2I (Shen *et al.*, 1996). Although BRCA1 and BRCA2 have been shown to co-localise with Rad51 in damaged induced foci (Scully *et al.*, 1997;Chen *et al.*, 1998b), it has since become clear that it is BRCA2 and Rad51 which interact (see section 1.5.1.1), and it is through the interaction of BRCA1 and BRCA2 that allows BRCA1 and Rad51 to associate (Sharan *et al.*, 1997;Wong *et al.*, 1997). It is important to note that damage induced Rad51 foci fail to form in BRCA2 deficient cells (Yu *et al.*, 2000), but S-phase foci remain unaffected, indicating that these foci must be distinct from each other (Tarsounas *et al.*, 2003). In addition, other major HR proteins, including Rad52 and Rad54, also co-localise to DNA damage induced Rad51 foci (Haaf *et al.*, 1995;Tan *et al.*, 1999;Liu and Maizels, 2000;Essers *et al.*, 2002b;Tarsounas *et al.*, 2004). Indeed, the damaged induced Rad51 foci are predicted to be sites of DNA damage, not only due to their formation shortly after DNA damage induction, but also due to the presence of ssDNA within a focus (Raderschall *et al.*, 1999). Despite this, the exact size, composition and number of breaks per foci remains unclear (van Gent *et al.*, 2001;West, 2003;Rouse and Jackson, 2002). However, it is currently presumed that a single focus represents more than one DSB (West, 2003;Essers *et al.*, 2002b) and potentially these represent 'repair centres' containing multiple catalytic or regulating factors.

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

In order to investigate the process of antigenic variation in *T. brucei*, a number of proteins have been identified and studied, which in other eukaryotes function in DNA repair and recombination pathways. This line of research began in 1999 with the generation of *rad51-/-* mutants in *T. brucei* (McCulloch and Barry, 1999). These parasite mutants displayed an impaired growth phenotype, sensitivity to a DNA damaging agent, an impaired ability to perform homologous recombination and, more importantly, a defect in the ability to switch VSG coat. Although VSG switching and DNA recombination were both reduced, neither process was completely abolished, with DNA recombination being found to occur using short lengths of sequence homology (Conway *et al.*, 2002c). These results therefore suggest the presence of one or more pathways that can compensate for the absence of RAD51.

A number of other DNA repair proteins have subsequently been investigated, and have produced some interesting and sometimes surprising results. Despite the demonstration of the presence of Ku70 and Ku80 in the *T. brucei* genome, and their function in telomere maintenance, no evidence has yet been provided that NHEJ actually functions *in vivo* (Conway *et al.*, 2002b). The potential absence of NHEJ in *T. brucei* has recently been supported by bioinformatics analyses, revealing that homologues of neither DNA ligase IV or XRCC4 could be detected (Burton *et al.*, 2007). Instead, it appears that microhomology-based repair occurs, indicated by the fact that *T. brucei* cell extracts can support the end-joining of linear DNA molecules in reactions that take place independently of the Ku heterodimer. These reactions are further distinguished from NHEJ by their use of short stretches of sequence microhomology (5-15 bp in length) (Burton *et al.*, 2007). Furthermore, these microhomology mediated reactions observed *in vitro* are highly reminiscent of reactions observed *in vitro* in *rad51-/-* mutants (Conway *et al.*, 2002c).

Research has revealed that *T. brucei* possesses a functional MMR system, with homologues of MSH2, MSH3, MSH6, MLH1 and PMS1 present (Bell *et al.*, 2004). Mutation of either MSH2 or MLH1 caused an increased frequency of homologous recombination between both perfectly matched and diverged DNA sequences (Bell *et al.*, 2004). The same mutation also resulted in an increase in the rate of sequence variation at a number of microsatellite loci, and an increased tolerance to the alkylating agent N-methyl-N'-nitro-N-nitrosoguanidine, both of which are consistent with mismatch repair impairment (Bell *et al.*, 2004). Despite this, no detectable difference was observed on

VSG switching, indicating either that MMR acts too subtly to be detected by the assays or that mismatch selection does not act on VSG recombination reactions (Bell and McCulloch, 2003).

Further insights into homologous recombination were discovered through the generation of *mre11-/-* mutants (Robinson *et al.*, 2002). These mutants displayed an impaired growth phenotype and an impaired ability to perform homologous recombination, phenotypes highly comparable to those of *rad51-/-* mutants. However, notable differences to *rad51-/-* mutants were discovered, including a lack of sensitivity to MMS and an accumulation of gross chromosomal rearrangements. These results indicate the importance of Mre11 in the repair of chromosomal damage and DSBs in *T. brucei*. Perhaps the most surprising result was that *mre11-/-* mutants did not display any defects in VSG switching, despite the clear importance of Mre11 in homologous recombination (Robinson *et al.*, 2002).

More recently, further insight has been provided into homologous recombination and antigenic variation in *T. brucei* with the discovery of five *RAD51*-related genes: *DMC1, RAD51-3, RAD51-4, RAD51-5* and *RAD51-6* (Proudfoot and McCulloch, 2005). To date, however, the interactions of these proteins with each other and with RAD51 remain unknown. Although damage-induced RAD51 foci have been demonstrated to form in *T. brucei* (Proudfoot and McCulloch, 2005;Proudfoot and McCulloch, 2006) their precise nature and composition also remains unclear. *RAD51-3* and *RAD51-5* mutants have been generated and display growth impairment, sensitivity to DNA damaging agents and an impaired ability to perform homologous recombination. However, only *RAD51-3* was seen to have an effect on VSG switching, with the results being highly reminiscent of those obtained for *rad51-/-* mutants whereby VSG switching events still occurred at a low level. These results suggest that the family of RAD51 proteins present in *T. brucei* have assumed specialized functions in homologous recombination, comparable with related proteins in metazoan eukaryotes (Proudfoot and McCulloch, 2005). *dmc1-/-* mutants, however, behaved quite similarly to that of wild type cells, suggesting that DMC1 does not have an important role in homologous recombination or *VSG* switching, at least in the bloodstream stage of *T. brucei* (Proudfoot and McCulloch, 2006).

In summary, therefore, to date, only two proteins have been identified in *T. brucei* that have been shown to function in VSG switching, despite the characterisation of a range of repair proteins. It is highly unlikely, however, that these are the only proteins to influence antigenic variation in this organism.

1.8 Aims of the thesis

The overall aim of this thesis was to further examine the factors that regulate antigenic variation in *Trypanosoma brucei*, with the hope that this would shed further light in the relationship between VSG switching and homologous recombination.

The first aspect examined in this thesis was the *T. brucei* homologue of BRCA2. In part, this stemmed from the suggestion that the *T. brucei BRCA2* homologue has a highly unusual organisation, in that it is proposed to contain 15 BRC repeats, a much higher number than observed in any other organism (Lo *et al.*, 2003).

A clear hypothesis is that this unusual structural organisation is a consequence of the high levels of recombination needed by *T. brucei* during antigenic variation. This was tested in a number of approaches:

(i) Examination of the *BRCA2* structure in a number of *T. brucei* and Trypanosome strains.

(ii) Generation of *T. brucei BRCA2* knockout mutants in order to determine the proteins function.

(iii) Generation of *T. brucei BRCA2* mutants with decreased numbers of BRC repeats in order to determine why *T. brucei* BRCA2 has so many BRC repeats.

The second area of investigation in this thesis was the role and molecular composition of *T. brucei* RAD51 sub-nuclear foci (Proudfoot and McCulloch, 2005). The hypothesis was tested that the foci are repair centres containing multiple homologous recombination factors and specific sites of DNA lesions. To do this, the tandem affinity purification (TAP) method (Rigaut *et al.*, 1999) was used to attempt to identify RAD51 interacting factors, before and after induced DNA damage.

CHAPTER 2

Materials and Methods

2.1 Trypanosome culture

2.1.1 Trypanosome strains and their growth

2.1.1.1 Bloodstream stage cells

The bloodstream form *Trypanosoma brucei* strains used in this thesis are Lister 427 MITat 1.2a (McCulloch *et al.*, 1997;Rudenko *et al.*, 1996) and its transgenic derivative, 3174.2. The Lister 427 strain MITat1.2a (expressing VSG221), was derived from many syringe passages through rodents over a number of years, although its exact derivation is uncertain (Melville *et al.*, 2000). This is a monomorphic strain, which usually only displays the long slender bloodstream form. The switching frequency of the VSG being expressed is approximately 1 x 10^6 to 1 x 10^7 switches/cell/generation. The 3174.2 strain contains hygromycin and G418 resistance genes in the expression site containing *VSG221*, which allows the analysis of VSG switching. *In vitro* growth of *T. brucei* bloodstream stage cells was carried out using HMI-9 growth medium (Hirumi and Hirumi, 1989) at 37 °C in a humidified 5 % CO_2 incubator. The population doubling time of this strain is approximately 8 hours (Proudfoot and McCulloch, 2006). To keep a working culture of *T. brucei* bloodstream stage cell lines, cells were passaged three times weekly by the addition of 20 μl of a log-phase culture (at a density of ~ 4 x 10^6 cells.ml^{-1}) to 1.5 ml HMI-9 medium in a 24-well plate. Bloodstream stage *T. brucei* were grown in petri dishes in volumes of 25 mls to obtain large numbers of cells for experiments. *In vivo* growth was carried out using adult female ICR mice (approximately 25 g) infected by intraperitoneal injections.

2.1.1.2 Procyclic form cells

The procyclic form *Trypanosoma brucei* cells used in this study are of strain East African Trypanosomiasis Research Organisation (EATRO) 795. *In vitro* growth of procyclic form trypanosomes was carried out using SDM-79 growth medium (Brun and Schonenberger, 1979) at 27 °C. To keep a working culture of *T. brucei* procyclic form cell lines, cells were passaged twice weekly by addition of approximately 1000 μl of a log-phase culture (at a density of ~ 8 x 10^6 cells.ml^{-1}) to 9 mls SDM-79 medium in a 25 cm^2 tissue culture flask. Procyclic form *T. brucei* were grown in 75 cm^2 tissue culture flasks in volumes of up to 100 mls to obtain large numbers of cells for nuclear extracts.

2.1.2 Stabilate preparation and retrieval

For the long term storage of trypanosomes, stabilates were prepared by adding 100 μl of sterile 100 % glycerol to 900 μl of *T. brucei* culture at a density of ~ 2 x 10^6 cells.ml^{-1} (bloodstream stage cells) or ~ 7 x 10^6 cells.ml^{-1} (procyclic form cells). These 1 ml aliquots were placed in 1.2 ml cryotubes (Nunc), before freezing at - 80 ˚C overnight and then transferring to liquid nitrogen. For retrieval of stabilates from liquid nitrogen, the cells were defrosted at 37 ˚C (bloodstream stage cells) or 27 ˚C (procyclic form cells), and placed in 10 mls HMI-9 growth medium (bloodstream stage cells) or 5 mls SDM-79 growth medium (procyclic form cells) overnight; the cells were then passaged normally as described above.

2.1.3 Transformation of trypanosomes

2.1.3.1 Transformation of bloodstream stage trypanosomes

T. brucei bloodstream stage cultures were grown to a density of 1-2 x 10^6 cells.ml^{-1} and centrifuged at room temperature for 10 minutes at 583 x g. The cells were resuspended in Zimmerman post-fusion medium (5 M NaCl, 1 M KCl, 1 M Na_2HPO_4, 1 M KH_2HPO_4, 1 M MgOAc, 0.2 M $CaCl_2$, pH 7.0) supplemented with 1 M D-glucose (ZMG), at a concentration of 1 x 10^8 cells.ml^{-1}. 5 x 10^7 cells per transformation were electroporated in 0.5 mls ZMG at 1.5 kV and 25 μF capacitance using a BioRad Gene Pulser II. Approximately 5 μg of purified DNA that had been restriction digested, phenol-chloroform extracted and ethanol precipitated was routinely used for transformations. After electroporation, cells were placed in 10 mls of HMI-9 for three population doubling times (normally 24 hours) before being subjected to antibiotic selection. For this, the recovered cells were centrifuged at room temperature for 10 minutes at 583 x g and resuspended in HMI-9 containing the appropriate antibiotic at a concentration of 5 x 10^5 cells.ml^{-1}. 1-2 x 10^7 cells (unless otherwise stated) were plated out in 1.5 ml aliquots over 24 well plates. Transformants were counted after 7-10 days by looking at the plates under a light microscope (Leitz) and counting the number of wells that contained growing cells. The population of cells in a well should have descended from a single transformant, and could therefore be considered as clonal, so long as less than 80 % of the wells contain living cells (Wickstead *et al.*, 2003b).

2.1.3.2 Transformation of procyclic form trypanosomes

T. brucei procyclic form cultures were grown to a density of 1-2 x 10^6 cells.ml^{-1} and centrifuged at room temperature for 10 minutes at 580 x g. The cells were resuspended in

Zimmerman post-fusion medium (5 M NaCl, 1 M KCl, 1 M Na_2HPO_4, 1 M KH_2PO_4, 1 M MgOAc, 0.2 M $CaCl_2$, pH 7.0) (ZM), at a concentration of 1 x 10^8 cells.ml^{-1}. 5 x 10^7 cells per transformation were electroporated twice in 0.5 mls ZM at 1.5 kV and 25 μF capacitance using a BioRad Gene Pulser II. Approximately 5 μg of purified DNA that had been restriction digested, phenol-chloroform extracted and ethanol precipitated was routinely used for transformations. After electroporation, cells were placed in 10 mls of SDM-79 for three population doubling times (normally 24 hours) before being subjected to antibiotic selection. The recovered cells were resuspended in SDM-79 containing the appropriate antibiotic at concentrations of 10^4, 10^5, 10^6 and 10^7 cells in 10 ml cultures. Cultures containing less than 10^6 cells were either supplemented with 10^6 wild type cells or were placed in conditioned media (75 % SDM-79, 10 % FBS, 15 % SDM-79 conditioned by growth of procyclic form cells to approximately 8 x 10^6.ml^{-1}, centrifuged and filter sterilised to remove trypanosomes). Transformants typically grew through after 7-14 days. After this period, the transformant polyclonal population was cloned by plating out 1 cell per well over 96 well plates containing conditioned media and the appropriate antibiotic. These were left to grow for 10-14 days before identifying clonal wells under a light microscope (Leitz).

2.1.4 Analysis of growth

2.1.4.1 Analysis of *in vitro* growth

In vitro growth analysis was carried out on bloodstream stage *T. brucei* by inoculating 2 ml cultures with 5 x 10^4 cells.ml^{-1}, previously grown in culture to a density of 1-2 x 10^6 cells.ml^{-1}. For procyclic form *T. brucei*, a 2 ml culture was inoculated at 5 x 10^5 cells.ml^{-1}, previously grown in culture to a density of 7 x 10^6 cells.ml^{-1}. The numbers of cells were counted at 24, 48, 72 and 96 hours subsequently using a haemocytometer (Bright-line, Sigma). Three or four repetitions of each cell line were carried out and the results plotted on a semi-logarithmic scale. The population doubling times were calculated by examining the linear phase of the graph and represented as a mean of the population doubling times calculated.

2.1.4.2 Analysis of *in vivo* growth

In vivo growth rates were examined by intraperitoneally injecting ICR mice with 1 x 10^6 trypanosomes, previously grown in culture to a density of 1-2 x 10^6 cells.ml^{-1}. The density of trypanosomes was determined every 24 hours up to a maximum of 120 hours before sacrificing the mice. A small volume of blood was removed from the tail of each mouse and placed into heparin-coated capillary tubes (Hawksley). 1 μl samples of blood were

diluted in 99 μl of 0.85 % ammonium chloride. This solution preferentially lyses red blood cells, therefore allowing the *T. brucei* to be visualised and counted in a haemocytometer (Bright-line, Sigma). The results were plotted on a semi-logarithmic scale and population doubling times calculated.

2.1.5 Cell cycle analysis

In order to examine the cell cycle, trypanosomes were prepared for microscopy analysis by DAPI staining (4, 6-diamidino-2-phenylindole) (Vector Laboratories Inc.) (section 2.11.1). Differential interface contrast (DIC) was used to visualise intact cells and UV to visualise DAPI. Cells were counted according to the number of nuclei and kinetoplast they contained. Cells in G1 phase of the cell cycle contain 1 nucleus and 1 kinetoplast (1N 1K). Kinetoplast division then occurs resulting in cells with 1 nucleus and 2 kinetoplasts (1N 2K). After this, the nucleus divides leading to cells containing 2 nuclei and 2 kinetoplasts (2N 2K). Completion of cell division forms two daughter cells in G1 phase, containing 1 nucleus and 1 kinetoplast (1N 1K). Any cells that were observed not to be in these cell cycle phases were noted as being aberrant cell types and were described as 'others'.

2.1.6 Analysis of DNA damage sensitivity

Sensitivities of *T. brucei* cell lines to methyl methane sulphonate (MMS) were assayed by a clonal survival assay and an Alamar blue assay. The Alamar blue assay was also utilised to measure the sensitivities of *T. brucei* cell lines to phleomycin.

2.1.6.1 Clonal survival assay

The clonal survival assay was performed by growing cultures to a maximum density of 2 x 10^6 cells.ml^{-1} and plating out one cell per well over five 96 well plates, containing an MMS concentration of 0, 0.0001, 0.0002, 0.0003 or 0.0004 %. Four repetitions for each strain were carried out and the number of wells containing a viable parasite population after 20 days of growth was counted. The number of wells growing on the plate without MMS was taken as being 100 % and the number of wells growing through on the MMS containing plates calculated relative to this, thereby removing any errors due to plating efficiency and growth rates.

2.1.6.2 Alamar blue assay

Reduction of Alamar blue (resazurin) was examined by growing cultures to a maximum density of 2 x 10^5 cells.ml^{-1} and placing 100 μl into 11 wells each containing 100 μl of media with serially decreasing amounts of drug (either MMS or phleomycin). After 48

hours of growth, 20 μl of Alamar blue (12.5 mg.ml^{-1} resazurin, Sigma) was added. The plates were left for a further 24 hours for the cells to metabolise the resazurin, which is blue and non-fluorescent. When resazurin is reduced to resorufin is becomes pink and highly fluorescent (O'Brien *et al.*, 2000;Raz *et al.*, 1997;Onyango *et al.*, 2000). This fluorescence was then measured on a Perkin Elmer LS55 Luminometer at 539 nm excitation and 590 nm emission. Three repetitions were performed, IC50's calculated and mean IC50s plotted graphically.

2.1.7 Transformation efficiency assay

To examine the ability of *T. brucei* cells to undergo recombination, a transformation assay was used. This assay involves the transformation of an antibiotic resistance marker (for the work in this thesis it was hygromycin), flanked by tubulin intergenic sequences into the cell lines. The construct, named *tubHYGtub,* targets to the *TUBULIN* array, replacing an alpha tubulin gene and conferring hygromycin resistance.

In each transformation, 5 x 10^7 cells were electroporated with 5 μg of construct DNA as described in section 2.1.3. The transformed cells were recovered in 10 mls of media for three generations before being plated out in selective media containing 5 μg.ml^{-1}of hygromycin (Roche). 5 x 10^6 cells were plated out over 24 wells for the wild type and heterozygous cells, whilst 2 x 10^7 cells was plated out over 48 wells for homozygous cells, where fewer transformants were expected. The number of wells containing antibiotic resistant transformants were counted after 14 days and expressed as the number of transformants per 10^6 cells plated out.

2.1.8 VSG switching analysis

The method used in this study to analyse the frequency and mechanism of VSG switching is based upon that used by McCulloch *et al* (1997), McCulloch and Barry (1999) and Proudfoot and McCulloch (2005, 2006).

2.1.8.1 Analysis of VSG switching frequency

Mice were generated with acquired immunity primarily against *VSG221* by injecting intraperitoneally 2 x10^5 wild type 3174.2 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. The trypanosomes were allowed to proliferate in the mice for 3 to 4 days before curing the mice by injection of cymerlarsan (Rhone Merieux; 5 mg.kg^{-1}). In order to generate switched variants, cell lines previously grown on hygromycin (5 μg.ml^{-1}) and G418 (2.5 μg.ml^{-1}) for a period of 5

days were removed from antibiotic selection for a period of 9 generations before injecting

$4-8 \times 10^7$ cells into the immune mice. After 24 hours, the mice were exsanguinated by

cardiac puncture, and the blood withdrawn into 5 % sodium citrate anticoagulant in Carters

Balanced Salt Solution (CBSS - 0.023 M HEPES, 0.12 M NaCl, 5.41 mM KCl, 0.55 mM

$CaCl_2$, 0.4 mM $MgSO_4$, 5.6 mM Na_2HPO4, 0.035 M glucose, 0.04 mM phenol red, pH

adjusted to 7.4) (0.15 ml CBSS/5 % sodium citrate per 0.85 ml blood). The surviving *T.*

brucei cells were recovered by centrifuging 2 x 0.4 ml of exsanguinated mouse blood at

5000 rpm in a micro-centrifuge for 5 minutes. The centrifugation separated the blood into

red blood cells in the bottom layer, white blood cells and *T. brucei* cells in the middle layer

and plasma in the top layer. In order to isolate clonal switched variants, the top and middle

layers were removed with a 1 ml syringe and a 19 gauge needle and added to 40 mls HMI-

9, before plating out over 2 x 96 well plates. Cells were allowed to grow for up to 2 weeks

before identifying the number of wells that had grown under a light microscope (Leitz).

Variations of this assay, using Lister 427 MITat1.2a cells are discussed in the results.

2.1.8.2 Analysis of VSG switching mechanism

The mechanisms of VSG switching that had been used in the switched variants were

determined through growth on antibiotic selection and through PCR-amplification of

resistance cassettes. 10 μl of the cells that were recovered from the immunised mice were

passaged into 1.5 mls HMI-9 containing either hygromycin (5 $\mu g.ml^{-1}$), G418 (2.5 $\mu g.ml^{-1}$)

or no drug. After a period of 10 days the cells were scored for their antibiotic resistance.

The cells grown on no drug were then passaged into 5 ml cultures of HMI-9 and grown for

a further 2 days before preparing genomic DNA. PCR analysis was performed on the

genomic DNA to determine the presence or absence of the hygromycin and G418

resistance genes (primers are described in the text and listed in appendix 1).

2.2 Isolation of material from trypanosomes

2.2.1 Isolation of genomic DNA

Genomic DNA that was to be utilised for PCR alone was prepared using the Qiagen

DNeasy® Tissue kit. 5 mls of bloodstream stage *T. brucei* grown to a density of ~ 4×10^6

cells.ml^{-1} or 2 mls of procyclic form *T. brucei* grown to a density of ~ 8×10^6 cells.ml^{-1}

were harvested by centrifugation at 1620 x g for 10 minutes at room temperature. The cell

pellet was resuspended in 200 μl PBS, before 20 μl proteinase K (>600 $mAU.ml^{-1}$) and 200

μl Buffer AL were added to the sample and mixed by vortexing. Following an incubation

at 70 °C for 10 minutes, 200 μl of 100 % ethanol was added to the sample and mixed by

vortexing. The sample was subsequently added to a DNeasy Mini spin column placed in a 2 ml collection tube. The samples were centrifuged at 6000 x g in a micro-centrifuge for 1 minute and the flow through discarded. 500 μl of Buffer AW1 was added to the column and centrifuged at 6000 x g in a micro-centrifuge for 1 minute. The flow through was discarded and 500 μl of Buffer AW2 added to the column. The column was then centrifuged at 14000 x g in a micro-centrifuge for 3 minutes, to dry the DNeasy membrane. The flow through was discarded and the column centrifuged again at 14000 x g in a micro-centrifuge for 1 minute to ensure that all residual ethanol was removed. The DNeasy column was placed in a fresh eppendorf and 200 μl of Buffer AE added directly onto the membrane. After a 1 minute incubation the DNA was eluted from the column by centrifuging at 6000 x g in a micro-centrifuge for 1 minute.

Genomic DNA that was to be subsequently used for restriction digestion and Southern analysis was prepared using the following protocol. 25 mls of bloodstream stage *T. brucei* grown to a density of ~ 4 x 10^6 cells.ml^{-1} or 10 mls of procyclic form *T. brucei* grown to a density of ~ 8 x 10^6 cells.ml^{-1} were harvested by centrifugation at 1600 x g for 10 minutes at room temperature, and resuspended in 500 μl of trypanosome lysis buffer (1 mM ethylenediaminetetraacetic acid (EDTA), 100 mM NaCl, 50 mM Tris-HCl, pH 8). 50 μl of 10 % sodium dodecyl sulphate (SDS) and 2.5 μl of a 20 μg.μl^{-1} proteinase K solution were then added, and the solution incubated at 37 ˚C overnight to lyse the trypanosomes and digest the proteins. The DNA was recovered from the lysis reaction by phenol/chloroform extraction and ethanol precipitation (section 2.2.11).

2.2.1.1 Phenol: Chloroform extraction and ethanol precipitation

An equal volume of a 1:1 mixture of phenol/chloroform (Sigma) was added to the lysis reaction and mixed by gentle inversion. The phenol and aqueous phases were then separated by centrifugation at 16000 x g in a micro-centrifuge for 1 minute at room temperature. The upper aqueous phase containing the DNA was transferred to a new eppendorf tube, where 2 volumes of 100 % ethanol and 1/10 volume 3 M sodium acetate (pH 5.2) were added. This solution was mixed by inverting the tube several times, and incubated at - 20 ˚C for 30 minutes to overnight. The DNA was harvested by pelleting through centrifugation at maximum speed in a micro-centrifuge for 30 minutes at 4 ˚C. The 100 % ethanol was removed by aspiration and the nucleic acid pellet washed by addition of 100 μl 70 % ethanol, followed by centrifugation at 16000 x g in a micro-centrifuge for 2 minutes at room temperature. The 70 % ethanol was removed by aspiration and the pellet air-dried. The genomic DNA was resuspended typically in a volume of 30 μl of sterile dH$_2$O or TE buffer (100 mM Tris, 10 mM EDTA, pH 7.4).

DNA was quantified spectrophotometrically at 260 nm and multiplied by 50 to give the approximate concentration in $\mu g.ml^{-1}$ of double-stranded DNA. When the concentration of DNA was not sufficient to quantify spectrophotometrically, the amount of DNA was visualised under UV induced fluorescence emitted by ethidium bromide. The DNA quantification was estimated by comparing the sample to that of a known standard. In this case, the 1 kb DNA ladder was used (Invitrogen), where the 1.6 kb band contains 10 % of the mass applied to the gel.

2.2.2 Isolation of total RNA

25 mls of bloodstream stage *T. brucei* grown to a density of ~ 4 x 10^6 cells.ml^{-1}, or 10 mls of procyclic form *T. brucei* grown to a density of ~ 8 x 10^6 cells.ml^{-1}, were harvested by centrifugation at 1600 x g for 10 minutes at room temperature and removing the supernatant. Total RNA was isolated using the RNeasy® mini kit (Qiagen) following the manufacturer's instructions. The pelleted cells were resuspended and lysed in 600 µl of Buffer RLT (containing the appropriate amount of 2-mercaptoethanol). The sample was homogenised by passing the lysate 5 times through a 25 gauge needle fitted to a RNase-free 1 ml syringe. 600 µl of 70 % ethanol was added to the sample and mixed by pipetting. 700 µl of this solution was applied to a RNeasy column placed in a 2ml collection tube, and centrifuged for 15 seconds at 16000 x g in a micro-centrifuge, discarding the flow-through. The column was washed by applying 700 µl of Buffer RW1 and centrifuging for 15 seconds at 16000 x g. The column was transferred to a new collection tube before applying 500 µl of Buffer RPE and centrifuging for 15 seconds at 16000 x g. The flow-through was discarded and a final wash step carried out by applying another 500 µl of Buffer RPE to the column and centrifuging for 2 minutes at 16000 x g. RNA was eluted from the column by placing it in an RNase-free eppendorf tube, adding 30 µl of RNase-free dH$_2$O and centrifuging for 1 minute at 16000 x g. RNA was quantified spectrophotometrically at 260 nm and multiplied by 40 to give the approximate concentration in $\mu g.ml^{-1}$ of single-stranded RNA.

2.2.3 Isolation of protein extract

25 mls of bloodstream stage *T. brucei* grown to a density of ~ 4 x 10^6 cells.ml^{-1}, or 10 mls of procyclic form *T. brucei* grown to a density of ~ 8 x 10^6 cells.ml^{-1}, were harvested by centrifugation at 1620 x g for 10 minutes at room temperature and washed twice in PBS. The pelleted cells were resuspended in 1 ml of PBS before centrifuging at 2400 x g for 10 minutes in a micro-centrifuge. The pellet was resuspended in SDS-PAGE sample buffer (0.5 M Tris-HCL, pH 6.8, 10 % Glycerol, 10 % SDS, 5 % 2-mercaptoethanol, 0.05 %

(w/v) bromophenol blue) to a concentration of 10^9 cells.ml^{-1} and PBS to a concentration of 5×10^8 cells.ml^{-1}. Protein extracts were denatured at 95 °C for 5 minutes prior to loading.

2.2.4 Preparation of genomic plugs

Each genomic agarose plug prepared for the work in this thesis contained 5×10^7 bloodstream stage *T. brucei*. Bloodstream stage *T. brucei* were grown to a density of ~2 x 10^6 cells.ml^{-1}, centrifuged at 583 x g for 10 minutes at room temperature and washed twice in 10 mls PSG (1 x PBS, 1 % w/v glucose). The pellet was then resuspended in 50 μl PSG and warmed at 37°C for 1 minute, before adding an equal volume of pre-warmed 1.4 % low melting agarose (agarose for PFGE sample preparation, Sigma) made with dH$_2$0. This mixture was swirled to mix before filling disposable plug moulds (BioRad) with 50 μl of the agarose/trypanosome solution and placing at 4 °C for ~ 4 h to set. The agarose plugs were then removed from the moulds, incubated in NDS buffer, pH 9.0 (0.5 M EDTA, 10 mM Tris base and 34.1 mM lauroyl sarcosine) containing 1 mg.ml^{-1} proteinase K at 55 °C for ~ 24 h and then transferred into NDS buffer pH 8.0 containing 1 mg.ml^{-1} proteinase K at 55 °C for ~ 24 h. The plugs were finally transferred into NDS buffer pH 8.0 for storage at 4 °C.

2.2.5 Trypanosome nuclear extract preparation

Nuclear extracts were prepared by the method described in Bell and Barry (1995). 3×10^9 procyclic form trypanosomes were harvested by centrifugation at 1600 x g for 10 minutes at 4°C and washed twice in ice cold PBS. The pelleted trypanosomes were resuspended in two packed cell volumes of Buffer A (20 mM Tris, pH 7.9, 10 mM NaCl, 0.5 mM DTT and protease inhibitors). The cells were lysed with 60 strokes of a Dounce homogeniser (Type A pestle). To pellet the nuclei, the homogenate was centrifuged at 3700 x g for 5 minutes at 4°C. The nuclei were then resuspended in Buffer C (50 mM Tris, pH 7.9, 400 mM NaCl, 0.2 mM EDTA, 0.5 mM DTT, 25 % glycerol and protease inhibitors) to a concentration of 10^{10} cells.ml^{-1}, and the pellet homogenised with a further 50 strokes of a Dounce homogeniser. This homogenate was then mixed by rotation for 30 minutes at 4 °C. This nuclear lysate was then centrifuged at 25000 x g for 30 minutes at 4 °C. The supernatant was then dialysed against the 50 volumes of the buffer being used in the TAP purification.

2.3 Electrophoresis

DNA and RNA electrophoresis gels were visualised using a trans-UV illuminator and Gel Doc software (BioRad).

2.3.1 DNA electrophoresis

Standard DNA separations were performed on 1.0 % agarose gels (Seakem LE agarose, BioWhittaker Molecular Applications) made with 1 x TAE buffer (40 mM Tris, 19 mM acetic acid, 1 mM EDTA, pH 8.0) and containing 0.2 $\mu g.ml^{-1}$ ethidium bromide (Sigma) or 1.25 $X.ml^{-1}$ SYBRSafe (Invitrogen). Typically, the separations were run in 1 x TAE buffer at 100 V. A commercial 1 kb DNA ladder was used as a size marker (Invitrogen) and apparatus was supplied by Gibco BRL, BioRad or Sigma. Separating genomic DNA digests for Southern blotting analysis was carried out on 0.8 % agarose gels, made with 1 x TAE buffer, electrophoresed in 1 x TAE buffer at ~ 30 V overnight.

2.3.2 RNA electrophoresis

RNA molecules were separated by electrophoresis on 1 % agarose gels (Seakem LE agarose, BioWhittaker Molecular Applications) made with 0.4 x MNE buffer (MOPS/Sodium acetate/EDTA buffer: 1 x: 0.024M MOPS, 5mM NaOAc, 1mM EDTA, pH 7.0) and containing 2.46 M formaldehyde. Gels were typically run for ~16 h at 30 V in 1 x MNE buffer, using a commercial 0.5 – 9.0 kb (New England Biolabs) ladder as a size marker. RNA samples (typically 10-20 μg) were added to 20 μl RNA loading buffer (7.38 M formaldehyde, 20 % v/v formamide, in 1 x MNE buffer) and 1 μl ethidium bromide at 0.2 $\mu g.ml^{-1}$, and incubated at 65 ˚C for 5 minutes before loading.

2.3.3 Pulsed field gel electrophoresis

Prior to electrophoretic separation, the pulsed field gel electrophoresis (PFGE) apparatus (CHEF-DR III, BioRad) was cleaned by the circulation of 2 litres of 0.1 % SDS overnight at 20 ˚C. The tank was then rinsed twice by circulating dH_2O for ~ 1 h at 15 ˚C, and once by circulating the appropriate electrophoresis buffer for ~ 1 h at 15 ˚C. 1 x TB1/10E (90 mM Tris base, 90 mM boric acid, 2 mM EDTA) was used for the separation of mega-base chromosomes, whereas 0.5 x TBE (45 mM Tris base, 45 mM boric acid, 10 mM EDTA, pH 8.0) was used for the separation of intermediate and mini chromosomes. Gels were electrophoresed in 2 litres buffer, which was circulated in the tank for at least 30 minutes at 15 ˚C before the gel was run.

All separations were conducted using 1.2 % agarose (Seakem LE, BioWhittaker Molecular Applications). Agarose was dissolved in 150 mls of the appropriate electrophoresis buffer, and 140 mls used to prepare a gel using the tray provided with the PFGE system, keeping the remainder at 37 ˚C. After the agarose gel had set, the comb was removed, agarose genomic plugs placed into the wells, and the wells sealed with the remaining agarose. The

agarose genomic plugs had been prepared by 3 rounds of dialysis in the appropriate electrophoresis buffer. Gels were electrophoresed at 15 ˚C, either at 2.5 V.cm^{-1} for 144 hours with an initial switch time of 1400 seconds and final switch time of 700 seconds for the separation of megabase chromosomes, or at 5.8 V.cm^{-1} for 24 hours with initial and final switch times of 20 seconds for the separation of intermediate and mini chromosomes. Chromosomes were visualised by placing agarose gels in 200 mls of electrophoresis buffer containing 4 µl ethidium bromide at 10 µg.ml^{-1} and placing on a rocking table for ~ 30 minutes. They were then de-stained in dH$_2$O for ~ 30 minutes, or until they could be visualised clearly by UV illumination.

2.3.4 Protein electrophoresis

Protein samples were fractionated on either Bio-Rad Ready Gels (10 % Tris-HCL), 12 % NuPAGE® Novex® Bis-Tris mini gels (Invitrogen), or on SDS-polyacrylamide gels made up to the desired percentage using 37 % acrylamide (Sigma), 10 % APS (ammonium persulphate) and TEMED (N, N, N', N' – Tetramethylethylenediamine, Sigma) to facilitate the polymerisation of the acrylamide between 2 glass plates. The gels were electrophoresed in 1 x SDS running buffer (0.19 M Glycine, 0.025 M Tris, 0.03 M SDS) at 175 V, using either the Mini-PROTEAN 3 Cell system (Bio-Rad) or the XCell Surelock™ Mini-Cell (Invitrogen).

The gels were either prepared for Western blots by transfer to a nylon membrane (see section 2.4.3), or proteins were visualised by Coomassie or Sypro Ruby (Bio-Rad) staining. For Coomassie staining, gels were placed in Coomassie stain solution (0.25 g Coomassie brilliant blue R [Sigma] in 90 ml of methanol: water [1:1 v/v] and 10 mls glacial acetic acid), and placed on a rocker for 45 minutes to 4 hours. Visualisation of protein bands was achieved by placing the gels in destaining solution (10 % glacial acetic acid, 40 % methanol) for 1-3 hours.

For Sypro Ruby staining, gels were placed in 50 ml Sypro Ruby and left overnight on a rocker. Gels were destained by washing in 10 % methanol, 7 % acetic acid and visualised using a trans-UV illuminator and Gel Doc software (BioRad).

2.4 Blotting

2.4.1 Southern blotting

Agarose gels to be Southern blotted were photographed on a UV transilluminator, alongside a ruler, parallel to the gel in order to allow calculation of the sizes of fragments hybridised by radioactively labelled DNA (see section 1.5). To depurinate the DNA, the gels were soaked in 125 mM HCl for 15 minutes and then rinsed with distilled water. The DNA was then denatured by placing the gel in denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 30 minutes. Following rinsing with distilled water, the gel was placed in neutralisation solution (1 M Tris base, 1.5 M NaCl, 186 mM HCl, pH 7.2) for 30 minutes. The gel was rinsed again in distilled water, before rinsing in 20 x SSC transfer buffer (3 M NaCl, 300 mM NaOAc, pH 7.0). The DNA was subsequently transferred to a nylon membrane (Hybond XL, Amersham Biosciences) by overnight capillary blotting (Sambrook *et al.*, 1989) using 20 x SSC transfer buffer. Following transfer, the DNA was cross-linked to the membrane using the auto-crosslink setting on a UV Stratalinker (Stratagene).

Pulsed field gels were Southern blotted essentially as described above, but with slightly different wash treatments, due to the chromosomes being tightly bound within the agarose. After ethidium bromide staining, the chromosomes were nicked by soaking the gels twice in 125 mM HCl for 7 minutes. After rinsing in distilled water the chromosomes were denatured by soaking in denaturation solution twice for 15 minutes. The treatment then resumed as with the above protocol, apart from the capillary blotting, which was usually performed for at least 48 hours.

2.4.2 Northern blotting

Agarose gels to be northern blotted were photographed on a UV transilluminator, alongside a ruler, parallel to the gel to allow calculation of the sizes of fragments hybridised by radioactively labelled DNA (see section 1.5). Gels were soaked in sodium phosphate (10 mM Na_2HPO_4/NaH_2PO_4, pH 6.5) for 15 minutes to remove any residual formaldehyde, before the transfer of RNA to a nylon membrane (Hybond XL, Amersham Biosciences) by overnight capillary blotting (Sambrook *et al.*, 1989) using sodium phosphate as the transfer buffer. The RNA was then cross-linked to the membrane using the auto-crosslink setting on a UV Stratalinker (Stratagene).

2.4.3 Western blot transfer

Western blotting of protein gels was carried out using the Mini Trans-Blot®
Electrophoretic Transfer Cell (Bio-Rad). Gels and Trans-Blot® nitrocellulose membrane
were equilibrated in transfer buffer (0.19 M Glycine, 0.025 M Tris base, 200 ml methanol,
800 ml dH_20, pH 8.0), before assembling the gel sandwich. The sandwich consisted of the
gel and the nitrocellulose membrane, surrounded by filter paper and foam, sandwiched
between a plastic cassette. An ice block was placed alongside the cassette to prevent
overheating. Transfer was carried out by electrophoresing at 100 V for 1 hour.

2.5 Radiolabelling and hybridisation of DNA probes

2.5.1 Probe manufacture by random hexamer labelling of DNA

The DNA fragments used for probe manufacture were specific PCR products amplified as
described in section 2.7.1, separated on an agarose gel and purified using the Qiagen gel
extraction kit, following the manufacturer's protocol (section 2.7.1.2).

Radio-labelling of these fragments was performed using the Prime It II kit (Stratagene).
Approximately 25 ng of DNA was mixed with 10 μl random hexameric oligonucleotides
(27 OD units.ml^{-1}) and dH_20 in a final reaction volume of 36 μl. The DNA was denatured
by incubation at 95 ˚C for 5 minutes. 10 μl of 5 x dATP or dCTP primer buffer, 2 μl of
α^{32}P-labelled dATP or dCTP (~ 0.74 MBq) and 1 μl Klenow DNA polymerase (5 U.μl^{-1})
were added and the reaction incubated at 37 ˚C for 4-10 minutes. The probes were then
purified from any unincorporated nucleotides by size exclusion chromatography using
Microspin columns (Amersham Biosciences) according to the manufacturer's protocol.
After purification, the probes were denatured at 95 ˚C for 5 minutes before hybridisation.

2.5.2 Hybridisation of radiolabelled DNA probes

Nylon filters blotted with DNA or RNA (sections 2.4.1 and 2.4.2) were placed in
hybridisation tubes (Hybaid) with approximately 50mls of pre-warmed 0.5 M Church
Gilbert solution (342 mM Na_2HPO_4, 158 mM NaH_2PO_4, 7 % SDS, 1 mM EDTA, pH 7.2)
and pre-hybridised for a minimum of 1 hour at 65 ˚C in a rotating hybridisation oven. The
denatured, radiolabelled probe (section 1.5) was then added to the Church Gilbert solution
in the hybridisation tube and allowed to hybridise to the blot overnight at 65 ˚C in a
rotating hybridisation oven. Following hybridisation, the filters were washed in a rotating
hybridisation oven with 50 mls of 2 x SSC, 0.1 % SDS for 30 minutes at 65 ˚C and then 50
mls of 0.2 x SSC, 0.1 % SDS for another 30 minutes. After washing, the filters were

blotted with filter paper (Whatman) to remove any excess liquid, before sealing in plastic and exposing to a phosphorimaging screen (Fuji) at room temperature for 4-72 hours (depending on the strength of the signal). The phosphorimaging screen was then visualised using a Typhoon 8600 phosphorimager (Amersham Biosciences).

2.5.3 Stripping of hybridised nylon membranes

To strip nylon membranes of hybridised probe DNA, membranes were placed in a heat-proof container, with boiling 0.1 % SDS. After allowing the solution to cool to room temperature, the SDS solution was poured off and the procedure repeated. Successful stripping was checked by exposure to a phosphorimage screen (Fuji) for 24 h and visualisation using a Typhoon 8600 (Amersham Biosciences).

2.6 Western blot detection

2.6.1 Binding and detection of antibodies

Membranes were placed in blocking buffer (PBS, 5 % Milk, 0.1 % Tween), for 1 hour to overnight, on a rocker. This blocking step avoided the unspecific binding of antibodies. Membranes were rinsed in blocking buffer before placing in blocking buffer containing the primary antibody for 1 hour. Membranes were rinsed three times in blocking buffer for 15 minutes, before placing in blocking buffer containing the secondary antibody for 1 hour. In this thesis, all secondary antibodies were horse radish peroxidise conjugated. Membranes were rinsed twice in blocking buffer for 15 minutes, and once in PBS-T (PBS, 0.1 % Tween). The membranes were finally rinsed in PBS for a few seconds before applying the SuperSignal® West Pico Chemiluminescent Substrate. The substrate was applied to the membrane and placed in the dark for 5 minutes before exposing the membrane to an X-ray film (Kodak) for 5 seconds to overnight. X-ray films were visualised by developing in a Kodak M-35-M X-omat processor.

2.6.2 Stripping Western blots

To strip the nitrocellulose membranes of bound antibodies, membranes were placed in a container with 20 mls of Restore™ Western Blot Stripping Buffer (Pierce) and rocked for 30 minutes. Successful stripping was checked by applying SuperSignal® West Pico Chemiluminescent Substrate to the membrane and exposing it to an X-ray film (Kodak). X-ray films were visualised by developing in an X-omat (Kodak). Membranes were finally rinsed in PBS before being re-probed.

2.7 Polymerase chain reaction (PCR)

2.7.1 Standard PCR

PCRs were normally set up in volumes of 25 μl for diagnostic reactions and 50 μl for reactions intended to amplify DNA fragments for cloning or transformations. The amount of reagents used in 25 μl reactions were exactly half those used in the 50 μl reactions. The 50 μl reactions contained either 1 μl of Taq (ABGene, at $5U.μl^{-1}$) or Herculase (Stratagene, at $5U.μl^{-1}$) DNA polymerase, 5 μl of the manufacturer's 10 x reaction buffer, 2 μl of 10 mM dNTPs and 2 μl of forward and reverse oligonucleotide primers (5 mM). For Taq-based PCRs, $MgCl_2$ was typically added to a final concentration of 1.5 mM, although this was occasionally increased to improve efficiency. For Herculase-based PCRs, the reaction buffer provides 2.0 mM Mg^{++}, although this was also occasionally increased by addition of $MgCl_2$, up to a maximum concentration of 6 mM. In both reactions, dH_2O was added to a final volume of 50 μl. PCR was conducted either in Robocycler (Stratagene) or PCRSprint (Hybaid) machines. Reaction conditions were typically 95 ˚C for 5 minutes, followed by 30 cycles of 95 ˚C for 1 minute, 50-60 ˚C for 1 minute, and 72 ˚C for 1 minute per kb of expected product, and a final cycle of 72 ˚C for 10 minutes. PCR products were routinely purified using the Qiagen PCR Purification and Gel Extraction kits, following manufacturer's instructions.

A list of oligonucleotides used for PCRs are displayed in the appendix, and specific primers are referred to in the text.

2.7.1.1 PCR purification

PCR products were purified using the PCR purification kit (Qiagen). Five volumes of Buffer PB was added to one volume of pooled PCR samples and mixed. 750 μl of this sample was applied to a QIAquick spin column in a 2 ml collection tube and centrifuged at 16000 x g in a micro-centrifuge for 1 minute. The flow-through was discarded and the spin column re-used for the remaining sample. In order to wash the column, 750 μl of Buffer PE was added and centrifuged at 16000 x g in a micro-centrifuge for 1 minute. The flow-through was discarded and the column centrifuged again at 16000 x g in a micro-centrifuge for an additional minute. This step was performed in order to remove any residual ethanol. The DNA was eluted by placing the column in a clean eppendorf and adding 30 μl of dH_2O or EB buffer (10mM Tris-Cl, pH 8.5). The column was left to stand for 1 minute and then centrifuged at 16060 x g in a micro-centrifuge for 1 minute.

2.7.1.2 Gel extraction

PCR products were extracted from agarose gels using the gel extraction kit (Qiagen). DNA fragments to be purified were excised from the agarose using a scalpel and dissolved in 3 volumes of Buffer QG (*e.g.*, a gel fragment weighing 100 mg was dissolved in 300 µl of buffer) by incubation at 50 °C for 10 minutes. In order to increase the yield of DNA fragments between 500 bp and 4 kb, one gel volume of isopropanol was added to the solution and mixed. 750 µl of the sample was applied to a QIAquick spin column in a 2 ml collection tube and centrifuged at 16000 x g in a micro-centrifuge for 1 minute. The column was washed by addition of 750 µl of Buffer PE and centrifuged at 16000 x g in a micro-centrifuge for 1 minute. The flow-through was discarded and the column centrifuged again at 16000 x g in a micro-centrifuge for an additional minute to ensure that all the ethanol was removed. The DNA was eluted from the column by adding 30 µl of dH_20 or EB buffer and centrifuging at 16000 x g in a micro-centrifuge for 1 minute.

2.7.2 MVR-PCR

For each *T. brucei* strain or subspecies, the complete *BRCA2* ORF was initially PCR-amplified from genomic DNA with Taq DNA Polymerase (ABgene) and the primers *Tb BRCA2 for* and *Tb BRCA2 rev* to provide a substrate for MVR mapping. These PCR products were then used in 25 µl MVR PCR reactions, which contained 5 µM of the primers *TbBRCrepfor* and *TbBRCreprev*, 2.5 µl of 10 x Taq buffer and 3 mM $MgCl_2$ and 5 U of Taq DNA polymerase (ABgene). PCR was performed for 18, 21 or 28 cycles of 95 °C for 1 min, 55 °C for 1 min and 72 °C for 4 min, and the products separated by electrophoresis on a 1.5 % agarose gel.

2.7.3 Reverse transcriptase PCR (RT-PCR)

Total RNA was treated with DNAaseI to remove any genomic DNA contamination prior to cDNA preparation. To do this, 1 µg of RNA was incubated with 1 µl of DNAaseI (Invitrogen, $1U.µl^{-1}$) and 1 µl of 10 x DNAaseI buffer in a final reaction volume of 10 µl. The reaction was terminated by the addition of 1 µl 0.25 mM EDTA pH 8.0 and incubation at 65 °C for 20 minutes.

cDNA was prepared from DNase-treated RNA using the Superscript First-Strand Synthesis System for RT-PCR kit (Invitrogen), according to the manufacturer's instructions. 50 ng of random hexamers and 1 µl of dNTPs were added to 5 µl of DNAase-treated RNA and the mixture incubated at 65 °C for 5 minutes and on ice for 1 minute. 4 µl of 25 mM $MgCl_2$, 2 µl of 0.1 M DTT, 2 µl of 10 x RT buffer and 1 µl of RNaseOUT recombinant

ribonuclease inhibitor were added, and incubated for 2 minutes at 25 °C. 1μl of Superscript II reverse transcriptase (RT; 200 U.μl^{-1}) was then added, and the reaction incubated at 25 °C for 10 minutes, followed by 42 °C for 50 minutes. For each RT reaction, a duplicate reaction was set up using the same RNA, but without RT, thereby acting as a control for DNA contamination in downstream experiments. Following cDNA generation, RT was heat-inactivated at 70 °C for 15 minutes. Finally, 1 μl RNaseH (3.8 U.μl^{-1}) was added and the reaction incubated at 37 °C for 20 minutes to remove any remaining single-stranded RNA. cDNA prepared in this way was used directly in PCR reactions, with 1 μl of undiluted cDNA routinely acting as a substrate in 25 μl reaction volumes.

2.8 Restriction enzyme digestion of DNA

Routinely, restriction digestions were carried out in a final reaction volume of 30 μl, containing 1-10 μg of DNA, 3 μl of restriction enzyme (NEB at 10 or 20 U.μl^{-1}) and 3 μl of 10 x buffer (NEB) as recommended by the manufacturer. Digests were incubated at the appropriate temperature for the enzyme(s) for approximately 2 hours for plasmid DNA, or overnight for genomic DNA.

If larger quantities of digested DNA were required, the reactions were scaled up to a maximum of 50 μl per 1.5ml eppendorf and were subsequently phenol: chloroform extracted and ethanol precipitated (section 2.2.1.1).

2.9 Cloning of DNA fragments

2.9.1 Cloning using T4 DNA ligase

DNA fragments for cloning were prepared either by PCR-amplification, purification (section 2.7.1) and restriction digestion, or by restriction digestion from a plasmid (section 2.8). When vectors were restriction digested using a single enzyme, self-ligation was prevented by the treatment of calf intestinal phosphatase (CIP; Roche), which removes the 5' phosphate groups. To do this, 1 μl of CIP (10 U.μl^{-1}) was added to the restriction digestion reaction and incubated at 37 ˚C for 1 hour. After CIP treatment, vectors were purified by agarose gel extraction following electrophoresis using the Qiagen gel extraction kit according to manufacturer's instructions as described in section 2.7.1.2. Inserts for cloning, either derived by PCR or by plasmid digestion, were also purified by gel extraction following agarose gel electrophoresis.

Ligation of DNA fragments into a plasmid vector were carried out in a 20 μl reaction volume, containing 1 μl of T4 DNA ligase (400 U.μl^{-1}, New England Biolabs) and 2 μl of ligase buffer (New England Biolabs), and were incubated at room temperature for 4 hours at room temperature or 16 °C overnight. 2 μl of the 20 μl ligation reaction was used to transform 60-120 μl of *E. coli* XL-1 blue MRF' cells (section 2.10).

2.9.2 Cloning into the TOPO vector

Cloning DNA fragments into the TOPO TA vector (Invitrogen) occurs using the 3' single adenosine overhang that is present on all PCR products generated by Taq DNA polymerase. PCR products generated by Herculase DNA polymerase do not generate these 3' adenosine overhangs and therefore needed to be treated with the addition of 1μl of Taq DNA polymerase per 50 μl reaction and incubated at 72 °C for 10 minutes prior to TOPO TA cloning. For either Taq or Herculase PCRs, 0.5 - 4 μl of PCR product was incubated with 1 μl of salt solution (provided with the vector) and 1 μl TOPO TA vector, made up to a final reaction volume of 6 μl with dH$_2$O and incubated for 5 minutes at room temperature. 2 μl of this reaction was then used to transform 25 μl TOP10 F' *E. coli* cells (Invitrogen) (section 2.10).

2.10 Transformation of E.coli and plasmid retrieval

Transformation of XL-1 blue MRF' (Stratagene) *E. coli* cells was carried out using either heat shock or electroporation, whilst TOP10 F' (Invitrogen) *E. coli* cells was carried out using heat shock. Transformations by heat shock were performed by incubating 2 μl of the ligation reaction and 80 μl of cells on ice for 20 minutes. The cells were then heat-shocked at 42 °C for 45 seconds before transferring to ice for 2 minutes. Cells were allowed to recover before antibiotic selection by adding either 900 μl SOC (XL-1 blue MRF') or 250 μl SOC (TOP10 F') to the transformed cells and incubating at 37 °C for 1 hour. Since all plasmids used in this study encode ampicillin resistance, transformants were therefore selected by spreading 150 μl of recovered cells onto L-agar plates containing ampicillin at a final concentration of 100 μg.ml^{-1} (Sigma) and incubated overnight at 37 °C.

Transformations by electroporation were performed by incubating 2 μl of the ligation reaction and 40 μl of cells on ice for 5 minutes. The cells were then placed in a pre-chilled 0.1 cm gene pulser® cuvette (Bio-Rad) and electroporated at 1.2 kV using a Bio-Rad micro-pulser. Cells were allowed to recover before antibiotic selection by immediately adding 900 μl SOC to the transformed cells and incubating at 37 °C for 1 hour. Transformants were then selected as above.

2.10.1   Small scale plasmid retrieval

Single colonies from bacterial plates were picked and used to inoculate 3 mls of L-broth containing ampicillin (Sigma) at a final concentration of 100 µg.ml^{-1} and grown up overnight at 37 ˚C in a shaking incubator. Plasmids were purified from 1.5 ml of the overnight culture using the Qiagen miniprep kit®, following the manufacturer's instructions. Cells were pelleted by centrifugation at 16000 x g in a micro-centrifuge for 1 minute. The supernatant was discarded and cells were resuspended in 250 µl of Buffer P1. 250 µl of Buffer P2 was added to lyse the cells and the solution mixed by inverting the eppendorf tube 4-6 times. 350 µl of Buffer N3 was added to neutralise the solution and mixed as above, before centrifuging at 16000 x g in a micro-centrifuge for 10 minutes. The supernatant was applied to a QIAprep column in a collection tube, centrifuged at top speed in a micro-centrifuge for 1 minute, and the flow-through discarded. 750 µl of Buffer PE was added to the column and centrifuged at 16000 x g in a for 1 minute, the flow-through discarded, and the column centrifuged for an additional 1 minute to remove residual ethanol. Plasmid DNA was eluted from the column by addition of 50 µl of dH$_2$O or EB buffer to the column, which was placed in a clean eppendorf tube and centrifuged at 16000 x g for 1 minute.

2.10.2   Large scale plasmid retrieval

When larger amounts of DNA were required, plasmids were purified from 150 ml of an overnight culture using a maxi prep kit (Sigma), according to the manufacturer's instructions. Cells were pelleted by centrifugation at 5000 x g for 10 minutes, then resuspended in 12 mls of resuspension solution by pipetting or vortexing. Cells were lysed by the addition of 12 mls of lysis solution and mixed by inversion of the tube several times. The solution was neutralised by the addition of 12 mls of neutralisation solution, before the addition of 9 mls of binding solution. The tube of cells was inverted twice and immediately applied to the barrel of a filter syringe, and left for 5 minutes. During this incubation step, the binding column was prepared by adding 12 mls of column preparation solution to the column, which was subsequently centrifuged at 3000 x g for 2 minutes. Half of the cleared lysate was expelled through the filter syringe into the binding column, and centrifuged at 3000 x g for 2 minutes. The eluate was discarded, and this step was repeated with the other half of the lysate. The column was washed by the addition of 12mls of wash solution 1 and centrifuging at 3000 x g for 2 minutes. The eluate was discarded before applying 12 mls of wash solution 2 and centrifuging at 3000 x g for 5 minutes. The plasmid DNA was eluted by transferring the binding column to a clean 50

ml collection tube and applying 3 mls of elution solution or dH$_2$O, before centrifuging at 3000 x g for 5 minutes.

2.11 Microscopy.

2.11.1   DAPI staining

DAPI (4, 6-diamidino-2-phenylindole) (Vector Laboratories Inc.) stain binds to DNA and fluoresces under UV light, allowing the DNA content of fixed cells to be analysed.

10 mls of a *T. brucei* bloodstream form culture, grown to a density of 1-2 x 10^6 cells.ml^{-1}, or 5 mls of a *T. brucei* procyclic form culture, grown to a density of 5 x 10^6 cells.ml^{-1}, were centrifuged at room temperature for 10 minutes at 583 x g. The cells were washed twice in PBS before being resuspended in 1 ml PBS. 10 µl samples were spotted onto microscope slides (C.A.Hendley Ltd) and allowed to air dry. The cells were then fixed by soaking in methanol for 10 minutes at room temperature. The slides were allowed to air dry before placing 2 drops of vectashield with DAPI (Vector Laboratories Inc.) onto the slide, positioning a cover slip and sealing the slide with clear nail varnish (Boots 17). Differential interface contrast (DIC) was used to visualise intact cells and UV to visualise DAPI using a Zeiss Axioskop microscope.

2.11.2   Immunofluorescence

Two methods of fixation were utilised for performing immunofluorescence on trypanosome cells. The fixation methods utilised were either methanol or formaldehyde fixing. For methanol fixation, trypanosomes from culture were harvested by centrifugation at 583 x g for 10 minutes at room temperature. The pellet was washed 10 mls of PBS, resuspended in 1 ml PBS, centrifuged at 5000g for 10 minutes and resuspended in 40 µl of PBS. 10 µl of the resuspension was smeared across the slides and allowed to air dry. The slides were subsequently fixed by submersion in methanol for 10 minutes and allowed to air dry. The slides were re-hydrated by placing in PBS for 5 minutes, before blocking in PBS containing 1 % Tween-20 and 3 % BSA (PBS-T-BSA) for 10 minutes. The slides were drained of the blocking solution and transferred to a dark humid chamber before adding the primary antibody, diluted in PBS containing 1 % Tween-20 and 3 % BSA (PBS-T-BSA). The incubation occurred at room temperature for 90 minutes before the slides were washed three times with PBS-T-BSA for 5 minutes. The slides were then returned to the humid chamber where the secondary antibody diluted in PBS-T-BSA was added and left to incubate for 30 minutes. Following this incubation, the slides were

washed twice with PBS containing 1 % Tween-20 (PBS-T). The slides were drained of washing buffer and allowed to air dry at room temperature. Once dried, the slides were mounted by adding 2 drops of vectashield with DAPI (Vector Laboratories Inc.) onto the slide, positioning a cover slip and sealing the slide with clear nail varnish (Boots 17). Fluorescence microscopic analysis was performed using an Axioskop 2 microscope (Zeiss) and images obtained with Openlab software (Improvision).

For formaldehyde fixation, trypanosomes from culture were harvested by centrifugation at 583 x g for 10 minutes at room temperature. The pellet was resuspended in 1 ml of PBS and centrifuged at 5000 x g in a micro-centrifuge for 10 minutes. 900 µl of the supernatant was removed and the cells resuspended completely in the remaining 100 µl of PBS. 1 ml of 1 % (v/v) formaldehyde in PBS was added and the eppendorf inverted several times, before incubating at 4 °C for 1 hour. Following the incubation, the cells were centrifuged at 6000 x g for 1 minute. The pellet was washed twice in 1ml of chilled PBS, followed by a wash in 500 µl of chilled 1 % BSA diluted in dH20. The cells were resuspended in 30 µl of 1 % BSA/dH20 solution. 10 µl of the resuspension was smeared across the slides and allowed to air dry for 3 hours.

The slides were re-hydrated by placing in PBS for 5 minutes, before blocking the slides in 50 % Foetal Bovine Serum in PBS (FBS/PBS) for 15 minutes. The blocking solution was removed and the slides transferred to a dark humid chamber before adding the primary antibody diluted in 3 % FBS/PBS for 45 minutes. The slides were subsequently washed twice in PBS for 5 minutes. The slides were returned to the humid chamber where the secondary antibody, diluted in 3 % FBS/PBS was added to the slides and incubated for 45 minutes at room temperature. Following the incubation, the slides were washed twice in PBS for 5 minutes and allowed to air dry. Once dried, the slides were mounted by adding 2 drops of vectashield with DAPI (Vector Laboratories Inc.) onto the slide, positioning a cover slip and sealing the slide with clear nail varnish (Boots 17). Fluorescence microscopic analysis was performed as above.

CHAPTER 3

Identification of a putative BRCA2 homologue from *T. brucei*

3.1 Introduction

The maintenance of genomic stability is a well conserved process, which most likely occurs in all organisms, both eukaryotic and prokaryotic. Genomic stability is maintained via a number of different pathways, with homologous recombination being an essential process for providing error free repair (Pastink *et al.*, 2001). An array of proteins are involved in homologous recombination, with RAD51 being the central eukaryotic repair enzyme, performing the role of DNA strand exchange (Sonoda *et al.*, 2001). At least *in vivo*, a number of different factors mediate and regulate RAD51-catalysed strand exchange. Of these, five RAD51-related genes have been discovered to date in *T. brucei* (Proudfoot and McCulloch, 2005): DMC1, RAD51-3, RAD51-4, RAD51-5 and RAD51-6. RAD51-3 and RAD51-5 have been shown to function with RAD51 in homologous recombination (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005) in bloodstream stage *T. brucei*, whilst DMC1 appears not to act in this life cycle stage (Proudfoot and McCulloch, 2006).

BRCA2 has recently emerged as an important regulator of homologous recombination (Venkitaraman, 2002;Davies *et al.*, 2001), at least part of whose function is to sequester RAD51 until DNA damage occurs, when it transports the repair enzyme to the sites of damage. At least *in vivo*, it appears that in the absence of BRCA2, RAD51 may not be able to target these sites, thereby preventing homologous recombination from progressing. As a means to begin asking if *T. brucei* utilises the same repertoire of RAD51 regulators as other eukaryotes, it was decided to search the *T. brucei* genome for a homologue or homologues of BRCA2. This would develop further our understanding of homologous recombination in *T. brucei* and advance our understanding of VSG switching.

This chapter describes the identification of a single BRCA2 homologue in the trypanosomatids, and characterisation of the protein in terms of conserved functional motifs, genomic organisation and *T. brucei* life cycle expression.

3.2 Identification of BRCA2 in the trypanosomatids

The second hereditary breast cancer susceptibility gene, *BRCA2*, was localized in *H. sapiens* to chromosome 13q12-q13 in 1994 (Wooster *et al.*, 1994). Shortly after its localisation, it was identified as being the gene responsible for germline mutations in breast cancer families (Wooster *et al.*, 1995). In recent years, homologues of human BRCA2

have been uncovered in many different eukaryotes, including vertebrates, plants, fungi and nematodes (Bignell *et al.*, 1997;Siaud *et al.*, 2004;Kojic *et al.*, 2002;Martin *et al.*, 2005).

Putative trypanosomatid BRCA2 proteins were identified through BLAST searches of the *T. brucei, T. congolense, T. vivax, T. cruzi and L. major* genome databases using the Gene DB server (Sanger; http://www.genedb.org/). Initially, a BLASTp search was performed using *H. sapiens* BRCA2 (AAB07223) as the query protein sequence against the *T. brucei* database. This revealed a hypothetical gene (Tb927.1.640), situated on chromosome 1, encoding a predicted protein of 1648 amino acids, that showed substantially greater homology than any others. Though the overall sequence identity of this protein with human BRCA2 is low (see section 3.5), the likelihood that this is indeed a BRCA2 orthologue is strengthened by the prediction of a BRCA2 repeat region detected between residues 80 and 686, corresponding to fourteen putative BRC repeats. Our confidence in this protein being a legitimate orthologue of BRCA2 is further enhanced by the previous identification of this protein by Lo, T. *et al* (2003) and by Warren *et al.*, (2003) predicting fifteen BRC repeats in this *T. brucei* protein.

BLASTp searches were next performed using the putative *T. brucei* BRCA2 homologue protein as the query protein sequence against the *T. congolense, T. cruzi, T. vivax,* and *L. major* genome databases. The *T. congolense* database revealed a hypothetical gene (congo695a05.p1k_18), situated on chromosome 1, encoding a protein of 1179 amino acids. Here, the genome annotation of this protein predicted a BRCA2 repeat region between residues 80 and 211 with three BRC repeats. The *T. cruzi* database revealed a putative DNA repair gene, inferred from homology as BRCA2 (Tc00.1047053505999.40), encoding a protein of 1030 amino acids, though no BRC repeats were annotated. The *T. vivax* database revealed a hypothetical gene (tviv192h02.q1k_9) encoding a protein of 1179 amino acids, with a predicted BRCA2 repeat region between residues 62 and 96 containing one BRC repeat. Finally, the *L. major* database revealed a hypothetical gene (LmjF20.0060), situated on chromosome 20, encoding a protein of 1165 amino acids in which a BRCA2 repeat region was annotated between residues 106 and 178, predicting two BRC repeats. Together, these preliminary data suggest that each of the above trypanosomatids contains a likely BRCA2 homologue, though with variable numbers of BRC repeats predicted, and no other functional domains identified.

3.3 The genomic environments of the putative trypanosomatid BRCA2 genes

The genomic environments of the putative *BRCA2* genes were investigated by examining the surrounding sequences on Gene DB. Analysis of the genes predicted around the putative *T. brucei BRCA2* gene (Hall *et al.*, 2003;Berriman *et al.*, 2005) is shown in figure 3.1 and reveals a *T. brucei* anti-silencing gene, encoding ASF1-like protein, downstream of *BRCA2*. Further downstream is a phosphate repressible phosphate permease gene (Hall *et al.*, 2003;Berriman *et al.*, 2005). The surrounding sequence also contains a number of conserved hypothetical ORFs, which exhibit similarities to hypothetical proteins from other organisms and unlikely hypothetical ORFs, which are predicted to encode a protein of less than 150 amino acids. BLAST searches of nucleotide and protein databases retrieve only insignificant alignments. Analysis of the genomic environments surrounding the putative *BRCA2* gene from the other trypanosomatids reveals likely gene synteny with that of *T. brucei* (figure 3.1), since the 2 downstream conserved genes are found in similar positions in 3 of the 4 other genomes, and broadly similar ORF size and positioning is seen upstream and downstream of *BRCA2* in all. The only notable exceptions are in *T. cruzi,* which has a putative duplication of the phosphate repressible phosphate permease gene, and in *L. major*, which lacks the phosphate repressible phosphate permease gene in this region. This confirms that the genes identified in each trypanosomatids by BLAST searches with the putative *T. brucei* BRCA2 polypeptide are orthologues.
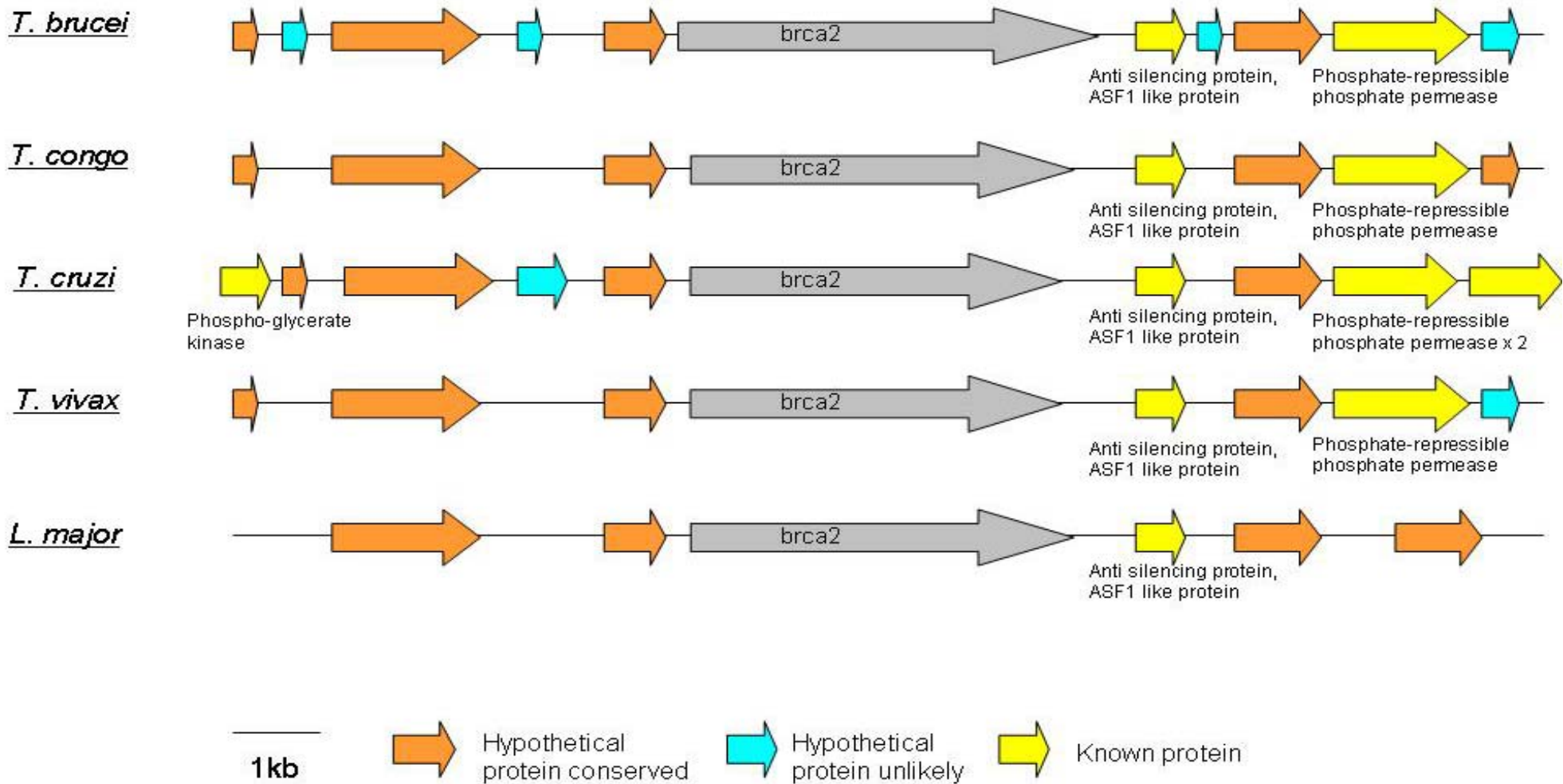
Figure 3.1 – The genomic environment of the *BRCA2* orthologues of the trypanosomatids. The putative *BRCA2* genes are displayed in grey, whilst genes encoding known proteins are shown in yellow, conserved hypothetical proteins, which exhibit similarities to hypothetical proteins from other organisms are shown in orange and unlikely hypothetical proteins, which are predicted to encode a protein of less than 150 amino acids are shown in blue (GeneDB).

3.4 Phylogenetic analysis

Phylogenetic analysis was carried out using the polypeptide sequences of BRCA2 homologues from a wide range of organisms. This analysis considered the proteins which have already been functionally examined and a number which remain uncharacterised. The functionally examined proteins were *Arabidopsis thaliana* (Siaud *et al.*, 2004)*, Caenorhabditis elegans* (Martin *et al.*, 2005)*, Canis familiaris* (Ochiai *et al.*, 2001)*, Felis catus* (Oonuma *et al.*, 2003)*, Gallus gallus* (Takata *et al.*, 2002)*, Homo sapiens* (Wooster *et al.*, 1994)*, Mus musculus* (Sharan and Bradley, 1997)*, and Ustilago maydis* (Kojic *et al.*, 2002)*. *The uncharacterised proteins came from *Drosophila melanogaster, Dictyostelium discoideum, Entamoeba histolytica, Giardia lamblia, Leishmania major, Toxoplasma gondii, Trichomonas vaginalis, Trypanosoma brucei, Trypanosoma congolense, Trypanosoma cruzi, Trypanosoma vivax and Plasmodium falciparum.* The uncharacterised proteins were obtained through searching the NCBI database (http://www.ncbi.nlm.nih.gov/) and a series of BLASTp searches on Gene DB (http://www.genedb.org/), Toxo DB (http://www.toxodb.org/) and Giardia DB (http://www.mbl.edu/Giardia/). The polypeptide sequences used in this analysis are presented in table 3.3 and accession numbers are indicated in the appendix.

The polypeptide sequences were compared by Clustal W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) to generate a phylogenetic tree, which was then visualised using Treeview (Page, 1996). The results from the phylogenetic analysis are shown in figure 3.2 and display that the putative BRCA2 polypeptides from the trypanosomatids form a discrete grouping, suggesting a level of conservation within the kinetoplastida.
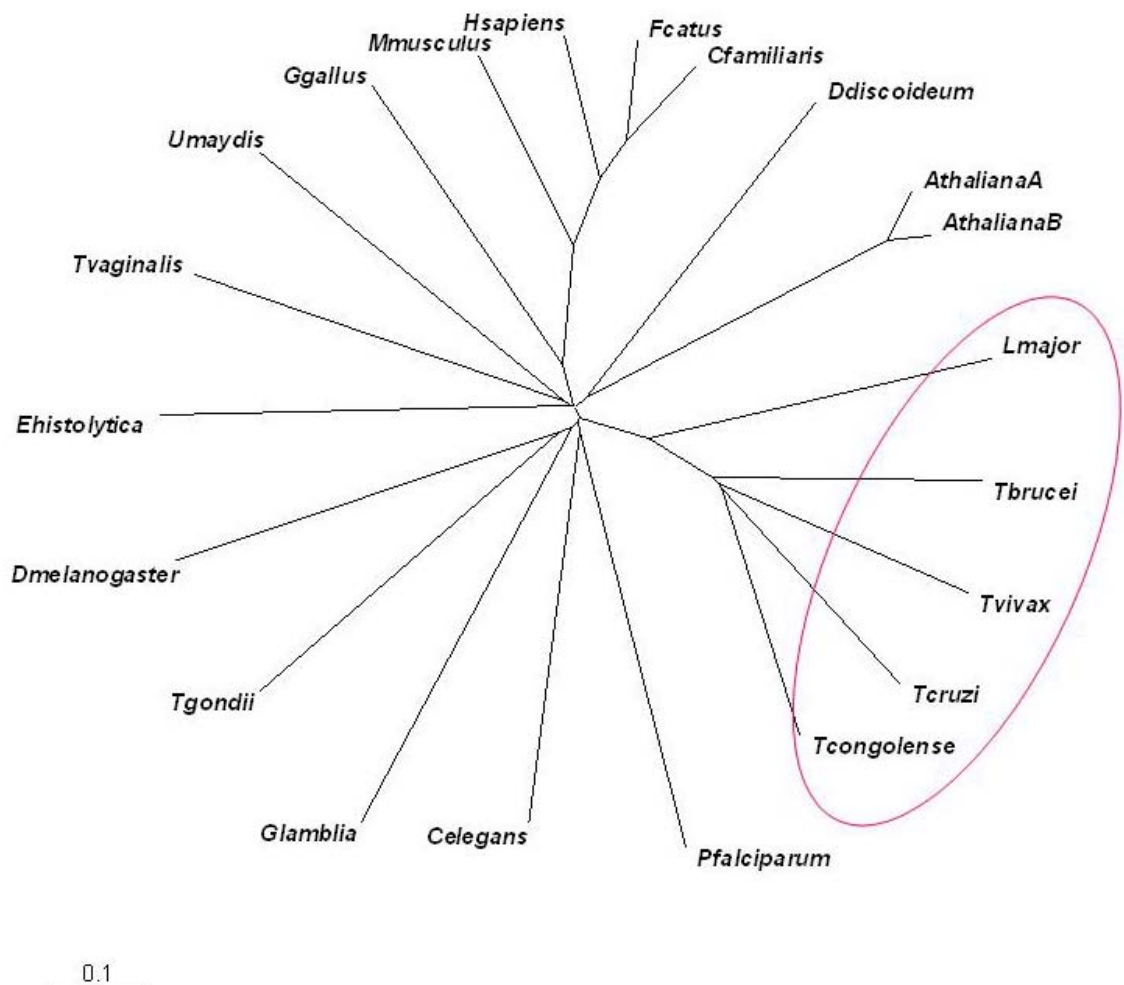
Figure 3.2 – Phylogenetic tree of BRCA2 proteins. The polypeptide sequences of BRCA2 from *Arabidopsis thaliana* (Athaliana)*, Caenorhabditis elegans* (Celegans)*, Canis familiaris* (Cfamliliaris)*, Drosophila melanogaster* (Dmelanogaster)*, Felis catus* (Fcatus)*, Gallus gallus* (Ggallus)*, Homo sapiens* (Hsapiens)*, Mus musculus* (Mmusculus)*, Dictyostelium discoideum* (Ddiscoideum)*, Ustilago maydis* (Umaydis)*, Entamoeba histolytica* (Ehistolytica)*, Giardia lamblia* (Glamblia)*, Leishmania major* (Lmajor)*, Toxoplasma gondii* (Tgondii)*, Trichomonas vaginalis* (Tvaginalis)*, Trypanosoma brucei* (Tbrucei)*, Trypanosoma congolense* (Tcongolense)*, Trypanosoma cruzi* (Tcruzi)*, Trypanosoma vivax* (Tvivax) and *Plasmodium falciparum (Pfalciparum)* were compared by Clustal W (Chenna *et al.*, 2003). The sequence comparison was then used to generate a phylogenetic tree and visualised using Treeview (Page, 1996). The red oval highlights the clustering of BRCA2 from the trypanosomatids.

3.5 Alignments of the putative T. brucei BRCA2 polypeptide with eukaryotic BRCA2 orthologues

A global multiple alignment of the putative *T. brucei* BRCA2 polypeptide with characterised BRCA2 orthologues from *G. gallus, H. sapiens, A. thaliana* and *U. maydis*, as well as the predicted orthologues from *T. cruzi* and *L. major* (section 3.2) was produced using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003). This was then visualised using the Boxshade server (http://www.ch.embnet.org/software/BOX_form.html), as shown in figure 3.3. The alignment shows that there is little sequence conservation observed throughout the BRCA2 polypeptides from these eukaryotes. This was confirmed by determining the sequence identities of BRCA2 from these eukaryotes by pair-wise comparisons, which were performed using AlignX (Vector NTI) and the percentage sequence identities calculated (see table 3.1); a graphical representation of this analysis is shown in figure 3.4. The putative *T. brucei* BRCA2 polypeptide shares between only 6.4 % and 11.7 % sequence identity with the BRCA2 proteins from *H. sapiens*, *G. gallus*, *A. thaliana* and *U. maydis*. Such low levels of homology are true also for the *T. cruzi* and *L. major* proteins when compared with these eukaryotes, indicating that this divergence is not peculiar to *T. brucei*. In fact, such low level global identity is true throughout, with only *H. sapiens* and *G. gallus* sharing substantially higher levels (30.1 %). The trypanosomatids proteins, as might be expected, are more closely related to each other than to other eukaryotes.

The results from the multiple alignment and pair-wise comparisons reveal little about the basis for the conservation of function between BRCA2 from different organisms. One of the reasons for this low level of conservation is likely to be due to the large differences in the size of the proteins (table 3.3). For example, the smallest trypanosomatid BRCA2 homologue, at 1030 amino acids from *T. cruzi*, is less than 1/3rd the size of the human protein, so homology between sequences would be difficult to observe. Indeed, even smaller proteins have been identified, the most extreme of which is around 10 % of the human BRCA2 size in *C. elegans*, but functions analogously (Martin *et al.*, 2005). For this reason, we decided to consider the conservation of functional domains of the protein, rather than the protein as a whole (section 3.7).

```
Tbrucei       1  ---------------------------------------------------------------------------------------------
Tcruzi        1  ---------------------------------------------------------------------------------------------
Lmajor        1  ---------------------------------------------------------------------------------------------
Ggallus       1  MAYKSGKRPTFFEVFKAHCSDSDLGPISLDWFEELSSEAPPYEPKLLGEPEGPIGWFDQT FKTPKAKSSTDSQLASTPLIFKEQNT-MPPFSSPGKELDQKKMETSRENLLSPSMAGRK
Hsapiens      1  MPIGSKERPTFFEIFKTRCNKADLGPISLNWFEELSSEAPPYNSEPAEESEHKNNNYEPN LFKTPQRKPSYNQLASTPIIFKEQGLTLPLYQSPVKELDKFKLDLGRNVPNSRHKSLRT
AthalianaA    1  ---------------------------------------------------------------------------------------------
Umaydis       1  ---------------------------------------------------------------------------------------------

Tbrucei       1  ---------------------------------------------------------------------------------------------
Tcruzi        1  ---------------------------------------------------------------------------------------------
Lmajor        1  ---------------------------------------------------------------------------------------------
Ggallus     119  T DQENQILASPHGICHNYTAASPAIVRNPCRTPQRSNIPGPYGSLFCTPKFLEIP-TPKRI SESLGAEVDPEMSWTSSLATPPTLGVTVIIARENDSISGAKQQDERAEIVLHNFLSE
Hsapiens    120  V KTKMDQADDVSCPLLNSCLSESPVVLQCTHVTPQRDKSVVCGSLFHTPKFVKGRQTPKHI SESLGAEVDPDMSWSSSLATPPTLSSTVLIVRNEEASETVFPHDTTAN--VKSYFSN
AthalianaA    1  ---------------------------------------------------------------------------------------------
Umaydis       1  ---------------------------------------------------------------------------------------------

Tbrucei       1  ---------------------------------------------------------------------------------------------
Tcruzi        1  ---------------------------------------------------------------------------------------------
Lmajor        1  ---------------------------------------------------------------------------------------------
Ggallus     236  DDG YTAKNDTSLLSIPETVKLNARDDIKDLESEVLDGLFGETNSFEDSFNLPAESSGILLSPR ALDAIEKCEIKIDEAQEKSDVLSEQHMRRKSTISQEVKAANWTEKSCCVEVKDSI
Hsapiens    236  HDE SLKKNDRFIASVTDSENTNQR-------EAASHGFGKTSGNSFKVNSCKDHIGKSMPNV LEDEVYETVVDTSEEDSFSLCFSKCRTKNLQKVRTSKTRKKIFHEANADECEKSK
AthalianaA    1  ---------------------------------------------------------------------------------------------
Umaydis       1  ---------------------------------------------------------------------------------------------

Tbrucei       1  ---------------------------------------------------------------------------------------------
Tcruzi        1  ---------------------------------------------------------------------------------------------
Lmajor        1  ---------------------------------------------------------------------------------------------
Ggallus     354  IQNTN EDIMDSKDNCLLGHEKELEYLRIAGNLQDNRTQKSSVNEKLVKDVLSSSSQWSQLNLSGL ECNSSGMSICSSPQSDSCREKGLERESVLMTKDDAVETSLLNTSGLRNAQELS
Hsapiens    346  NQ--- ----VKEKYSFVSEVEPNDTDPLDSNVAHQKPFESGSDKISKEVVPSLACEWSQLTLSGL N----GAQMEKIPLLHISSCDQNISEKDLLDTENKRKKDFLTSENSLPRISSL
AthalianaA    1  ---------------------------------------------------------------------------------------------
Umaydis       1  ---------------------------------------------------------------------------------------------

Tbrucei       1  ---------------------------------------------------------------------------------------------
Tcruzi        1  ---------------------------------------------------------------------------------------------
Lmajor        1  ---------------------------------------------------------------------------------------------
Ggallus     472  SASLSEN GSDTKISKNNPMSEITPVKPVCASPKLVKGYAHEDVSGMSFLNCSSFLIESTNVMEYSVV YNSTFSTHLKATSQSVVTDVLSHPLICSAASPDNCSDLHLRNSENTLRKSN
Hsapiens    453  PKSEKPL NEETVVNKRDEEQHLESHTDCILAVKQAISGTSPVASSFQGIKKSIFRIRESPKETFNAS FSGHMTDPNFKKETEASESGLEIHTVCSQKEDSLCPNLIDNGSWPATTTQN
AthalianaA    1  ---------------------------------------------------------------------------------------------
Umaydis       1  ---------------------------------------------------------------------------------------------
```

```
Tbrucei       1 ----------------------------------------------------------------------------------------------------
Tcruzi        1 ----------------------------------------------------------------------------------------------------
Lmajor        1 ----------------------------------------------------------------------------------------------------
Ggallus     590 FKSLN---M LSRLRKKSKRFIYTINNTLVYQEENVQKEVTSESPDNPVLTHLESDLHEFKDCQVATDGN QDCLLSAERQSNIKENNLNTLTIKVDIMDNSSDNSVNN-----RLKQEL
Hsapiens    571 SVALKNAGL ISTLKKKTNKFIYAIHDETFYKGKKIPKDQKSELINCSAQFEANAFEAPLTFANADSGLL HSSVKRSCSQNDSEEPTLSLTSSFGTILRKCSRNETCSNNTVISQDLDY
AthalianaA    1 ----------------------------------------------------------------------------------------------------
Umaydis       1 ----------------------------------------------------------------------------------------------------

Tbrucei       1 ----------------------------------------------------------------------------------------------------
Tcruzi        1 ----------------------------------------------------------------------------------------------------
Lmajor        1 ----------------------------------------------------------------------------------------------------
Ggallus     700 SESGKNAREYQ PATSFKCLKASHTESEDTTDCLNSGRISNIKHKVLTSAYLMARRHSRLFPEDCCL----- ---------------RKGKNDTYTVSNVNSRAAVPWSPKGQPPQSS
Hsapiens    689 KEAKCNKEKLQ LFITPEADSLSCLQEGQCENDPKSKKVSDIKEEVLAAACHPVQHSKVEYSDTDFQSQKSL LYDHENASTLILTPTSKDVLSNLVMISRGKESYKMSDKLKGNNYESD
AthalianaA    1 ----------------------------------------------------------------------------------------------------
Umaydis       1 ----------------------------------------------------------------------------------------------------

Tbrucei       1 ----------------------------------------------------------------------------------------------------
Tcruzi        1 ----------------------------------------------------------------------------------------------------
Lmajor        1 ----------------------------------------------------------------------------------------------------
Ggallus     797 PSCSDCLIDMHHG TAFVTNSKFNNTLSHIKFGMNRVSSNSCNKILADKRRASDQLSVAECREIVAPLGIN--- -------------CLENNSTSLKQRGKEDVDENQETLS-------
Hsapiens    807 VELTKNIPMEKNQ DVCALNENYKNVELLPPEKYMRVASPSRKVQFNQNTNLRVIQKNQEETTSISKITVNPDS EELFSDNENNFVFQVANERNNLALGNTKELHETDLTCVNEPIFKN
AthalianaA    1 ----------------------------------------------------------------------------------------------------
Umaydis       1 ----------------------------------------------------------------------------------------------------

Tbrucei       1 ----------------------------------------------------------------------------------------------------
Tcruzi        1 ----------------------------------------------------------------------------------------------------
Lmajor        1 ----------------------------------------------------------------------------------------------------
Ggallus     892 --------IKSSENP QAAAWNNESIEVAEEFLDCIDNSLNEVVSEEDRQVAPVYFNTKPIESLEHKGKSSGDLNA CSSSLSFGGFQTASNKQIKFSESSIAKGKMLFKDIENEFFEAS
Hsapiens    925 STMVLYGDTGDKQAT QVSIKKDLVYVLAEENKNSVKQHIKMTLGQDLKSDISLNIDKIPEKNNDYMNKWAGLLGP ISNHSFGGSFRTASNKEIKLSEHNIKKSKMFFKDIEEQYPTSL
AthalianaA    1 ----------------------------------------------------------------------------------------------------
Umaydis       1 ----------------------------------------------------------------------------------------------------

Tbrucei       1 ----------------------------------------------------------------------------------------------------
Tcruzi        1 ----------------------------------------------------------------------------------------------------
Lmajor        1 ----------------------------------------------------------------------------------------------------
Ggallus    1002 SMERVRNFSNRVQKENI FSSDLESKTGSTSSGLQTRCMQYIPRKVDLCKNSPRNQLSVQEPNQSLTASQEAEIAELS NILEETGSQFEFTQFRKQSNMIQSHIQQFGATNVENASEAG
Hsapiens    1043 ACVEIVNTLALDNQKKL S----KPQSINTVSAHLQSSVVVSDCKNSHITPQMLFSKQDFNSNHNLTPSQKAEITELS TILEESGSQFEFTQFRKPSYILQKSTFEVPE----------
AthalianaA    1 ----------------------------------------------------------------------------------------------------
Umaydis       1 ----------------------------------------------------------------------------------------------------
```

```
Tbrucei        1  ----------------------------------------------------------------------------------------------------
Tcruzi         1  ----------------------------------------------------------------------------------------------------
Lmajor         1  ----------------------------------------------------------------------------------------------------
Ggallus     1120  EDTNFYSTLKSENHVINDE YCSKSKNENECKMVEYEKEDTVVFHKNKREVTFTNLDRNESRISSHESCPVPLRDSFSNF VGFTSAGGKKINISKAALTRSAELFKDLDDDNFLFKSSG
Hsapiens    1147  ---NQMTILKTTSEECRDA DLHVIMNAPSIGQVDSSKQFEGTVEIKRKFAGLLKNDCNKSASG------YLTDENEVGF RGFYSAHGTKLNVSTEALQKAVKLFSDIENISEETSAEV
AthalianaA     1  ----------------------------------------------------------------------------------------------------
Umaydis        1  ----------------------------------------------------------------------------------------------------

Tbrucei        1  ----------------MSHKKGRQGSNSGARQNSDTPQRNRTKCRSDAPKRQRSRSGES VQGKSPLQERETRIQPRRDRTYGTEN----GQESTAVQGNSTDVPTLFVSAAGKPITVS
Tcruzi         1  ----------------------------------------------------------------------------------------------------
Lmajor         1  ----------------------------------------------------------------------------------------------------
Ggallus     1238  TNTRCCNSDERVSSNWNFLRC QTKEDEGGILCVPNIKSIGPISHHSEKKYAENISSPCEENTENWTEILSDNVDFSCTNGG YSASGMRNSPSSFKKPHQNCKNSDQF----LNQGNSE
Hsapiens    1256  HPISLSSSKCHDSVVSMFKIE NHNDKTVSEKNNKCQLILQNNIEMTTGTFVEEITENYKRNTENEDNKYTAASRNSHNLEF DGSDSSKNDTVCIHKDETDLLFTDQHNICLKLSGQFM
AthalianaA     1  --------------------------------------------------------------------------------------MSTASPSVAHAFPFGSADPLFDDD
Umaydis        1  --------------------------------------------------------------------------------------MSTASPSVAHAFPFGSADPLFDDD

Tbrucei       99  E SSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQ ESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARAR-MNTENGQESTAVQGNSTDV
Tcruzi         1  ----------------------------------------------------------------------------------------------------
Lmajor         1  ----------------------------------------------------------------------------------------------------
Ggallus     1352  VEGCLQEDTSYLICLGDNITSAE EHDLNVSDEMENLSPNQKEDRKQEDEHLLLNRQAADTDAVSISDSSFRSSLRDLNVQCGE RDTGVSEKSSKQKTNSVSVEGEDSTYKNLFVSESE
Hsapiens    1374  KEGNTQIKEDLSDLTFLEVAKAQ EACHGNTSNKEQLTATKTEQNIKDFETSDTFFQTASGKNISVAKESFNKIVNFFDQKPEE LHNFSLNSELHSDIRKNKMDILSYEETDIVKHKIL
AthalianaA     1  ----------------MSTWQLFPDSSGDGFRWEVAGRILQSVSDSTPTKALESTAPLPSMADLLIQGCSKLIAREEAMPGEIP-------------------------
Umaydis       25  IAATQQSILEELHTISEEALSANSDHSESHINSHIIDQSYGAETQGEHDGIHSDASSSGL SQLLMSRFASQQGAQLSIPASSEHNMEHSPAAPHIAE--------------------

Tbrucei      216  P-- -----TLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPI TVSESSLQVARARMN----------------------------TENGQEST
Tcruzi         1  ----------------------------------------------------------------------------------------------------
Lmajor         1  ----------------------------------------------------------------------------------------------------
Ggallus     1470  IK-IGSNQRHQVPSEQEMDVDKNKV KGTYLTGFCTASGKKITIADGFLAKAEEFFSENNVDLGKDDNDCFEDCLRKCNKSYVKDR DLCMDSTAHCDADVLNFKDKLIPQEPGDRLKQT
Hsapiens    1492  KESVPVGTGNQLVTFQGQPERDEKI KEPTLLGFHTASGKKVKIAKESLDKVKNLFDEKEQGTSEITSFSHQWAKTLKYREACKDL ELACETIEITAAPKCKEMQNSLNNDKNLVSIET
AthalianaA    69  - ------MFRTGLGKSVVLKESSIAKAKSILAEKVTYSDLRNTNCSIPQMRQVDTAETLP- --------------------------------------------------
Umaydis      122  -- --------RSGFEQEAPSPTPPIMADGSEITSQTADDGTNSNVVKITPLQADIEESVVTL ES-------------------------------------------------

Tbrucei      295  AVQGN STDVPTLFVSAAGKPITVSESSLQVARAR---------MNTENGQESTAVQGNSTDVPTL FVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPI
Tcruzi         1  ----------------------------------------------------------------------------------------------------
Lmajor         1  ----------------------------------------------------------------------------------------MKPIACPHCTF
Ggallus     1587  IEESPIIQA---VNHDSIKVGAFINVD EDCERNLAAPCANKEAYVRPGRSEVESLP---------VHGNNSLSRTLLLEDRKRFAER DVEYSATKRDNPESKPDSSLKCATSLHITKV
Hsapiens    1610  VVPPKLLSDNLCRQTENLKTSKSIFLK VKVHENVEKETAKSPATCYTNQSPYSVIENSALAFYTSCSRKTSVSQTSLLEAKKWLREG IFDGQPERINTADYVGNYLYENNSNSTIAEN
AthalianaA   122  -----------------------------------------MFRTASGKSVPLKESSIAKAM SILGSDKIIDSDNVLPRESGFGVSNSLFQTASNKKVN----------VSSAGLARA
Umaydis      176  ---- --------------------------------LPQRSDPQSATPISSSILAPTQ TLNTTPEPPDAAPSQMDASFELDHADLFDGIEPDAFDDIELSPPTRRHVALPLK
```

```
Tbrucei     404  TVSESSL QVARARMN--------------TENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQ VARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMN
Tcruzi        1  ----------------------------------------------------------------MGLEPDNVCAHCT----FINAPGR-----------VRCSMCFRN------
Lmajor       12  INPPSKVK--------------CGVCLRLLRKRERTVDDASPATPSRGRKGTPSPSSTP PEHDAP----ATGASQVRCDTTKSPMPLAAPESS----------------------
Ggallus    1693  SSHLADNSVPGGIIQTVSAEDSCKSNQSF LLPRGSVPRSTSPY-------LNCGNKEIDLKRLNEPCSNTDSFTNTVDNAHQEQSEFDL PEDETNLTCLQETSLNAESQKSDLKQVFS
Hsapiens   1728  DKNHLSEKQDTYLSNSSMSNSYSYHSDEV YNDSGYLSKNKLDSGIEPVLKNVEDQKNTSFSKVISNVKDANAYPQTVNEDICVEELVTS SSPCKNKNAAIKLSISNSNNFEVGPPAFR
AthalianaA  189  KALL GLEEDDLN----------------GFNHVNQSSSSSQQHGWSGLKTHEEFDATVVKHHSG TPGQYEDVSGKRSEVLNPSLKVPPTKFQTAGGKSLSVS---------------
Umaydis     252  EPPQSL AGLDSGLD--------------SDEFINDESPQLPPGSQTMSFLQPCFVGFQTGHGKQV KLSDKALEKARKLMMQLDDTTDLLPPAQTSQSSLH----------------

Tbrucei     507  TENGQESTA VQGNSTDV--------PTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDV PTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFV
Tcruzi       30  -------------------------------------IRKRIR-----EGAAEIGSELHL PHLSVSNQ------------------SEKQQE--------GCVATLFSTASGQPVVV
Lmajor       87  -- ----------------------------------STASAQAVQPPDVEAMAAAPPLV PTLFSTASGKPVTVRRESLQKVAERLGDLAAPD------MEARVPTLFETGRGKTV
Ggallus    1804  TAKGKAVSVSESALASIRQMFQTDCDASVKS EIETKSGTNQTAIAG-SSSFSIHAGGPGFATFLDTRKSEMNVAAPHFINGNGNLIENNHQ GANMFADADSVPGFQMQCFEQKSKLLG
Hsapiens   1846  IASGKIVCVSHETIKKVKDIFTDSFSKVIKE NNENKSKICQTKIMAGCYEALDDSEDILHNSLDNDECSTHSHKVFADIQSEEILQHNQNM SGLEKVSKISPCDVSLETSDICKCSIG
AthalianaA  276  ------ -----------------------------------AEALKRARNLLGDPELGSFFDDVA GGDQFFTPEKDERLSDIAINNGSANRGYIAHEEKTSNKHTPNSFVSPLWSSS
Umaydis     338  -------- -----------------------------------KRIHTTGSLPQALQSFGNAPSM LSTVTRTPMQEIVPKQRTAAINESEKCALAEEDKVASVQATSQVTALPAA

Tbrucei     617  SAAGKPITVSE SSLQVARARMNTENGQEST-------------------------------------AVQGN STDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAV------
Tcruzi       79  SE KSLQAARERLDADDAQIPL-------------------------------------THGNG DAAVATLFSTASGQPVVSEKSLQAARERVDADN---------------------
Lmajor      160  TVQK RSLVKAKASMDSLG------------------------------------ADGAP CTSAPALSATPGGAVVRPTVPMQPPAEHLRLRSLSDK----------------
Ggallus    1921  HFPVPDKQMEQSGPSGNLGFFSTASGKPVQLSE ESLKKARQLFSEMEGSHSSGLQDAHLL-EDVEKSTNHGEVFPREMQLLLPRGKENASTDK ISSPALGFSTASGKQVTISESAYQK
Hsapiens   1964  KLHKSVS-----SANTCGIFSTASGKSVQVSD ASLQNARQVFSEIEDSTKQVFSKVLFKSNEHSDQLTREENTAIRTPEHLISQKGFSYNVV NSSAFSGFSTASGKQVSILESSLHK
AthalianaA  352  KQFSSVNL ENLASGGNLIKKFD-------------------------------------AAVDE TDCALNATHGLSNNRSLASDMAVNNSKVNGFIP-----------------
Umaydis     410  QAPTTRRIEP HPFTTPKQTRNGRLP-------------------------------------VRQN LASPMRTPATAPGLRFTTPQPSKRISLGMLPRAEIG-----------

Tbrucei     693  ----------- ------------------------------------------------------------------------------------------------------------
Tcruzi      139  ---- -----------------------------------------------------------------------------------------------------------------
Lmajor      221  ------ --------------------------------------------------------------------------------------------------------------
Ggallus    2038  AMAILKEADGFLSSELGVTNELCEIKESGQHAEYL TGKVISESKTEKSCSEELDLKSIHPEKMKSLPSTHRVKITEYVPHSKRNSQSAPFKNSFE QEETRFFRKGELNLGIKTESESD
Hsapiens   2076  VKGVLEEFDLIRTEHSLHYSPTSRQNVSKILPRVD KRNPEHCVNSEMEKTCSKEFKLSNNLNVEGGSSENNHSIKVSPYLSQFQQDKQQLVLGTK VSLVENIHVLGKEQASPKNVKME
AthalianaA  412  ---------- ----------------------------------------------------------------------------------------------------------
Umaydis     475  ------------ --------------------------------------------------------------------------------------------------------

Tbrucei     693  --------------------------------------------------------------------------------------------------------------------QG
Tcruzi      139  -------------------------------------------------------------------------------------------------------------------
Lmajor      221  -------------------------------------------------------------------------------------------------------------------
Ggallus    2156  L----------------CSATSKAEINIFQTPKDYLK TEAVESAKAFMEDDLSDSGVQVKSAQSFGKMSDNFQNKPFGKRHLDEKDSHGEPPIKRQL LLEFEKMKIPPKSVKPLKSTP
Hsapiens   2194  IGKTETFSDVPVKTNIEVCSTYSKDSENYFETEAVEI AKAFMEDDELTDSKLPSHATHSLFTCPENEEMVLSNSRIGKRRGEPLILVGEPSIKRNLL NEFDRIIENQEKSLKASKSTP
AthalianaA  412  -------------------------------------------------------------------------------------------------------------------
Umaydis     475  -------------------------------------------------------------------------------------------------------------------
```

```
Tbrucei     695 NSTDVPTLFVSAAG KTVTVSESSLQVASANAASSAKPISGAGASLSKRTPRTHRKSASSSPLSSSKLARKPFVV PFAKNKGAVAKGVGEAVPSASHMPSSEGEGSEVGRTPRHLS---
Tcruzi      139 ---------------CATPCDKECGARGMS---------------DSSSAASPSFVGPRRGCFVV PFARSP-----------QPRQVVPPNRFQDPPARREMRSLS-------------
Lmajor      221 -------------DDTADTAPVLRTGAPRQTPLAEAPSTGLTTSTSERSLRALAPHRIGGQRRGFVP PQQRPT-----------AQPLAKLTHAMPLVKTTGLTILR----------
Ggallus    2258 DGIFKDRRKFMYHVPLKPVTCQPLGTTKERQEVRNPTLT LPDQDLKGFKSIPAVFQHCALRQSSSGASGLFTPHKAVAKDSEETRSLCKSGKAVKTFVP PFKTKLTLSTGEQDGSKRC
Hsapiens   2312 DGTIKDRRLFMHHVSLEPITCVPFRTTKERQEIQNPNFT APGQ--EFLSKSHLYEHLTLEKSSSNLAVSGHPFYQVSATRNEKMRHLITTGRPTKVFVP PFKTKSHFHRVEQCVRNIN
AthalianaA  412 ------------------RGRQPGRPADQPLVDITNRRDTAYAYNKQDSTQKKRLGKTVSVSPFKRPRIS SFKTPSKKHALQASSGISVVSCDTLTSKKVLSTRYPEKSPR-------
Umaydis     475 ----------------------GSSSTGSKRTLPRFVTPFKGGKRPRTEDLQDLASPLRRLDRAQAQSLSRAS PISPRQSFSMRQASSNISKGSAVFCMQHDGPRHKLAAVGRP-----

Tbrucei     810 ---------------- ----------------------FDIFTFRSLSMTVPPS------------------ --------------------------------IDEIVRG
Tcruzi      204 ------ ----------------------FDIFAFFSLPMSSLPS------------------ ------------------------------------SLEVVQSTFAFKCVGCS
Lmajor      304 -------- ----------------------FNPDACTCRSITPSPS------------------ ----------------------------------LSLITSLMFSFKGSD
Ggallus    2376 HSPIRNSVTEERELNQIPVEQNSAE--AQDHQSCILHAAVT DIENDNLITSNMMANLHCARDLQEMRIKKKYRQNISPQPGSLYVTKTSARNRISLKTAVE EETPSFHSTEKLYTYGV
Hsapiens   2428 LEENRQKQNIDGHGSDDSKNKINDNEIHQFNKNNSNQAAAV TFTKCEEEPLDLITSLQNARDIQDMRIKKKQRQRVFPQPGSLYLAKTSTLPRISLKAAVG GQVPSACSHKQLYTYGV
AthalianaA  505 ------------ ----------------------VYIKDFFGMHPTATTR------------------ --------------------------------------TMSEESN
Umaydis     567 ------------ ----------------------EYYSSMQMIAKG------------------ --------------------------------------VPDEV

Tbrucei     833 NFLFKQFGCSPELLKLLE IPAECEFIPSANFRKAMLTLGASPRGCPDAWCLQMLTSTLLKLRGLTLHIDPPLP--VFS VAHTLLHMCFKYNHEYVEGKRPALRLIAEG-DVQAASLVV
Tcruzi      237 PELLLLG LSAEDDSVPPRSFHTALVRLGACTSSCSENWCQQMLKSTLLKLRGLSLCCSPSLP--VFS VSHALLYLCFKYNREFVDGVRPALRVLTEG-DVPASSLLVVSVVCLSLEE
Lmajor      335 CGKVLAALLD VSG-GESVQPVHWHTMLLKLGASPKHCTIEWCRHALVSAMARVHEMTKAPSGAVPS-AFS PVTVLLCIMQYNAEMVNGGRPALRKMVEG-DISSASLVVLYMSSVR-
Ggallus    2492 SKHCIQVNSTNAESFQFLIEEFFSKEYLLAGNGIHLADGGWLI PTDEGKAGKKEFYRALCDTPGVDPKLITEAWVYNHYRWIVWKLAAMEVCFPHKFANRCLT PETVLLQLKYRYDLE
Hsapiens   2546 SKHCIKINSKNAESFQFHTEDYFGKESLWTGKGIQLADGGWLI PSNDGKAGKEEFYRALCDTPGVDPKLISRIWVYNHYRWIIWKLAAMECAFPKEFANRCLS PERVLLQLKYRYDTE
AthalianaA  528 QVMQINVYSVMSLL QTRLELKLFFKCWLSLVLPYNMHPEKKVCDRSFEACMWIVWKLACYDIYYPAKCRGNFLT ITNVLEELKYRYEREVNHGHCSAIKRILSGDAPASSMVVLCTSA
Umaydis     584 LVVLKDASQAARYAFE GPDSALLMQQQALEELHARGCSN---ADMPWVQNHWTTILWKLAAMVRLEPSSASN-RWS WNELIRQLLYRYEREVNLAQRSCLKRIQEHDSSAARPMVLMV

Tbrucei     948 VWVVSVSFEERLTP------ ------HTCTAVVSDGFYHVKVSLDIPLTNLVRNGTLRCGQKIVTCGARMLRR-DCCSPL ECKDE--VLLSINYNCTQPVGPSSPLGLYHT-CLPTLL
Tcruzi      352 RLAP------ ------HTSVGVISDGCYEIKVALDVPLTNLVREGILRCGHKLLVCGAKMLLK-NFCSPV DCRDD--VVLSINYNCTRPVDPTAALGFYQT-NPPVVSSAAVHPLGGL
Lmajor      449 EERSSP------ ------HMRIVTLSDGIYHLKVTCDIPLSNLIREGVLKPGQRMAVCGAKSLLH-RQCAPT ECEGQ--VVLSINYNCVRAVAQQTPLGVYHG-EPLPLPLSLVHPLG
Ggallus    2610 TDKSKRSAIKKITERDDAAGKTLVLCVSKVLSLNTAVSPSNSN-N NTEGEKAAAIIEVTDGWYGIRALLDPPLKAFLHRRRLTVGQKIIVHGAELIGSPNGCTPL EAPDS--LMLKIA
Hsapiens   2664 TDRSRRSAIKKIMERDDTAAKTLVLCVSDIISLSANISETSSNKT SSADTQKVAIIELTDGWYAVKAQLDPPLLAVLKNGRLTVGQKIILHGAELVGSPDACTPL EAPES--LMLKILS
AthalianaA  646 INPKTDNDSQE----A HCSDSCSNVKVELTDGWYSMNAALDVVLTKQLNAGKLFVGQKLRILGAGLSGWATPTSPL EAVISSTICLLNINGTYRAHWADRLGFCKE-IGVPLALNCI
Umaydis     698 SKILEEEIEVQSPSGE-- --IVSRICTILELSDGWYRILAQIDSVLTNACQRGRLRIGQKLAIMGATLDAHGEGKEVL SAYRMS--NLVLTANSVSLAPWDAKLGFAST--PFCASLR

Tbrucei    1050 PSAMDMLGGLVPCLKGRVERVL PPFFLEKTFKGAR--TGDTRGSTGGALKIVRSLLAQLSFQECMARGAVAPFEGKSD-RQL SRLTSFLLSCERQGDVLLQIWDDCGANCPAGDLEEH
Tcruzi      454 VPSIRGVVERIL PPIFVEQVVG-----NGDAVSTG---AKVVRNLLAQLKRLESVLREAAVPSESTNT-RRL VRLTSLVLTCEQKEDVMVQLWEDCAEGCGAESLEEDCGSEFPSEGS
Lmajor      551 GLVPAIEGVVARTL PSFFMSEEVTETSGTADGARQARNRVFKTVRNAHAQLQVTDRLRREAESRADGEAAPSKL LSRVTSLLIVKDNAEALVQQWETVDER--ALLADDDGGASLPVE
Ggallus    2725 ANSTRCARWYTKLGFHRDPRPFPLPLSSLYSEGGTVGCIDVVVQRTY PIQWMEKTSAG----SYVFRNSRAEEREAAKHAEDQQKKLEALFAKIQAEFEKHEE-RNC RRAPRSRIVTR
Hsapiens   2780 ANSTRPARWYTKLGFFPDPRPFPLPLSSLFSDGGNVGCVDVIIQRAY PIQWMEKTSSG----LYIFRNEREEEKEAAKYVEAQQKRLEALFTKIQEEFEEHEENTTK PYLP-SRALTR
AthalianaA  759 KCNGGPVPKTLAGTKRIY PILYKERLGEK----KSIVRSERTESRIIQLHNQRRSALVEGIMCEYQRGINGVHS---- -----------QNDTDSEEGAKIFKLLETAAEPEFLMAEM
Umaydis     808 SLTPEGGLISLMDVVITKVY PLAYVDVDKSN----AGAPRGEQEEAEQREAWLQRREDAMQQLELEAEAELG-------- ----------------RLYDLVEALNDLVGDAFLPS
```

```
Tbrucei     1165 S-CDFPPEGAEIVVFSVTPSRFRP GHPFQRTTVLYSRSPLRYSIVSPPRKGFVRQPLRSAEDVSPKTETGDAIDFAGLFVGTKS VDTVNSHIIVALNDGWKPGCVPASYFMIDVPHAT
Tcruzi       563 SVLIFALTPSRSRP SHPFQHAKVLYTRARLDYRVISSP-SGFVRPLRCSISDADLSMPAGAATDFAGLFLASAK NEAVNSFVIVMLKN---EETETPSFCIMDVPCATPVKEIAITMP
Lmajor       667 GSWVTLYAVNPAKSRT AAAPFTRAKLFFSSRKLYYVPSKNPPQHLRRIWMAATDNNSTTGVGDVADVCGLYVGSHR NEQGTFALLLCND---------TYALLQIPVPSAGRALSLP
Ggallus     2838 QQIHNLQDGAELYEAIQNAADPSYMEGYLSEDQLKALNAHKQLMNDKKQ TRIREEFKKAVESAEQEKHGFSKRDVSTVWKLFVVDYRKQEKHRGVILSIWRPLLDVCSL LKEGSRYRI
Hsapiens    2893 QQVRALQDGAELYEAVKNAADPAYLEGYFSEEQLRALNNHRQMLNDKKQ AQIQLEIRKAMESAEQKEQG-LSRDVTTVWKLRIVSYSKKE-KDSVILSIWRPSSDYSL LTEGKRYRI
AthalianaA   858 SPEQLRSFTTYKAKFEAAQQ MRKEKSVAETLEDAGLGERNVTPFMRIRLVGLTSLSYEGEHNPKEGIVTIWDPTERQRTE LTEGKIYMMKGLVP-------------INSDSEILYLH
Umaydis      896 IPDDPTGRLEAFANQLFDQLR- -------AQPNPASAVKERVVTAGHTSLVPWLHNLAKSALLQEDGIRGSSLSAELDRLCP PRKVREFRVVKFRD---------------ARLPPQ

Tbrucei     1282 GSKEIVLALPSIPFTPVIVQNASFIR -CAEDLGPDCIHVLANEYTK-VYSRPAEPLLRGVVESLGKIRGMAKSS-RPIIARSEELL RMR--------------TLSEEARADICRLS
Tcruzi       677 VASFTPVIVQNSSFIR FAVDDLGPDCLHVLANEFTK-VLQRPASSYLRTAISSLEIVREKAMGR-KSLMARAEEIL RLR--------------QLSVEARTDVRRYLGELNGRDVPL
Lmajor       776 LPTTERLSLVVLNATFIT GEDPVAGSDCCRLFANEYTA-VLQRSTQANLKGALETAAQLRGLVDAAPQKYAARKAEVF RCLDEGEPRGGLGLGSSTALTDRDNPGVPEVPLNVWRDAP
Ggallus     2956 CQLSTSQSKGRSDSTNVQLSATKKTRYLQLSVSQKMLQQIFFPRKALKFTS LLDPSYQPPCAEVDVVGVVI-SISRTGFSNMVYLSDESYNLVAIKIWADLRHFAIEDIVV RCS----
Hsapiens    3009 YHLATSKSKSKSERANLQLAATKKTQYQQLPVSDEILFQIYQPREPLHFSK FLDPDFQPSCSEVDLIGFVVSVVKKTGLAPFVYLSDECYNLLAIKFWIDLNEDIIKPHML IAA----
AthalianaA   963 ARGSSRWQPLS-PKDSENFQP FFNPRKPISLSNLGEIPLSS---EFDIAAYVVYVGNAYTDVLQKKQWVFVTDGS------ -------------------AQHSGEISNSLLAIS
Umaydis      990 PPATCLSTKTQQVGGSGATSKRKN AYARAVQLTVRDAAQLGDELREGRRFLVTNLVPMSKSAWRKPD--------------- ------------------------DQAEVFLS

Tbrucei     1382 RELVGGDELPNPAATAQP--------SP RYQLRQE----------------ASTPVEQSITVSETSAARTLSSEEEQVEDLRSSNVKA SPRR---HVFGNIVGFRLLKCQGSDKPECI
Tcruzi       778 AVASGSTS--------RL PYYMRAE---------------RKGPISKGVVTPTLHARSEMTEEQKSVFPP-SPAIHH GERS---HIFGNIIEMRLIRVFDSGRRESVLLPIGLAASD
Lmajor       893 LASPSAAATSPVLAQRDGRL PYYLRHDNGRVRAGNLTGVLLPSAALPTPAAPTVVQGPYSATTIHPGVREPLVQPANSVA GARS---RHYGNIADLMFLFDPRLNRRAWHPLTDPLQA
Ggallus     3069 ---------FIAASNLQWQSEFRSEIPVLLAGDLSAFSASPKENHLQEKFNEL RRMIENVDS-------------FCSDAESKLMNLLQRNCSLTPILPKRCGLECSSPSCNS GLYAE
Hsapiens    3123 -------------SNLQWRPESKSGLLTLFAGDFSVFSASPKEGHFQETFNKM KNTVENID------------ILCNEAENKLMHILHANDPKWSTPTKDCTSGPYTAQIIP GTGN-
AthalianaA  1050 FSTSFMDDSS------------- -----------------------VSHISHNLVGSVVGFCNLIKRAKDVTNEIWVAE AAEN---SVYFTNAEAAYSSHLKTSSAHIQTWAK
Umaydis     1065 TRRDTKWRPVA--------------- --------------------------------------------------- --------------------------------

Tbrucei     1473 EILGGRPSTLVSGSGK----FVVSPSDFSQ SLVYFEADIQFG-------ATAKQCAQTKVRSPSVLH-------------SLLEQCIPL KRACALTVDEIFADYYLARIKQLEDWQT
Tcruzi       868 VLTAG-----PAVPQETLLY DSICFDLVIQVG-------AVAEQQVKATMKSPAVLS-------------GMLEQRLAM RSVCAVAVDEGQVDYFMQRTKILETWQRAP--EESWWW
Lmajor      1008 TASAAVG-----------HGE GFRRAQLCWRLS-------ADSADDMTCRVEESSILG-------------TVLESVCPL QELCSVIADERHIDVSLARSERLVQWRRQDS-LSVW
Ggallus     3165 DRSISSKIETKHPSPLSASTPNTKLFPQGSAITPSSAVSS----ENHPRNSKKR KAVDFLSCIPAPPPLTPLCSIISPSLKKAFQPPRRLGSQHSKLSKETNPNAGCVTPSRRL REA
Hsapiens    3214 KLLMSSPNCEIYYQSPLSLCMAKRKSVSTPVSAQMTSKSCKGEKEIDDQKNCKKR RALDFLSRLPLPPPVSPICTFVSPAAQKAFQPPRSCG------------------TKY ETP
AthalianaA  1124 LSSSKSVIHELR------------- -------------------QRVLSIIGACKSPSC----------------------- --------------------------------
Umaydis          --------------------------------------------------- --------------------------------

Tbrucei     1566 PH--EECWWRLLTQSHVVEITSDVSGTPPEEL VGLQWLSNEWKMLLNILSGSLKHCLFM FSVEGS-EMVRATFIKEQCSVADLMRE------ --
Tcruzi       958 FLSLSHVVAGDSHP---PETEG TALLWLESEWRTLIDMLCEGLRDSLFKFSVDAAGEVT RAVFLKENCSLRELMKE------ --
Lmajor      1092 WRFFTDSRTLASPADLDG---ASS EHLWWLPAEWTEAMRTVSAKLQAAFFYFSLSGE-V LRHVRLISDCCSVAELPCD------ --
Ggallus     3279 VQLPDNDLVADEELAMINTQALINTVPEEKKMDYVNEDGTRATNLSGDTRATNLSRD TR APNLSGDLSSKNSSRSAKEANSSLKSSSEGADALQKDTEEPEGSLSIRRVLQRRKSRKCY
Hsapiens    3312 IKKKELNSPQMTPFKKFNEISLLESNSIADEELALINTQALLSGSTGEKQFISVSES TR TAPTSSEDYLRLKRRCTTSLIKEQESSQASTEECEKNKQDTITTKKYI----------
AthalianaA       --------------------------------------------------- --------------------------------------------------------------
Umaydis          --------------------------------------------------- --------------------------------------------------------------
```

Figure 3.3 – Global multiple alignment of the putative *T. brucei* BRCA2 polypeptide with a range of BRCA2 orthologues. Multiple sequence alignment of the putative *T. brucei* BRCA2 polypeptide with putative homologues from other eukaryotes: *T. cruzi*, *L. major*, *G. gallus*, *H. sapiens*, *A. thaliana* and *U. maydis*. Sequences were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and shaded using the BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html): residues that are identical in at least 50 % of the proteins are shaded in black and similarly conserved residues shaded in grey.

	T. brucei	T. cruzi	L. major	H. sapiens	G. gallus	A. thaliana	U. maydis
T. brucei	**100** 100	**26.6** 35.9	**17.2** 27.6	**6.4** 12.1	**6.8** 12.4	**11.7** 21.3	**8.8** 17.3
T. cruzi		**100** 100	**24.0** 38.5	**4.1** 7.7	**4.0** 7.6	**10.3** 20.6	**10.7** 21.8
L. major			**100** 100	**4.8** 8.6	**4.6** 8.6	**10.7** 19.9	**10.8** 20.8
H. sapiens				**100** 100	**30.1** 40.1	**6.3** 10.3	**5.1** 8.8
G. gallus					**100** 100	**5.7** 9.8	**5.4** 8.7
A. thaliana						**100** 100	**14.5** 25.6
U. maydis							**100** 100

Table 3.1 – Pair-wise comparison of the putative *T. brucei* BRCA2 polypeptide with a range of BRCA2 homologues. The full length putative *T. brucei* BRCA2 polypeptide was compared against polypeptide sequences from a range of eukaryotes. Pair-wise alignments were performed using AlignX (Vector NTI) and the percentage identities and similarities calculated. Percentage identities are displayed in bold.

Figure 3.4 – Graph displaying the % similarity and identity between *T. brucei* BRCA2 and BRCA2 from other organisms. Pair-wise alignments were performed as described in table 3.1 to compare the putative *T. brucei* BRCA2 polypeptide sequence with BRCA2 orthologues. Percentage identity is shown in blue and percentage similarity in maroon.

3.6 Alignments of the putative BRCA2 polypeptides in the trypanosomatids

A global multiple alignment of the putative *T. brucei* BRCA2 polypeptide with BRCA2 orthologues from *T. congolense*, *T. cruzi, T. vivax* and *L. major* was produced using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003). This was then visualised using the Boxshade server (http://www.ch.embnet.org/software/BOX_form.html), as shown in figure 3.5. The alignment shows that there is a high level of conservation observed throughout the BRCA2 polypeptides from trypanosomatids. This result was confirmed by determining the sequence identities of BRCA2 from these trypanosomatids by pair-wise comparisons. Pair-wise comparisons were performed using AlignX (Vector NTI), and the percentage sequence identities calculated (see table 3.2). A graphical representation of pair-wise alignments between *T. brucei* BRCA2 and orthologues from *T. congolense*, *T. cruzi, T. vivax* and *L. major* is shown in figure 3.6. The pair-wise comparisons show that the putative *T. brucei* BRCA2 polypeptide shares the highest level of sequence identity with the putative *T. congolense* BRCA2 polypeptide, with 32 % sequence identity and 41.9 % sequence similarity. This result demonstrates that the trypanosomatid proteins share a high level of homology throughout the protein, rather than just at the C terminus, which is what is observed when the *T. brucei* BRCA2 homologue is aligned with BRCA2 from other eukaryotes (see sections 3.7.1 and 3.7.2).

```
Tbrucei        1 MSHKKGRQGSNSGARQNSDTPQRNRTKCRSDAPKRQRSRSGESVQGKSPLQERETRIQPRRDRTYGTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAV
Tcongolense    1 ------------------------------------------------------------------------------------------------------------------------
Tvivax         1 ------------------------------------------------------------------------------------------------------------------------
Tcruzi         1 ------------------------------------------------------------------------------------------------------------------------
Lmajor         1 ------------------------------------------------------------------------------------------------------------------------


Tbrucei      121 QGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARAR
Tcongolense    1 ------------------------------------------------------------------------------------------------------------------------
Tvivax         1 ------------------------------------------------------------------------------------------------------------------------
Tcruzi         1 ------------------------------------------------------------------------------------------------------------------------
Lmajor         1 ------------------------------------------------------------------------------------------------------------------------


Tbrucei      241 MNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPIT
Tcongolense    1 -------------------------MVFSQKS------------------KGNICDVCSHVNKVD--QMRCDKCSHIIGRR--------SYSGSTRNSSKLSTPRKSNHG----
Tvivax         1 ------------------------------------------------------------------------------------------------------------------------
Tcruzi         1 ------------------------------------------------------------------------------------------------------------------------
Lmajor         1 -----------------------------------------------MKPIACPHCTFINPPSKVKCGVCLRLLRKRERTVDDASPATPSRGRKGTPSPSS---------


Tbrucei      361 VSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVP
Tcongolense   58 ---------------ANQSDEHEAAKG-----------------------------------AENDSSVERTTTGVATLFSTAAGKTVSVSESSLRAARMKLGQELCADGS----------
Tvivax         1 --------------------------------------------------MKQRQVGEKSPGAFHGRGSEELQVSYS----------
Tcruzi         1 ------------------------------------------------MGLEPDNVCAHCTFINAPGRVRCS-----------
Lmajor        55 -----------------------------------------------TPPEHDAPATGASQVRCDTTKSPMPLAAPESSTASAQAVQPPDVEAMA----------


Tbrucei      481 TLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARM----NTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTE
Tcongolense  119 -----------TLTEPPLQES------------------GPGVATLFSTAAGKTVSVSESSLRAARMKL----GQELCADGS--------------------TLTEPPLQES------
Tvivax        28 -----------EHNLNQR------------------------KRARASVCDLTETSGGSTEATI-----AQGDGQARK-----------
Tcruzi        25 -----------MCFRNIRKR-------------------IREGAAEIGSELHLPHLSVSN----QSEKQQEGC-----------
Lmajor       104 --------------AAPPLVPT----------------------LFSTASGKPVTVRRESLQKVAERLGDLAAPDMEAR----------


Tbrucei      597 NGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKTVTVSES
Tcongolense  178 ------------GPGVATLFSTAAGKTVSVSESSLRAARMKLGQELCADDEATVEN------------------------------TAQSESVGVPPPSTPVAGRRAKVSEP
Tvivax        66 ---------------MTMFSTAAGTKLSVSTDSLEKAKKKLEDIEWREEVQNNEAP--------------------------------LKQTALQSCASSVP------VNSDVKTSRV
Tcruzi        64 ----------------VATLFSTASGQPVVSEKSLQAARERLDADDAQIPLTHGNGD--------------------------------------AAVATLFSTASGQPVVSEK
Lmajor       147 ----------------VPTLFETGRGKTVTVQKRSLVKAKASMDSLGADGAPCTSAPA---------------------------------LSATPGCAAVRVPTVPMQPPAEHLR
```

```
Tbrucei      717  SLQVASANAASSAKPISGAGASLSKRTPRTHRKSASSSPLSSSKLARKPFVVPFAKNKGAVAKCVGEAVPSASHMPSSEGEGSEVGRTPRHLSFDIFTERSLSMTVPPSIDEIVRGNFLF
Tcongolense  248  PLRSVGEAFPPNDERLLLSGATPAGVTAVAGPNGQSKNVGACLRQLRKPFVVPFAKVAPDTGKAQEAERSSISNTLRAKRRFNDGVTSTKHISFDVCMRSMPLSSLPSIDDILNDSFSF
Tvivax       131  ETHIRANNVPSSMSARSSISDQRNASRLDTSKGSSSPSLTPSTSRPQRVFVVPYAKP-PLPCQGNGTEAKNHAQQQGTAAGGQRMWPLVRPLSFDISRFYSPPVSTVYSNDAILHGSFTF
Tcruzi       126  SLQAARERVDADN----CATPCDKECGARCMSDSSSAASPSFVGPRRGCFVVPFARS----PQPRQVVPNRFQDPPAR-------REMRSLSFDIFAFFSLPMSSLPSSLEVVQSTFAF
Lmajor       214  LRSLSDKDDTADTAPVLRTGAPRQTPLAEAPSTGLTTSTSERSLRALAPHRIGGQRRGFVPPQQRPTAQPLAKLTHAMP---LVKTTGLTILRFNPDACTCRSITPSPSLSLITSLMFSF

Tbrucei      837  KQFGCSPELLKLLEIPAECEFIPSANFRKAMLTLGASPRGCPDAWCLQMLTSTLLKLRGLILHIDPPLP-VFSVAHTLLHMCFKYNHEYVEGKRPALRLIAEGDVQAASLVVWVVSVSF
Tcongolense  368  KRLGCSLELLQLLEVQPGAEVVLPGSFRKALLSLGASAHWCTEEWCLQMMKSTLIKLRGLSLHCRPALP-VFSAEHTLLYMCFKYNHEFVDGQRPALRKVTEGDVPAGSVMVVFFVSLSK
Tvivax       250  SPFGCSEELLHLLEIPKNAESVPFTSFRKAMLKLGAVAHSCTEEWCMQMLASTLLKLRRLSLNCGMPLN-VFSVAHTLLYMCFKYNREFVDGSRPPLRLVTEDDVSAASLMVISLVSFSL
Tcruzi       231  KCVGCSPELLLLLGLSAEDDSVPPRSFHTALVRLGACTSSCSENWCQQMLKSTLLKLRGLSLCCSPSLP-VFSVSHALLYLCFKYNREFVDGVRPALRVITEGDVPASSLLVVSVVCLSL
Lmajor       331  KGSDCGKVLAALLDVSGG-ESVQPVHWHTMLLKLGASPKHCTIEWCRHALVSAMARVHFMTKAPSGAVPSAFSPVTVLLCIMQMYNAEMVNGGRPALRKMVEGDISSASLVVLYMSSVR-

Tbrucei      956  EERLTPHTCTAVVSDGFYHVKVSLDIPLTNLVRNGTLRCGQKIVTCGARMLRRDCCSPLECKDEVLLSINYNCTQPVGPSSPLGLYHTCLPTLLPSAMDMLGGLVPCLKGRVERVLPPFF
Tcongolense  487  DERLSPHTSVVSDGFYHVKVSLDVPLTNLTREGKLRGQKVMMCGAKMLKKDNCTPLECMGEVVLSLSYNCVKPVSDTALGLYHVCPPVVVPSAIDELGGMVPSLQGVVERVLPPFF
Tvivax       369  ADCLKPHTCTGTISDGCYHVKVAFDVPLTNMIRKGVIQCGQKLLVCGAKKLLRYSCSPLECKDEVVSIDYNCTKPVGPTPLGFYHINPPIVPLESIDTHGGLVPSIQGKVVRVLPPYF
Tcruzi       350  EERLAPHTSVGVISDGCYEIKVSLDVPLTNLIRGVRLRCGHKLLVCGAKMLLKNFCSPVDCRDVVLSINYNCTRPVDPTAALGFYQTNPPVVSSAAVHPLGGLVPSIRGVVERILPPIF
Lmajor       449  EERSSPHMRIVTLSDGIYLKVTCDIPLSNLIREGVLPGFRMAVCGAKSLLHRQCAPTECEGQVVLSINYNCVRAVAQQTPLGVYHGEPLPLPLSLVHPLGGLVPAIEGVVARTLPSFF

Tbrucei      1076 LEKTFKGA--RTGDTRGSTGGALKIVRSLLAQLSFQECMARGAVAPFE-GKSDRQLSRLTSFLLSCERQGDVLLQIWDDCGANCPAGDLEEHS-CDFPPEGAEIVVFSVTPSRFRPGHPF
Tcongolense  607  IEQPFKRT--RDTNARDGGSGGGKVVRNLLAQLKFHDRTSRGGNEEMC-EAGKRQLMRVTSFSLTCMHKGDVVLQLWDDFSNGCATEALEEDT-CAYPAEGSIIVVFALTPSRFRPSHPF
Tvivax       489  IQSSFT----NDGSTHGHRAGGTKVVRNMLAQLKSMEAS-RYAKSDDE-ASSHQRLSRVSSLVLTCSQKEDLLLQFWEDCGESCTAGSLEEYE-STFPPEGATITVFALTPSRSRPAHPF
Tcruzi       470  VEQVVG--------NGDAVSTGAKVVRNLLAQLKRLESVLREAAVPSE-STNTRRLVRLTSLVLTCEQKEDVMVQLWEDCAEGCGAESLEEDCGSEFPSEGSSVLIFALTPSRSRPSHPF
Lmajor       569  MSEEVTETSGTADGARQARNRVFKTVRNAHAQLQVTDRLRREAESRADGEAAPSKLLSRVISLLIVKDNAEALVQQWETVDERALLADDDGGA--SLPVEGSWVTLYAVNPAKSRTAAAP

Tbrucei      1192 QRTTVLYSRSPLRYSIVSPPRKGFVRQPLRSAEDVSPKTETGDAIDFAGLFVGTKSVDTVNSHIIVALNDGWKPGC--VPASYFMIDVPHATGSKEIVLALPSIPFTPVIVQNASFIRCA
Tcongolense  723  HHAKVLYSRSPLDYRVLSPPDKGGTRLPERAVDSVAAEIVTGGPVDVAGLFVATATVSNFNSHVIAMLLDDWCPGK--SSASYCVIDAPLATASKEITLAMPSAPFTPVIVQNASFIKRT
Tvivax       602  QQAKALHAKARLEYRTISSAREGDRREPCRSVKDVDLYTPAGVAMDFAGIFVKSARIDTVGSFVFVILEDGWATDLNTASQSYCLMDIPHDTPSKEIVLPTP-APFTPVVIQNASFIRIA
Tcruzi       581  QHAKVLYTRARLDYRVISSP-SGFVRPLRCSISDADLSMPAGAATDFAGLFLASAKNEAVNSFVIVMLKNEETETP-----SFCIMDVPCATPVKEIAITMPVASFTPVVQNSSFIRFA
Lmajor       687  FTRAKLFFSSRKLYYVPSKNPPQHLRRIWMAATDNNSTTGVGDVADVCGLYVGSHRNEQGTFALLLLCND-----------TYALLQIPVPSAGRAISLPLPTTERLSLVVLNATFLTGE

Tbrucei      1310 ED-LGPDCIHVLANEYTKVYSRPAEPLLRGVVESLGKIRGMAKSS-RPIIARSEELLRMR--------------TLSEEARADICRLSRELVGGDELPNPAATAQP--------SPRYQ
Tcongolense  841  ED-LGNDCIHTLANEFTRVLTRPATPALRSVVDSLEQLRHTAKTC-RVITSRCEELLRLR--------------TLSEEAQSDVRRLACESVGIDVVTITSSDTST--------SRLPY
Tvivax       721  HEGFGSDCAHALANEFTQVLQRPSAPFLRSVIGALEKLREKAKLT-ISISARAEELLRLR--------------GLSGEAQRDVHQFSDGLIN-DEVSITLSTRPS--------RVPYY
Tcruzi       695  VDDLGPDCLHVLANEFTKVLQRPASSYLRTALSSLELVREKAMGR-KSLMARAEEILRLR--------------QLSVEARTDVRRYLGEINGRDVPLAVASGSTS--------RLPYY
Lmajor       796  DPVAGSDCCRLFANEYTAVLQRSTQANLKGALETAAQLRGLVDAAPQKYAARKAEVFRCLDEGEPRGGLGLGSSTALTDRDNPGVPEVPLNVWRDAPLASPSAAATSPVLAQRDGRLPYY
```

```
Tbrucei      1405 LRQEA-STPVEQSITVSETSAARTLSSEEEQVEDLRSSNVKASP-RRHVFGNIVGFRLLKCQGSDKPECIEILGGRPSTLVSGSGKFVVSPSDFSQSL------VYFEADIQFGATAKQC
Tcongolense   936 YLREK-EVPFKQGVTLPKEKVTVSSSTARESLVNDTPMPLAAAP-LRHLFGNVVGFRLIKCHDSGRTERIDLFAVYVAGSSSESVSSDEPPAEDPQRS------SHFEIDVQFGATTQQR
Tvivax        816 MREGA-TTSAKQGVVIPCTNGSSPKEGKGAQKQGVATALLQQHGGRHHLFGNITELRLVRCYNIGKSESINLLKRSNGCSSLKQFGTDITVTADVQFS------SHFEIEIQFGAGEEQK
Tcruzi        791 MRAER-KGPISKGVVIPTLHAR--SEMTEEQKSVFPPSPAIHHGERSHIFGNIIEMRLIRVFDSGRRESVLLPIG-LAASDVLTAGPAVPQETLLYDS------ICFDLVIQVGAVAEQQ
Lmajor        916 LRHDNGRVRAGNLTGVLLPSAALPIPAAPTVVQGPYSATTIHPGVREPLVQPANSVAGARSRHYGNIADLMFLFDPRLNRRAWHPLTDPLQATASAAVGHGEGFRRAQLCWRLSADSADD
                  ----------------

Tbrucei      1517 AQTKVRSPSVLHSLLEQCIPLKRACALTVDEIFADYYLARIKQLEDWQTP-HEECWWRLLTQSHVVE---ITSDVSGTPPEELVGLQWLSNEWKMLINILSGSLKHCLFMFSVE-GSEMV
Tcongolense  1048 LTAQINVPRVLHALLEPRITLQRACALAVDEVIPEYSICRAKQLEKWRGA-REECWWRFLTRSCVVS---ARSHDLEAVLEPGDAVEWLASEWDILLEILSSEIRNCLFKFSVE-NGELT
Tvivax        929 KLVKLKNPCLLDALLERRVTLQVACSMAVDEECLDFVLRRNKLLKESQLP-PREQWWYLLTRSCIVNRDMSRSPTTLSQVATATAVEWLANEWEVLLGILITDALEDCLFKFSVNMEEELT
Tcruzi        901 VKATMKSPAVLSGMLEQRLAMRSVCAVAVDEGQVDYFMQRTKILETWQRA-PEESWWWFLSLSHVVAG------DSHPPETEGTALLWLESEWRTLIDMLCEGLRDSLFKFSVDAAGEVT
Lmajor       1036 MTCRVEESSILGTVLESVCPLQELCSVIADERHIDVSLARSERLVQWRRQDSLSVWWRFFTDSRTLAS------PADLDGASSEHLWWLPAEWTEAMRTVSAKLQAAFFYFSLS-GEVLR
                        --

Tbrucei      1632 RATFIKEQCSVADLMRE----
Tcongolense  1163 RAVFIREHCNIVNLMRK----
Tvivax       1048 QAVFIRENCSILELMQEESSV
Tcruzi       1014 RAVFLKENCSLRELMKE----
Lmajor       1149 HVRLISDCCSVAELPCD----
```

Figure 3.5 – Global multiple alignment of the putative trypanosomatid BRCA2 polypeptides. Multiple sequence alignment of the putative *T. brucei* BRCA2 polypeptide with homologues of BRCA2 from the other trypanosomatids: *T. congolense*, *T. vivax*, *T. cruzi* and *L. major*. Sequences were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and shaded using the BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html): residues that are identical in at least 50 % of the proteins are shaded in black and similarly conserved residues shaded in grey. The coloured dashes below the aligned sequence is representative of the domains highlighted in figures 3.9 and 3.10. Orange dashes: alpha helical domain; blue dashes: OB domains; lilac dashes: tower domain; red dashes: CDK phosphorylation domain.

	T. brucei	*T. congo*	*T. cruzi*	*T. vivax*	*L. major*
T. brucei	**100**	**32.0**	**26.6**	**26.2**	**17.2**
	100	41.9	35.9	35.2	27.6
T. congo		**100**	**39.4**	**38.1**	**25.3**
		100	52.9	49.4	36.2
T. cruzi			**100**	**41.6**	**24.0**
			100	56.0	38.5
T. vivax				**100**	**23.2**
				100	35.2
L. major					**100**
					100

Table 3.2 – Pair-wise comparison of the putative trypanosomatid BRCA2 polypeptides. The full length putative *T. brucei* BRCA2 polypeptide was compared against polypeptide sequences from putative Brca2 homologues from *T. congolense, T. cruzi, T. vivax* and *L. major*. Pair-wise alignments were performed using AlignX (Vector NTI) and the percentage identities and similarities calculated. The percentage identities are displayed in bold.
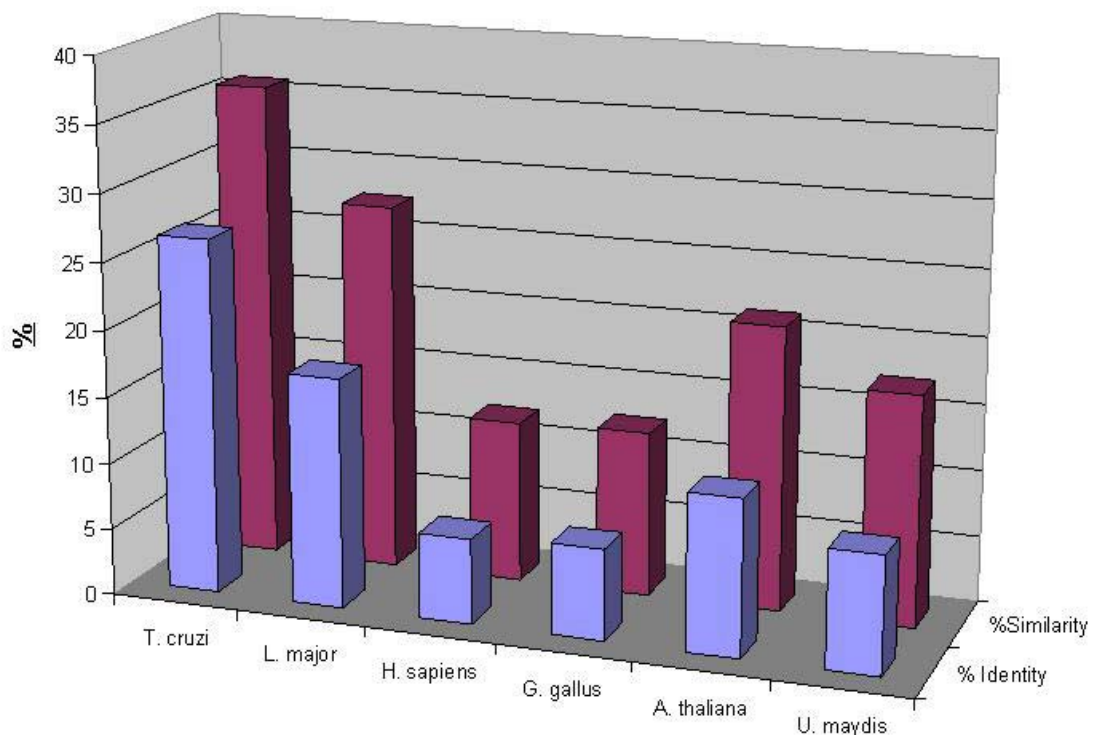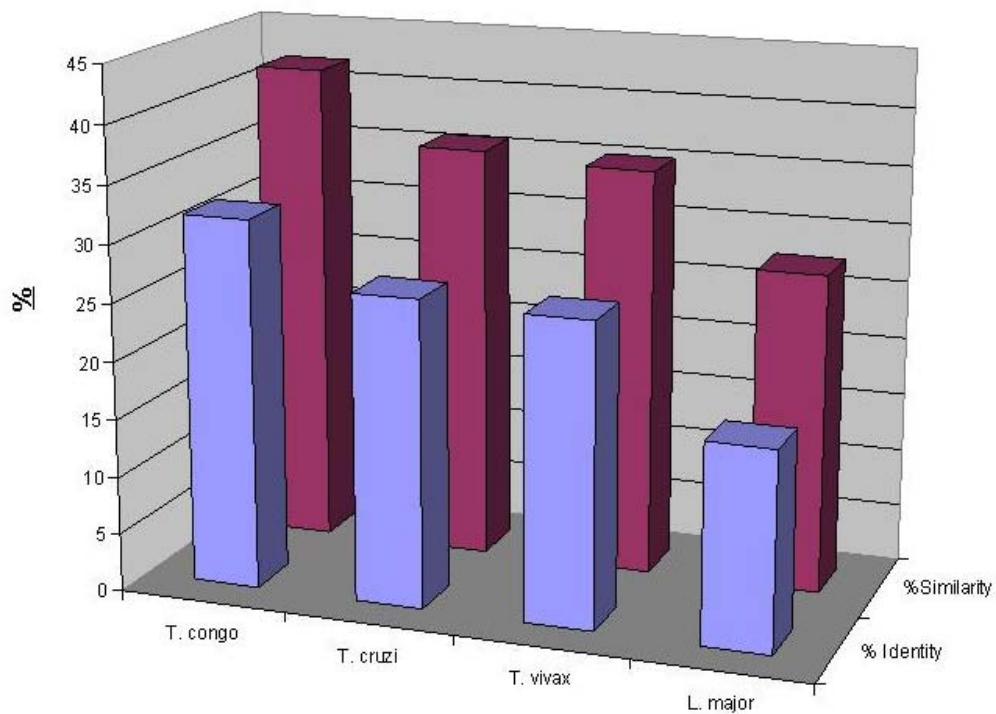

Figure 3.6 – Graph displaying the % similarity and identity between *T. brucei* BRCA2 and Brca2 from the trypanosomatids. Pair-wise alignments were performed as described in Table 3.4 to compare the putative *T. brucei* BRCA2 polypeptide sequence with Brca2 from *T. congolense, T. cruzi, T. vivax* and *L. major*. Percentage identity is shown in blue and percentage similarity in maroon.

3.7 Identifying the domains of T. brucei BRCA2

The human BRCA2 protein is a large multi-domain protein, composed of 3418 amino acids (Wooster *et al.*, 1995;Tavtigian *et al.*, 1996) and, as with many other DNA repair proteins it is localised to the nucleus (Bertwistle *et al.*, 1997). Sequence comparisons have proved to be relatively uninformative, due to the lack of homology with proteins of known function. However, crystallographic and functional studies have revealed a number of different domains in the protein. These are considered in turn below.

3.7.1 BRC repeats

In the human *BRCA2* gene, the large central exon 11 encodes eight sequence motifs, which have been termed the BRC repeats (Bork *et al.*, 1996). Each of these repeats, BRC1-BRC8, is composed of approximately 30 amino acids, and 6 out of the 8 motifs have been shown to interact with RAD51 (Wong *et al.*, 1997;Chen *et al.*, 1998b;Marmorstein *et al.*, 1998). A crystal structure of human BRC4 complexed with RAD51 has been solved (Pellegrini *et al.*, 2002), and shows that the RAD51 C terminal domain forms a mixed α-β-fold. This fold contains two loops, L1 (aa 230-236) and L2 (aa 269-292), which have been suggested to form DNA binding sites (Story *et al.*, 1992;Voloshin *et al.*, 1996). This structure had provided useful information in allowing BRC repeats to be predicted in different BRCA2 homologues (see figure 3.8) (Lo *et al.*, 2003). From this work, several critical residues for the interaction between a BRC repeat and RAD51 have been presented, constituting a BRC sequence fingerprint. This has further allowed the number of BRC repeats, and their functionality to be predicted, in many BRCA2 homologues. To examine this, BRCA2 proteins were identified from a range of organisms (used in the phylogenetic analysis described in section 3.4) and tabulated for their size and predicted number of BRC repeats (table 3.3).

Taxon	Organism	Size (amino acids)	No. of BRC repeats
Metazoa	*Homo Sapiens*	3418	8
Metazoa	*Canis familiaris*	3446	8
Metazoa	*Felis catus*	3372	8
Metazoa	*Mus musculus*	3328	8
Metazoa	*Gallus gallus*	3397	8
Metazoa	*Drosophila melanogaster*	947	3
Metazoa	*Caenorhabditis elegans*	394	1
Viridiplantae	*Arabidopsis thaliana-a*	1150	4
Viridiplantae	*Arabidopsis thaliana-b*	1155	4
Mycetozoa	*Dictyostelium discoideum*	1623	1
Fungi	*Ustilago maydis*	1075	1
Apicomplexa	*Toxoplasma gondii*	2741	8
Apicomplexa	*Plasmodium falciparum*	2668	6
Euglenozoa	*Leishmania major*	1165	2
Euglenozoa	*Trypanosoma cruzi*	1030	2
Euglenozoa	*Trypanosoma congolense*	1179	3
Euglenozoa	*Trypanosoma vivax*	1068	1
Euglenozoa	*Trypanosoma brucei*	1648	15
Diplomonadida	*Giardia lamblia*	1105	1
Parabasala	*Trichomonas vaginalis*	1664	14
Entamoebidae	*Entamoeba histolytica*	719	1

Table 3.3 – BRCA2 proteins from the eukaryotes used in the phylogenetic analysis. The sizes of the proteins and the number of predicted BRC repeats are indicated. The number of predicted BRC repeats were either inferred from the protein annotations in the sequence databases, were identified by Lo *et al*. (2003) or were identified manually in this work, using the BRC sequence fingerprint (Lo *et al.*, 2003).

It appears that a factor common to all BRCA2 homologues, is the presence of at least one BRC repeat motif (Lo *et al.*, 2003). For instance, BRCA2 from *U. maydis* and *C. elegans*, both of which are known to function (Kojic *et al.*, 2002;Martin *et al.*, 2005), contain one BRC repeat, and the predicted proteins from *D. discoideum, E. histolytica, G. lamblia* and *T. vivax* have the same. All the other BRCA2 proteins have multiple BRC repeats, though the predicted number in *T. brucei* appears to be exceptional since the majority of unicellular organisms were predicted to possess only one or two BRC repeats. In general terms, it appears that the simpler the organism, the smaller the number of BRC repeats (figure 3.9 summarises this). To illustrate this, 12 unicellular organisms are catalogued in table 3.2, and 8 have between 1 and 3 predicted BRC repeats. The exceptions to this are *T. brucei*, with 15 predicted repeats, *T. vaginalis*, with 14 predicted repeats and *P. falciparum* and *T. gondii*, each apicomplexans, with 6 and 8 predicted repeats, respectively. Conversely, 8 multicellular organisms are catalogued, of which 7 have 3 or more BRC repeats: each vertebrate has 8, *A. thaliana* has 4 and the insect *D. melanogaster* has 3. Again, there is an exception, with *C. elegans* having only a single BRC repeat. It seems possible that developmental complexity has generally selected for increased numbers of BRC repeats, perhaps due to more complex demands for homologous recombination

control in different tissues. It is also possible that the number of repeats correlates broadly with genome size, perhaps again with greater need to control homologous recombination. *T. vaginalis* provides an interesting example, as it is unusual amongst protozoans in having a genome estimated as ~100 Mb (Lyons and Carlton, 2004), around 3-4 fold larger than other protists examined, and 14 predicted BRC repeats.

Nevertheless, the BRC repeat number and arrangement in *T. brucei*, if correct, appears truly unusual. Though *P. falciparum* and *T. gondii* both have rather large numbers of BRC repeats, this appears to be true generally of apicomplexans, since the BRCA2 homologue in *Cryptosporidium hominis* (Chro.80593) is also predicted to contain 8 BRC repeats (www.cryptodb.org). In contrast, the closest relatives of *T. brucei; T. cruzi* and *L. major* in the kinetoplastids and even *T. congolense* and *T. vivax* (each belonging to the salivarian clade of the genus Trypanosoma), have far fewer repeats. Furthermore, 14 of the 15 (figure 3.7) predicted *T. brucei* BRC repeats are identical in sequence and all the repeats are present in a tandem array, separated by exactly 20 amino acids. In all other organisms with multiple repeats (with the exception of the 3 BRC repeats in *T. congolense*) they are found unevenly distributed in the polypeptide (figure 3.9) and have degenerated in sequence outside of the predicted functional residues (inferred from work by Lo *et al*., 2003). The most likely explanation for this is that *T. brucei* BRCA2 has undergone a recent BRC expansion, unique to this lineage.

To examine in more detail the potential functionality of the trypanosomatids BRC repeats, the sequence of those present in BRCA2 from *H. sapiens, T. brucei, T. congolense, T. cruzi, T. vivax* and *L. major* were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and visualised using the Boxshade server (http://www.ch.embnet.org/software/BOX_form.html) (figure 3.8) (Pair-wise comparisons are shown in appendix 7). This alignment displays a high level of conservation within the BRC repeat motifs and shows that the critical residues are conserved amongst the trypanosomatids compared with those in *H. sapiens*. Furthermore, it predicts that each repeat is likely to be capable of functioning in binding RAD51, as all repeats in each trypanosomatid retain the key residues identified by Lo *et al* (2003). Notably, this appears to include the 15th degenerate repeat in *T. brucei* BRCA2. For each protein, therefore, there is no bio-informatic evidence that some of the BRC repeats do not bind RAD51, as have been shown experimentally for *H. sapiens* BRC5 and BRC6 (Wong *et al.*, 1997;Chen *et al.*, 1998b;Chen *et al.*, 1999a).

MSHKKGRQGSNSGARQNSDTPQRNRTKCRSDAPKRQRSRSGESVQGKSPLQERETRIQPR

RDRTYGTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAV
82   116

QGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAG
126   160   170

KPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARAR
204   214

MNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNS
248   258   292

TDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPIT
302   336   346

VSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTE
380   390

NGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVP
424   434   468   478

TLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSES
512   522

SLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQE
556   566   600

STAVQGNSTDVPTLFVSAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFV
610   644   654

SAAGKPITVSESSLQVARARMNTENGQESTAVQGNSTDVPTLFVSAAGKTVTVSESSLQV
688   698

ASANAASSAKPISGAGASLSKRTPRTHRKSASSSPLSSSKLARKPFVVPFAKNKGAVAKG
732

VGEAVPSASHMPSSEGEGSEVGRTPRHLSFDIFTFRSLSMTVPPSIDEIVRGNFLFKQFG
CSPELLKLLEIPAECEFIPSANFRKAMLTLGASPRGCPDAWCLQMLTSTLLKLRGLTLHI
DPPLPVFSVAHTLLHMCFKYNHEYVEGKRPALRLIAEGDVQAASLVVVWVVSVSFEERLT
PHTCTAVVSDGFYHVKVSLDIPLTNLVRNGTLRCGQKIVTCGARMLRRDCCSPLECKDEV
LLSINYNCTQPVGPSSPLGLYHTCLPTLLPSAMDMLGGLVPCLKGRVERVLPPFFLEKTF
KGARTGDTRGSTGGALKIVRSLLAQLSFQECMARGAVAPFEGKSDRQLSRLTSFLLSCER
QGDVLLQIWDDCGANCPAGDLEEHSCDFPPEGAEIVVFSVTPSRFRPGHPFQRTTVLYSR
SPLRYSIVSPPRKGFVRQPLRSAEDVSPKTETGDAIDFAGLFVGTKSVDTVNSHIIVALN
DGWKPGCVPASYFMIDVPHATGSKEIVLALPSIPFTPVIVQNASFIRCAEDLGPDCIHVL
ANEYTKVYSRPAEPLLRGVVESLGKIRGMAKSSRPIIARSEELLRMRTLSEEARADICRL
SRELVGGDELPNPAATAQPSPRYQLRQEASTPVEQSITVSETSAARTLSSEEEQVEDLRS
SNVKASPRRHVFGNIVGFRLLKCQGSDKPECIEILGGRPSTLVSGSGKFVVSPSDFSQSL
VYFEADIQFGATAKQCAQTKVRSPSVLHSLLEQCIPLKRACALTVDEIFADYYLARIKQL
EDWQTPHEECWWRLLTQSHVVEITSDVSGTPPEELVGLQWLSNEWKMLLNILSGSLKHCL
FMFSVEGSEMVRATFIKEQCSVADLMRE

Figure 3.7 – Protein sequence of *T. brucei* BRCA2 with the BRC repeats highlighted. The fourteen identical BRC repeats are highlighted in red, whilst the fifteenth BRC repeat is highlighted in pink. Residue numbers are indicated.

```
                              FOTASG+OIOIOOOSIOOiOOIIO-
                                SGO           I
T.brucei1-14     1 -DVPTLFVSAAGKPITVSESSLQVARARMNTENGQE
T.brucei15       1 -DVPTLFVSAAGKTVTVSESSLQVASANAASSAKPI
T.congolense1-3  1 -GVATLFSTAAGKTVSVSESSLRAARMKLGQELCAD
T.vivax          1 -RKMTMFSTAAGTKLSVSTDSLEKAKKKLEDIEWRE
T.cruzi1         1 -CVATLFSTASGQPVVVSEKSLQAARERLDADDAQI
T.cruzi2         1 -AVATLFSTASGQPVVVSEKSLQAARERVDADNCAT
L.major1         1 -LVPTLFSTASGKPVTVRRESLQKVAERLGDIAAPD
L.major2         1 -RVPTLFETGRGKTVTVQKRSLVKAKASMDSLGADG
H.sapiens1       1 SFGG-SFRTASNKEIKLSEHNIKKSKMFFKDIEEQY
H.sapiens2       1 -VGFRGFYSAHGTKLNVSTEALQKAVKLFSDIENIS
H.sapiens3       1 -TSDTFFQTASGKNISVAKESFNKIVNFFDQKPEEL
H.sapiens4       1 -PTLLGFHTASGKKVKIAKESLDKVKNLFDEKEQGT
H.sapiens5       1 -NSALAFYTSCSRKTSVSQTSLLEAKKWLREGIFDG
H.sapiens6       1 -VGPPAFRIASGKIVCVSHETIKKVKDIFTDSFSKV
H.sapiens7       1 -NTCGIFSTASGKSVQVSDASLQNARQVFSEIEDST
H.sapiens8       1 -SAFSGFSTASGKQVSILESSLHKVKGVLEEFDLIR
```

Figure 3.8 – Multiple sequence alignment of the BRC repeat from trypanosomatids and humans. The polypeptide sequences of the BRC repeats were taken from *T. brucei, T. congolense, T. vivax*, *T. cruzi* and *H. sapiens*. The BRC repeat sequences highlighted in red indicate the BRC repeats in *H. sapiens* BRCA2 that do not bind RAD51. Sequences were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and shaded using the BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html): residues that are identical in greater than 50 % of sequences are shaded in black and similarly conserved residues shaded in grey. The structure based sequence fingerprint (Lo *et al.*, 2003) for the BRC repeat, with eight critical residues in red is indicated above the alignment. (O – polar; I – hydrophobic; i – small hydrophobic; (+) – positively charged; (-) – negatively charged.)
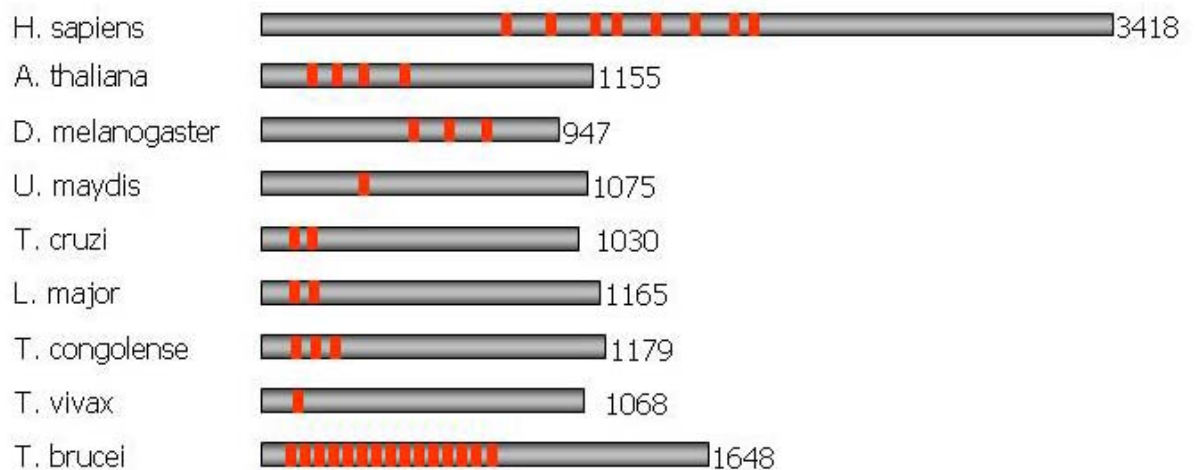
Figure 3.9 – Representation of the number of BRC repeats in BRCA2 proteins from trypanosomatids and other eukaryotes. BRC repeat motifs are displayed as red blocks, and their position within the BRCA2 polypeptides are shown. Protein sizes in amino acid residues are indicated.

3.7.2 DNA/DSS1 binding domains

The COOH-terminal region of BRCA2 corresponds to the greatest conserved region of the protein across orthologues from dog, mouse, rat and chicken (Yang *et al.*, 2002). This region is thought to have an important role for the function of BRCA2 since 27 % of the tumour derived mis-sense mutations in the breast cancer information core (BIC) database exist in this region (Szabo *et al.*, 2000). This region has been shown to interact with DSS1, a protein which is absent or mutated in split-hand split-foot syndrome (Marston *et al.*, 1999;Crackower *et al.*, 1996). Co-expression of DSS1 with the C terminal domain of *H. sapiens* BRCA2, allowed the structure to be crystallised and the X ray structure determined (Yang *et al.*, 2002). This structure revealed multiple domains similar to ssDNA and dsDNA binding motifs and was subsequently named the BRCA2DBD (DNA/DSS1 binding domain).

The BRCA2DBD has been discovered to contain five domains (see figure 3.10), four of which are arranged linearly, and one that protrudes out. The first domain is the alpha helical domain, with 190 residues consisting mainly of alpha helices. Following this are 3 structurally homologous domains containing oligonucleotide/oligosaccharide binding (OB) folds (OB1, OB2 and OB3). A tower domain is inserted into OB2, which contains a helix-turn-helix (HTH) motif that is similar to DNA binding domains of the bacterial site specific recombinases (Yang and Steitz, 1995;Feng *et al.*, 1994) and protrudes away from the OB fold. DSS1 interacts with the alpha helical domain, OB1 and OB2, characterised by hydrophobic interactions and a large number of acidic DSS1 residues which interact with basic residues on BRCA2 (figure 3.26).

The three OB folds of BRCA2 have been shown to have a high level of similarity to the OB folds of the ssDNA-binding protein RPA (Bochkarev *et al.*, 1999). RPA contains four OB folds (DBD-A, DBD-B, DBD-C and DBD-D), two of which bind to ssDNA with high affinity (DBD-A and DBD-B) (Bochkareva *et al.*, 2002). OB2 and OB3 in *H. sapiens* BRCA2 have been found to have the highest level of structural similarity to DBD-A and DBD-B and have therefore been attributed to binding to ssDNA. This was confirmed by the crystal structures of BRCA2DBD-DSS1-ssDNA displaying that ssDNA binds in the OB2-OB3 channel in a uniform manner. This result was supported by evidence from native gel electrophoretic mobility shift assays (EMSAs), which found OB2 and OB3 had a high affinity for ssDNA.
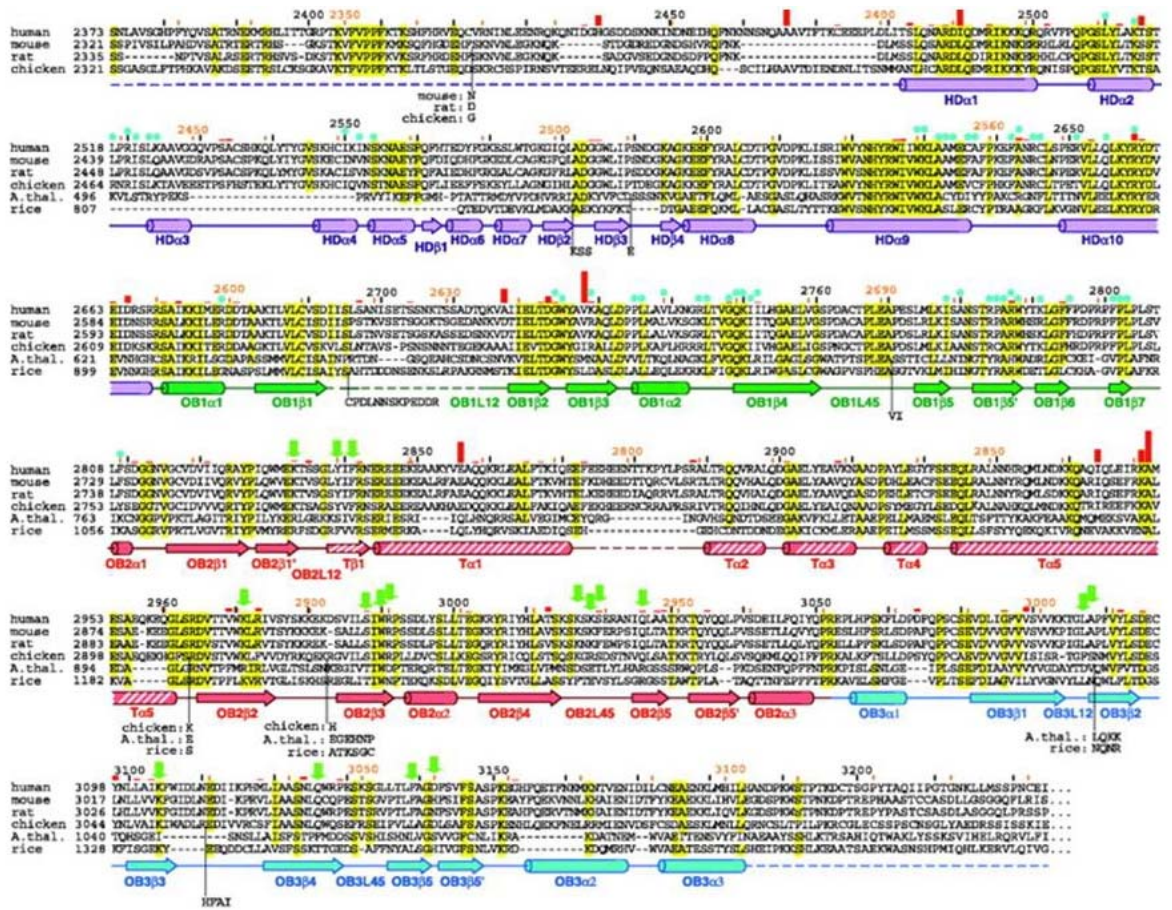
Figure 3.10 – Structures of the conserved BRCA2 COOH-terminal domain (figure taken from Yang *et al.*, 2002). Sequence alignment of the C terminal domains of BRCA2 from human, mouse, rat, chicken, arabidopsis (A. thal) and rice. Secondary-structure elements below the sequence are coloured in magenta for the helical domain (HD), green for OB1, red for OB2, hatched-red for the Tower insertion in OB2, and blue for OB3. Black dashed lines indicate gaps in the alignment. Insertions in orthologues are dropped below the sequence. Residues identical in five or more orthologues are highlighted in yellow, DSS1-interacting residues are indicated by blue dots, and ssDNA contacting residues by green arrows. Figure taken from *Yang et al*., 2002.

The tower region has been implicated to interact with dsDNA (Yang *et al.*, 2002), since a fourth DNA binding domain was predicted to exist due to the formation of fast complexes. It was unclear whether this was ssDNA or dsDNA binding, but the authors implicated dsDNA to be the likely candidate due to the tower domain containing 3 helix bundles, and most HTH motifs containing 3 helix bundle domains recognise dsDNA (Feng *et al.*, 1994;Yang and Steitz, 1995).

When Lo *et al.* (2003) identified a number of BRCA2 homologues, they noted that two features were common to the identified proteins; the BRC repeats and putative nuclear localisation signals. However, these authors noted that the different domains of the DSS1/DNA binding domains are conserved to different extents in different organisms, and it was predicted that in *T. brucei, L. major, E. cuniculi* and *U. maydis*, OB3 was completely absent, whilst no DBD was detectable in *D. melanogaster, C. elegans* or in the Plasmodium species.

The level of sequence homology was firstly investigated between BRCA2 from *T. brucei* and *H. sapiens* using AlignX (Vector NTI). The percentage sequence identity was calculated and a graphical representation generated displaying the level of sequence similarity for each amino acid residue (figure 3.11). This shows the lack of any identity between N termini of the two proteins. Instead, it highlights the fact that the homology that is present between the two BRCA2 homologues is limited to the C termini, indeed, at the DBD. This is consistent with the finding that this region was noted as being the best conserved across dog, mouse, rat and chicken orthologues (Yang *et al.*, 2002).

To examine this further, a global multiple sequence alignment of the region predicted to encompass the DBD of the *T. brucei* BRCA2 polypeptide with the DBD of BRCA2 orthologues from *T. cruzi, L. major, G. gallus, H. sapiens, A. thaliana* and *U. maydis* was produced using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003). This was then visualised using the Boxshade server (http://www.ch.embnet.org/software/BOX_form.html), as shown in figure 3.12. The DBDs were identified in *H. sapiens, G. gallus* and *A. thaliana* from the work produced by Yang *et al*., 2002. These domains allowed the identification of the potential DBDs in *T. brucei, T. cruzi, L. major,* and *U. maydis* through the previous alignments. The alignment of the DBD of these homologues predicts the presence of an alpha helical domain, OB1, OB2, OB3 and the tower domain in *T. brucei, T. cruzi* and *L. major*, and perhaps also in *U. maydis*. It should be noted that the greatest conservation is found in OB1 and the α helical domain. It is likely that this reflects conserved binding of DSS1. Despite the previous prediction that *T. brucei* BRCA2 does not contain an OB3 domain (Lo *et al.*, 2003), this alignment suggests that although the level of homology in this region is low, there is a recognisable conservation of numerous residues. The *U. maydis* protein does, however, appear to be truncated relative to the other BRCA2 proteins in at least the putative OB3 domain. Upstream, this alignment suggests that the α helical region is longer in the vertebrate proteins than any of the other eukaryotes. Nevertheless, despite the fact that we can predict that all domains appear to be present, biochemical studies will need to be undertaken to test this.

3.7.3 C terminal RAD51 binding domain

RAD51 binding has been described to occur at the BRC repeats in the centre of *H. sapiens* BRCA2, mapped to exon 11 (Bork *et al.*, 1996;Bignell *et al.*, 1997;Wong *et al.*, 1997;Chen *et al.*, 1998b). An additional, unrelated, RAD51 interaction domain has also been mapped to exon 27, at the C terminus of *H. sapiens* BRCA2 (Sharan *et al.*, 1997;Mizuta *et al.*,

1997). It has recently been shown that a site at the C terminus (serine 3291) is phosphorylated by cyclin-dependent kinases (CDK's) (Esashi *et al.*, 2005). Phosphorylation at this site increases at G2 phase and peaks at M phase, with the modification resulting in blocked C terminal interactions between BRCA2 and RAD51.

The C terminal region of BRCA2 has an important role for BRCA2 function, since cells expressing C terminal truncations of BRCA2 are hypersensitive to ionising radiation, are defective in their ability to perform recombination and have a reduced ability to form RAD51 foci upon DNA damage (Wang *et al.*, 2004). Recent work has shown that the C terminus of *H. sapiens* BRCA2 binds RAD51 filaments, but not monomers like the BRC repeat region (Esashi *et al.*, 2007;Davies and Pellegrini, 2007). It therefore appears that the C terminus stabilises RAD51 filaments, and the phosphorylation of S3291 at the C terminus, which blocks RAD51 binding could disassemble the filament.

Since evidence has been provided for possible conserved phosphorylation sites in the dog, rat and mouse BRCA2 orthologues (Yang *et al.*, 2002), it was decided to investigate whether the *T. brucei* orthologue might also possess a putative C terminal RAD51 binding domain and CDK target sites. A global multiple alignment of the C terminus of the *T. brucei* BRCA2 polypeptide with the C terminus of BRCA2 orthologues from *T. cruzi, L. major, G. gallus, H. sapiens, A. thaliana* and *U. maydis* was produced using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003). This was then visualised using the Boxshade server (http://www.ch.embnet.org/software/BOX_form.html), as shown in figure 3.13. Perhaps surprisingly, the CDK target site in human BRCA2 identified by Esashi *et al.* (2005) aligns with a serine, proline motif in *T. brucei, T. cruzi* and *G. gallus.* In contrast, no such obvious conservation is found in *L. major, A. thaliana* or *U. maydis*. Despite the potential conservation of a CDK target site, which will need to be confirmed through biochemical analyses, no obvious sequence homology between the RAD51 binding domain at the C terminus of *H. sapiens* and *G. gallus* BRCA2 is apparent with *T. brucei* BRCA2. It is interesting to note that BRCA2 in *U. maydis* does bind RAD51 at the C terminus despite a lack of conservational homology with the *H. sapiens* BRCA2 C terminal RAD51 binding motif (Zhou *et al.*, 2007). The *C. elegans* BRCA2 also binds RAD51 by a non BRC repeat motif, which also displays specificity for RAD51 filaments as opposed to monomers (Petalcorin *et al.*, 2007). It therefore appears, that a bimodal RAD51 binding function is conserved throughout BRCA2 homologues, though the position of a non BRC RAD51 binding motif is yet unknown.
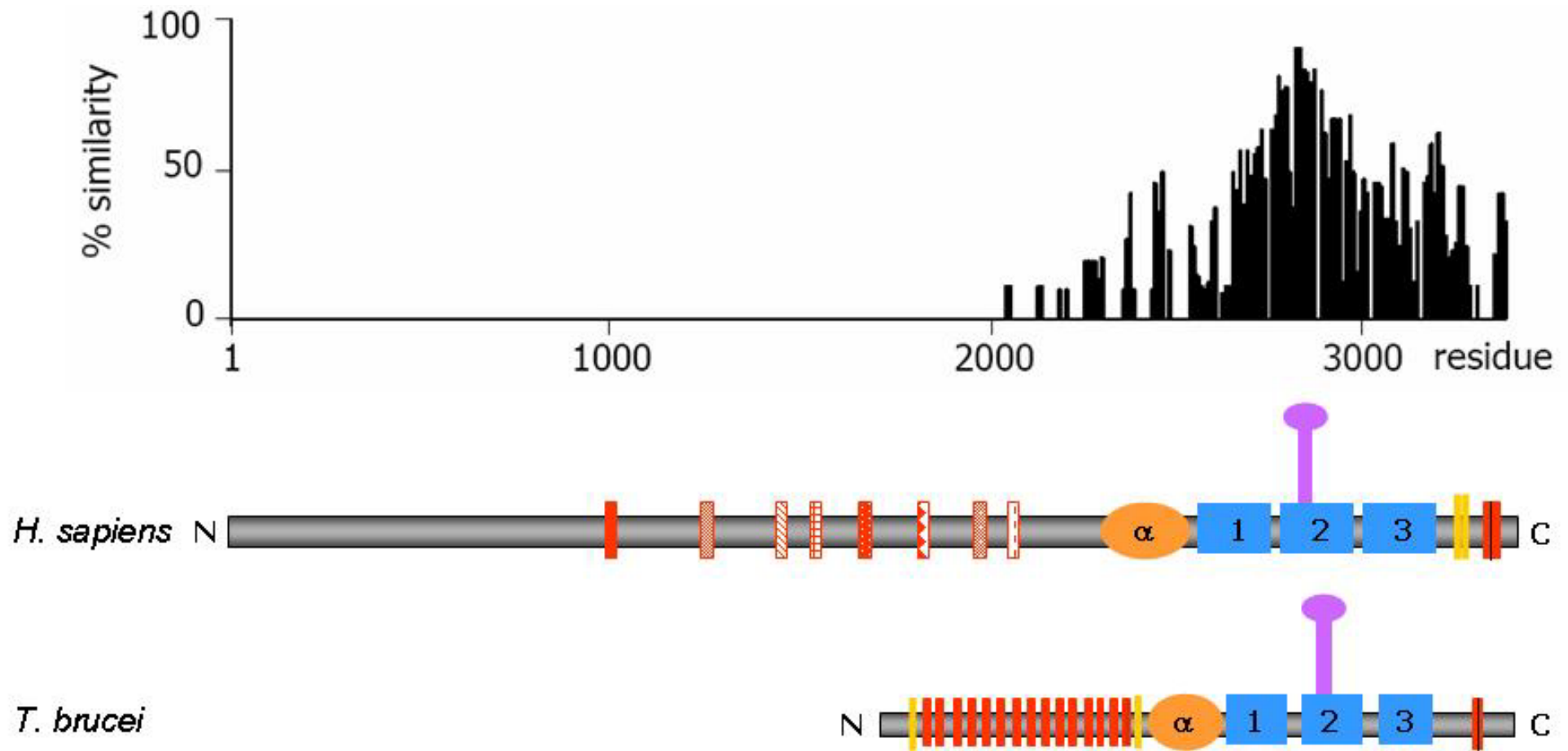
Figure 3.11 – Graph displaying the percentage similarity at the polypeptide level between *T. brucei* BRCA2 and *H. sapiens* BRCA2. The percentage similarity is displayed at each amino acid residue between *T. brucei* BRCA2 and *H. sapiens* BRCA2. Diagrams of the proteins are displayed underneath the graph and represent the length of the proteins and the positions of similarity between them. The BRC repeats are depicted by red blocks – differences in shading represents un-identical sequences; the α helical domain by the orange oval; the oligosaccharide binding domains by the blue squares; the tower domain by the lilac block extending from OB2; NLS sequences by yellow blocks and the C terminal RAD51 binding domain by a red block with a black line through it.

Figure 3.12 – Alignment of the C terminal DSS1/DNA binding domains of BRCA2. Multiple sequence alignment of the C terminal DSS1/DNA binding domains of the putative *T. brucei* BRCA2 polypeptide with homologues of Brca2 from other eukaryotes: *T. cruzi*, *L. major*, *G. gallus*, *H. sapiens*, *A. thaliana* and *U. maydis*. Sequences were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and shaded using the BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html): residues that are identical in at least 50 % of the proteins are shaded in black and similarly conserved residues shaded in grey. The coloured blocks within the alignment represent the corresponding domains in the diagram of the protein shown above.
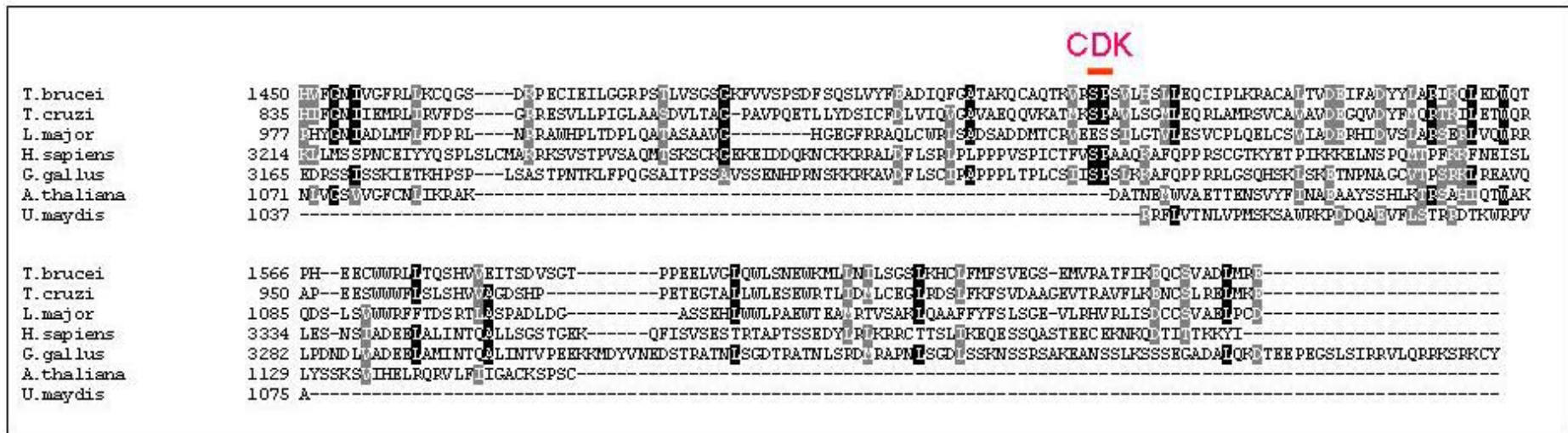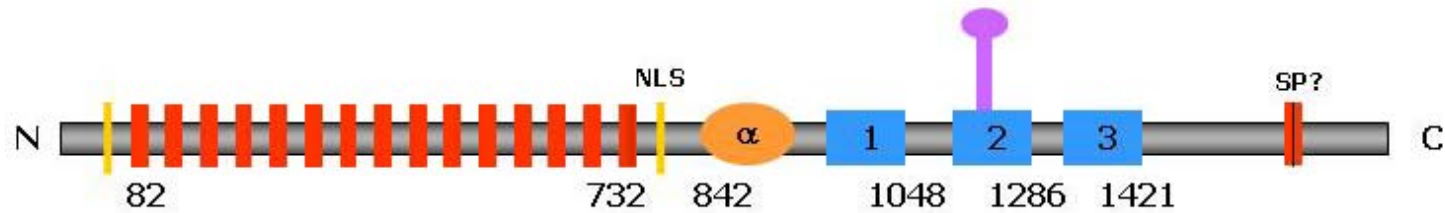
Figure 3.13 – C terminal alignment around a CDK phosphorylation site in human BRCA2. Multiple sequence alignment of the C terminus of the putative *T. brucei* BRCA2 polypeptide with homologues of Brca2 from other eukaryotes: *T. cruzi*, *L. major*, *G. gallus*, *H. sapiens*, *A. thaliana* and *U. maydis*. Sequences were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and shaded using the BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html): residues that are identical in at least 50 % of the proteins are shaded in black and similarly conserved residues shaded in grey. The putative CDK phosphorylation site is highlighted.

3.7.4 Locating the nuclear localisation signal sequences

BRCA2 from *H. sapiens* (Bertwistle *et al.*, 1997), *C. elegans* (Martin *et al.*, 2005) and *U. maydis* (Zhou *et al.*, 2007), has been shown to localise to the nucleus. It could therefore be presumed that the BRCA2 homologues identified in the trypanosomatids would also localise to the nucleus. Although it is possible that proteins without their own nuclear localization signal (NLS) enter the nucleus via co-transport with a protein that has one, the human MxB protein for example (Melen and Julkunen, 1997), many nuclear proteins have their own NLS and Lo *et al*. (2003) predicted that all the BRCA2 homologues they identified possessed their own NLS. NLSs are short regions within nuclear proteins that direct import into the nucleus, and these are currently classified into three categories (Hicks and Raikhel, 1995). The best studied NLS is that of the SV40 large T antigen (Kalderon *et al.*, 1984;Lanford and Butel, 1984), which is composed of a single peptide region containing basic residues. Another is the bipartite NLS, first found within the *Xenopus* nucleoplasmin (Robbins *et al.*, 1991;Dingwall *et al.*, 1988). The pattern of this NLS is two basic residues followed by a ten residue spacer and then another basic region consisting of at least three basic residues out of five residues. The last category of NLS is the type of N-terminal signal found in yeast proteins, such as Mat alpha2 (Hall *et al.*, 1984). This NLS possesses one or more hydrophobic residues in addition to the basic amino acids.

NLS sequences were predicted for the trypanosomatid BRCA2 homologues using PSORTII on the PSORT server (http://www.psort.org/). These results are summarised in table 3.4. PSORTII is a computer programme, which analyses the polypeptide sequences of proteins, predicts their localisation within the cell and identifies potential NLS sequences. The *T. brucei* BRCA2 homologue was predicted to be nuclear (82.6 % confidence) with three putative NLS sequences. The first and second NLS sequences, located at residues 33 and 59 respectively are of the SV40 T antigen type, whilst the third, a bipartite NLS is at residue 748. The *T. congolense* BRCA2 homologue was predicted to be cytoplasmic (47.8 % confidence), though a single NLS sequence was identified at residue 610. The *T. cruzi* BRCA2 homologue was predicted to be located in the endoplasmic reticulum (44.4 % confidence), with no NLS sequences being located. The *T. vivax* BRCA2 homologue was predicted to be nuclear (69.6 % confidence), but no NLS sequences are found with this programme. Finally, the *L. major* BRCA2 homologue was predicted to be nuclear (47.8 % confidence) with one NLS being located at residue 42.

Species	Residue	Sequence	NLS type
T. brucei	33	PKRQRSR	SV40 large T antigen
T. brucei	59	PRRDRTY	SV40 large T antigen
T. brucei	748	RKSASSSPLSSSKLARK	Bipartite
T. congolense	610	PFKRTRD	SV40 large T antigen
T. cruzi	-	-	-
T. vivax	-	-	-
L. major	42	PSRGRKG	SV40 large T antigen

Table 3.4 – Nuclear localisation signal (NLS) sequences located in the putative BRCA2 polypeptides in the trypanosomatids. The NLS sequences, amino acid residue position and type are indicated for all the trypanosomatids investigated.

The fact that all of these proteins are not predicted to be nuclear or to have NLS sequences by this programme does not necessarily mean that these proteins are not located to the nucleus or do not possess NLS sequences. This could be due to failings in the programme; for example, PSORTII does not examine sequences for the N-terminal signals found in yeast proteins, such as Mat alpha2, or for Nuclear Export Signals (NESs), and it is incongruous that the *T. congolense* BRCA2 homologue was predicted to be cytoplasmic, despite the presence of a putative NLS sequence. Equally, *T. vivax* BRCA2 homologue was predicted to be nuclear but the programme did not locate any NLS sequences. One possibility is that the *T. cruzi* and *T. vivax* BRCA2 homologues have diverged and interact with another protein, which transports them to the nucleus. Alternatively, another possibility is that the known evolutionary divergence of the kinetoplastida have diverged their NLS sequences so that they are undetectable by common eukaryotic programmes.

3.8 Examining BRCA2 *structure in the trypanosoma*

Given the potentially unusual organisation of the *T. brucei* BRCA2 protein, this section describes experimental analysis that was performed in order to determine that the prediction from the genome sequence is accurate, and to examine if this is conserved in other *T. brucei* strains and subspecies. To do this, the BRCA2 structure and expression was examined by mini-satellite variant repeat (MVR) mapping, DNA sequencing of the BRC repeat region and by Southern and northern analysis.

3.8.1 Determining the number of BRC repeats of BRCA2 *in the trypanosomatids*

Based upon the genome sequencing effort, which used the *T. b. brucei* strain TREU 927, the *T. brucei* BRCA2 protein was predicted to contain 15 BRC repeat motifs (Lo *et al.*, 2003). To verify the validity of this finding, it was necessary to investigate the number of BRC repeats directly and to examine if the number differs or is conserved between strains and subspecies of *T. brucei*. It was also important to investigate whether the sequence at

the nucleotide level was identical between repeats, since this would inform us of the mechanism of expansion, and if all strains possessed a 'degenerate' copy that differs from all other BRC repeats in the last eleven amino acids (section 3.7.1). The positioning of such a copy in the array is important to understand if it is functionally diverged, or is a potential flanking truncation in the array.

3.8.1.1 MVR mapping

Minisatellites are tandemly repeated DNA sequences normally between 10 and 100 bp, which show length variation due to differences in the number of repeat units (Jeffreys *et al.*, 1985). They also vary in the sequence of each repeat within an array. The first minisatellite variant repeat (MVR) mapping was developed in humans at locus D1S8 by Wong *et al*, in 1987 (Wong *et al.*, 1987). This minisatellite consists of a 29 base pair repeat unit showing two classes of MVR which differ by a single base substitution, resulting in the presence or absence of a *Hae*III restriction site (Jeffreys *et al.*, 1990). A much simpler PCR-based mapping system (MVR-PCR) has since been developed (Jeffreys *et al.*, 1991), which reveals length polymorphism by using an MVR primer specific to one repeat variant and a primer specific to a region flanking the minisatellite. The PCR usually contains the MVR specific primers at a low concentration to ensure that the primers will anneal to just one of the repeat units, yielding DNA fragments of sequentially increasing size. The progressive shortening of PCR products by internal priming of the MVR primers during cycling can be prevented by the use of 'tagged' primers.

MVR mapping techniques not only allows repeat length polymorphism to be identified but can also reveal information about the genetic relationships amongst different strains and subspecies of an organism. The MVR mapping technique has been applied to minisatellites in *Plasmodium falciparum* (Arnot *et al.*, 1993) to uniquely identify strains. More extensively, it has been used in *Trypanosoma brucei*, (MacLeod *et al.*, 2001a;MacLeod *et al.*, 2001b) where it has been utilised to determine population structure and to examine the relationships among *T. brucei* subspecies, providing evidence for multiple origins of human infectively.

MVR mapping was utilised here to determine the number of BRC repeats present in *BRCA2* genes from different strains and subspecies of *T. brucei*. To do this, *T. brucei brucei* strains TREU 927, Lister 427, ILTat 1.2 and EATRO 795 were analysed, as well as the related subspecies *T. brucei gambiense* (strain Eliane) and *T. brucei rhodesiense* (strain 222).

For each strain or subspecies, genomic DNA was used initially to PCR amplify the full length *BRCA2* as a template for MVR mapping. To do this, the primers *Tb BRCA2 for* and *Tb BRCA2 rev* were used, generating a DNA fragment of approximately 5 kb. From the full length *BRCA2* gene, MVR mapping was then done using a forward primer (*Tb BRC repfor*) specific to the first 20 bp of each BRC repeat and a reverse primer (*Tb BRC reprev*) specific to a region 1 bp downstream of the most 3' BRC repeat. This method is depicted in figure 3.14. PCR was performed for 18, 21 or 28 cycles, and the MVR-PCR products then separated by electrophoresis on a 1.5 % agarose gel, visualised by ethidium bromide staining and detected by Southern blotting with the hybridization of a 378 bp probe (see figure 3.14) from *BRCA2*.
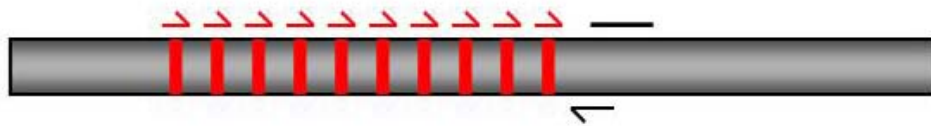
Figure 3.14 – Representation of the MVR-PCR method utilised to amplify the *T. brucei* BRC repeats. The forward primer is depicted in red and anneals to each of the BRC repeat motifs, whilst the reverse primer is depicted in black and anneals to a flanking region downstream of the BRC repeats. The black line represents the 378 bp region of *BRCA2* used as a probe.

The MVR mapping shown in figures 3.15 and 3.16 indicates that the number of BRC repeats is not constant between *T. brucei* strains or subspecies. Fifteen repeats were predicted to be present in *T. brucei BRCA2*, but this number could not be detected in the *T. brucei* strain TREU 927, which was used for the sequencing project. Instead, by counting the number of PCR products it appears that 11 or 12 BRC repeats are present in this strain. In the strain Lister 427, up to 12 BRC repeats were also detected, though the PCR ladder contained one noticeably stronger band below the largest suggesting the possible presence of two allelic variants, one with 10 BRC repeats and the other with 12. The strains ILTat 1.2 and EATRO 795, the former derived from the latter by passage in rodents, appeared to each possess 12 BRC repeats. For *T. b. gambiense* (strain Eliane) and *T. b. rhodesiense* (strain 222), the *BRCA2* genes were seen to possess fewer BRC repeats than was observed in *T. b. brucei*. A maximum number of 8 BRC repeats are predicted for both subspecies, but the phenomenon of a smaller stronger band was again observed. These results suggest that *T. b. gambiense BRCA2* possesses one allele containing 8 repeats and a second containing 5 repeats. *T. b. rhodesiense BRCA2* appears to possess one allele containing 8 repeats and a second containing 6 repeats.
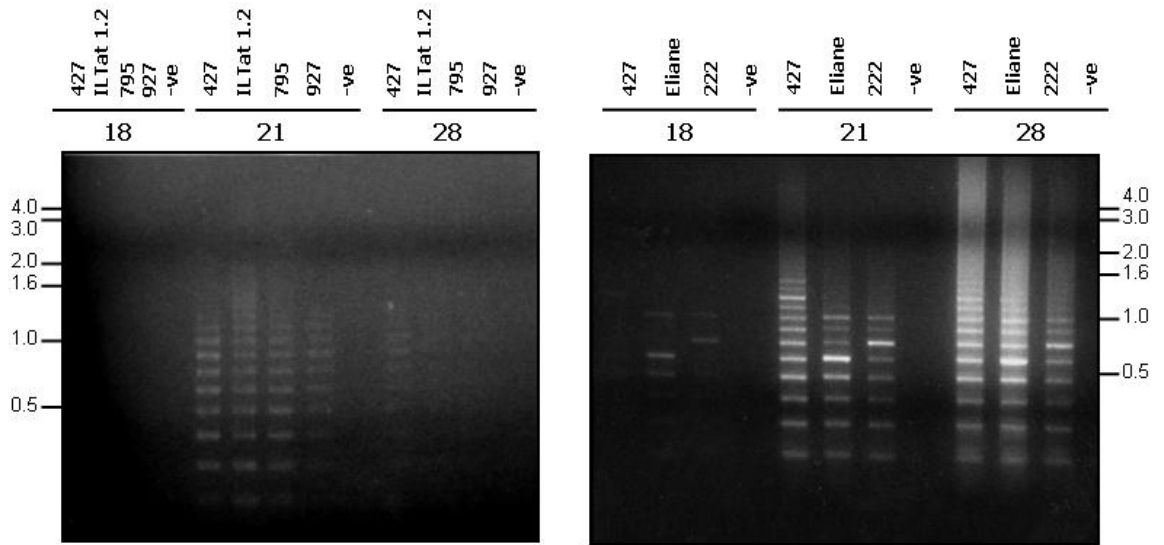
Figure 3.15 – Ethidium stained agarose gel depicting MVR mapping of *BRCA2* BRC repeat number. The gels show that the ladder of PCR products from genomic DNA from the following strains or subspecies: 427 - Lister 427; ILTat 1.2; 795 - EATRO 795; 927 - TREU 927; 222 - *T. brucei rhodesiense* and Eliane - *T. brucei gambiense*. The number of cycles undertaken in the MVR-PCR is indicated, as are DNA sizes (in kb).
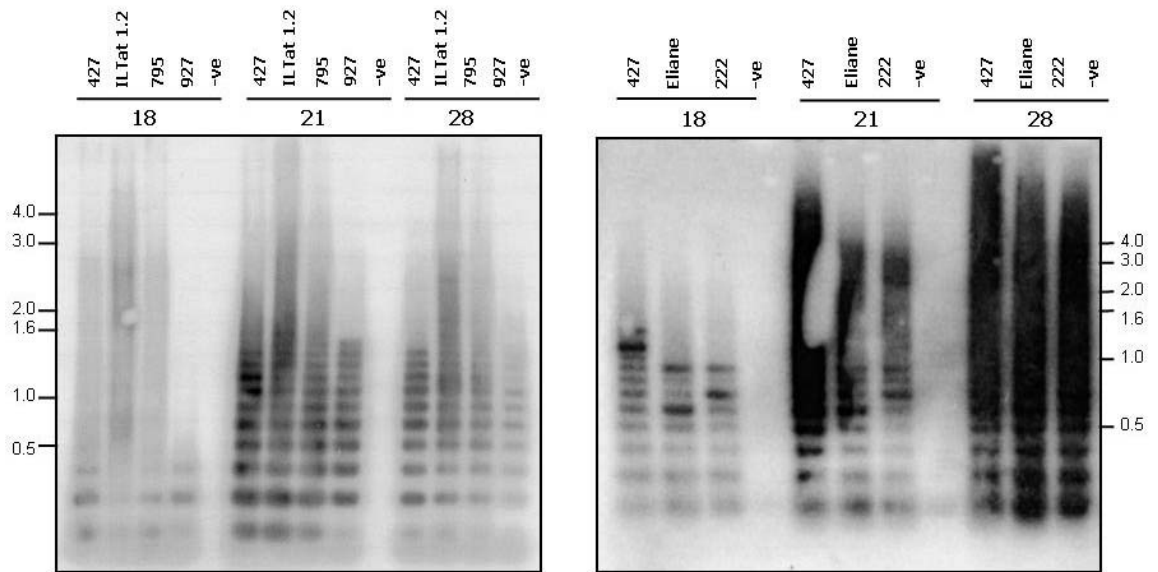
Figure 3.16 – Southern blot of the MVR mapping gel shown above. The agarose gel shown in figure 3.14 was Southern blotted and subsequently probed with a 378bp region of *BRCA2*. The strains displayed are 427 - Lister 427; ILTat 1.2; 795 - EATRO 795; 927 - TREU 927; 222 - *T. brucei rhodesiense* and Eliane - *T. brucei gambiense*. The number of cycles undertaken in the MVR-PCR is indicated, as are DNA sizes (in kb).

3.8.1.2 Topo cloning and sequencing

To attempt to confirm the above findings, and to examine the BRC repeat number variation in more detail, PCR primers were designed from the predicted genomic sequence of *BRCA2* to amplify the BRC repeat region in a number of *T. brucei* strains and subspecies and from other Trypanosome species. As before, PCR was performed on the *T. brucei brucei* strains TREU 927, Lister 427, ILTat 1.2, EATRO 795 and on the related subspecies *T. brucei gambiense* (strain Eliane) and *T. brucei rhodesiense* (strain 222) (PCR primers *Tb BRCA2 for* and *Tb BRC reprev*). In addition, similar reactions were performed with genomic DNA from *T. congolense* and *T. vivax* ILDAT2.1 (PCR primers *Tco BRCA2 for* and *Tco BRCA2 rev* for *T. congolense* and *Tviv BRCA2 5'* and *Tviv BRCA2 3'* for *T. vivax*). In all cases, a high fidelity DNA polymerase was used (Stratagene). Products ranging from 700 bp to 1.6 kb were generated, gel-extracted and cloned into the Topo TA 2.1 vector, before being sequenced by the Molecular Biology Sequencing Unit (MBSU), University of Glasgow. The produced sequences were assembled and analysed using Contig Express (Vector NTI).

The PCR-amplified BRC repeat region from the *T. brucei* strain Lister 427, and from *T. b. gambiense* and *T. b. rhodesiense*, revealed the presence of 2 predominant, large bands when run out on an agarose gel (excluding the PCR artefacts at ~500 bp) (see figure 3.17). This appears to confirm the prediction from the MVR mapping that two allelic forms of the *BRCA2* gene are present in these strains. In contrast, a single large product was generated in *T. b. brucei* strains ILTat 1.2 (excluding the PCR artefact at ~500 bp), EATRO 795 and TREU 927, again consistent with the MVR mapping.

Before examining the sequences of the PCR products from the different *T. brucei* strains and subspecies, the sequencing results from the genome sequencing effort were used to align each of the BRC repeats using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003). This was then visualised using the Boxshade server (http://www.ch.embnet.org/software/BOX_form.html), as shown in figure 3.18, allowing us to examine the predicted variation in repeat sequence. This alignment shows that each BRC repeat displays a high level of homology at the nucleotide level, with the exception of the most C terminal-coding BRC repeat, which was previously predicted to be distinct in the last eleven amino acids.

The sequencing results from the TOPO-cloned BRC repeat region clones revealed no significant differences in DNA composition between the genome sequence and the corresponding BRC repeats from the different *T. brucei* strains and subspecies (see figure

3.19). The most C-terminal-coding repeat from all samples aligned with the degenerate BRC repeat 15 from the predicted *T. brucei* genome sequence, with very small numbers of base pair differences, revealing that all the strains and subspecies retain the altered coding capacity of this repeat relative to the other BRC repeats. The BRC repeat immediately upstream of this degenerate copy, in all strains and subspecies, aligns with BRC repeat 14 from the genome prediction, and with the most N-terminal-coding repeat (BRC repeat 1). This suggests the first and last BRC repeats of the array are identical, as in the genome sequence. The differences in size and number of BRC repeats between the strains and subspecies, is most likely therefore a result of deletions and expansion of repeats in the conserved BRC repeat region, which comprises most of the sequence.

For *T. congolense* and *T. vivax* ILDAT2.1, PCR-amplification of the BRC repeat region generated a single band of 2.4 kb and 2.8 kb, respectively (not shown). Each of these was Topo-cloned and sequenced. The genome sequence prediction of the *T. congolense* BRCA2 protein suggest a BRC repeat sequence of 'GVATLFSTAAGKTVSVSESSLRAARMKLGQELCAD', with 3 BRC repeats predicted to exist. Sequencing of the PCR product, however, revealed the presence of 2 BRC repeats, one less than predicted by the genome sequencing effort (Figure 3.20). The 2 BRC repeats in the sequenced clone are not identical, but are highly related at both the amino acid and nucleotide level. This is true also of the 3 repeats predicted from the genome sequence. It is possible, therefore, that the *T. congolense* strain or isolate sequenced here has undergone a single BRC repeat deletion relative to the genome strain. It is also possible that this represents a truncated version of the BRC repeat array expansion and contraction seen in *T. brucei*.

The amino acid sequence prediction of the *T. vivax* BRC repeats is 'RKMTMFSTAAGTKLSVSTDSLEKAKKKLEDIEWRE', with just a single BRC repeat being predicted to exist. Sequencing of the *T. vivax* ILDAT2.1 BRC repeat region confirms that prediction, as shown in figure 3.21.

Taken together, these results suggest that the large number of BRC repeats represent an expansion that is a *T. brucei*-specific phenomenon. Since *T. vivax* and *T. congolense* are known to utilise a VSG-based system of antigenic variation (Richards *et al.*, 1981;Barry, 1986), it is possible that large numbers of BRC repeats is not related to antigenic variation, or is due to specific requirements of the immune evasion process in *T. brucei*.
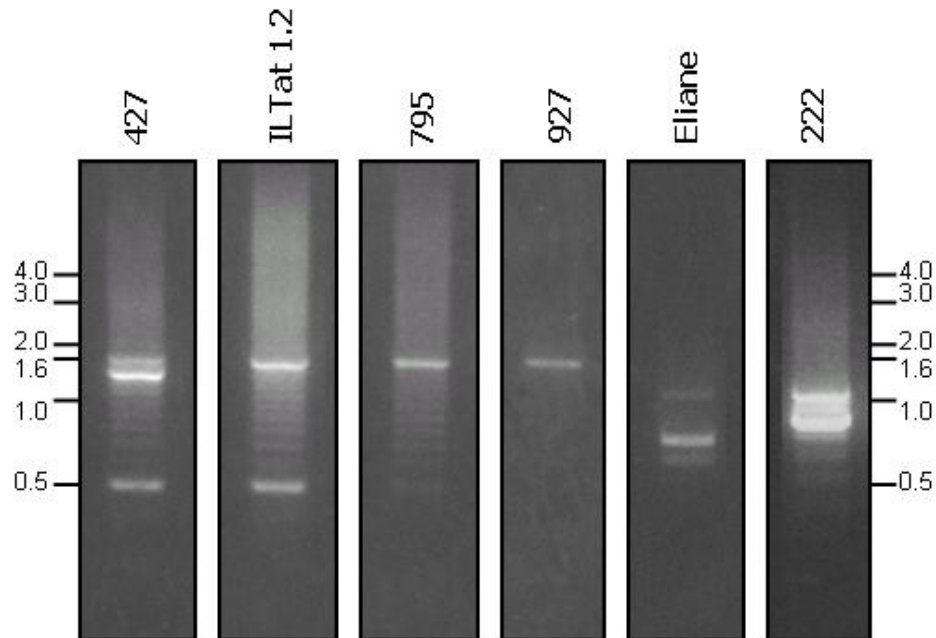
Figure 3.17 – PCR products of the BRC repeat region of *BRCA2* from different strains and subspecies of *T. brucei*. The PCR products are displayed from 427 - Lister 427; ILTat 1.2; 795 - EATRO 795; 927 - TREU 927; 222 - *T. brucei rhodesiense* and Eliane - *T. brucei gambiense*. The products were subsequently gel extracted, TOPO cloned and sequenced. In the case of 427, Eliane and 222, the two main bands were gel extracted, TOPO cloned and sequenced. DNA sizes are indicated in kilo-bases.
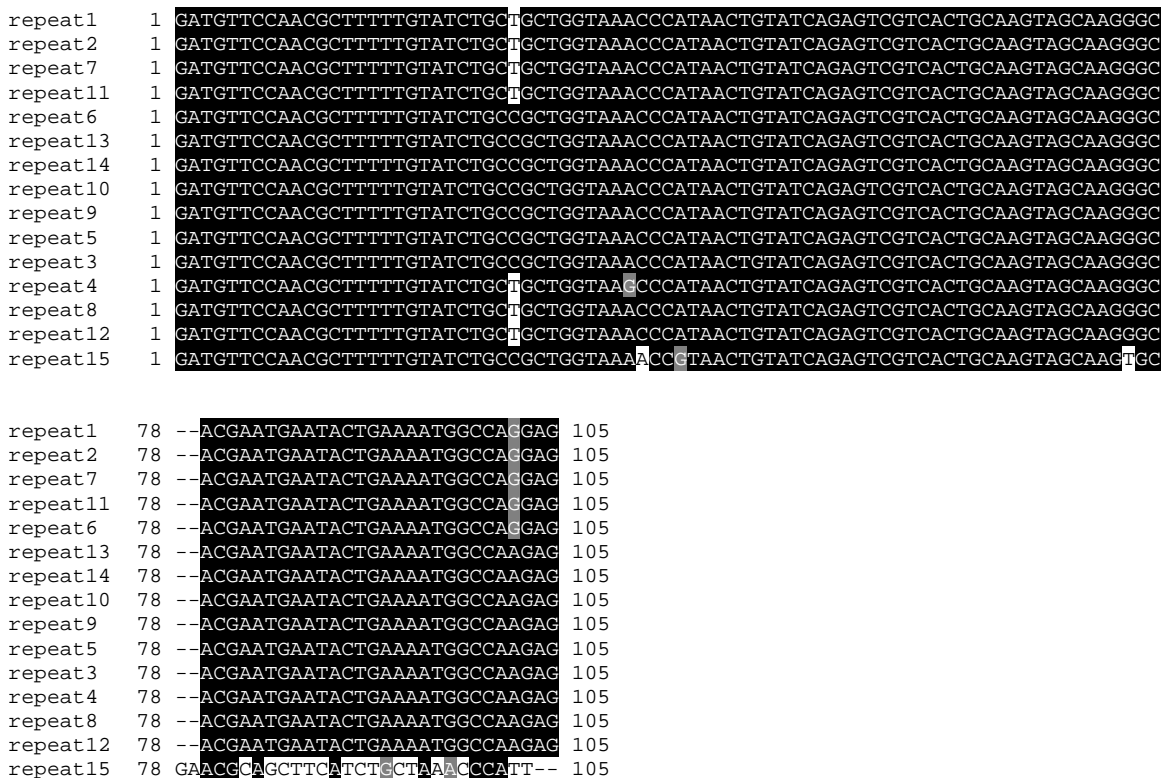
Figure 3.18 – Multiple sequence alignment of the BRC repeats from *T. brucei*. The BRC repeat nucleotide sequences were obtained from the genome project. Sequences were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and shaded using the BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html): residues that are identical in at least 50 % of the proteins are shaded in black and similarly conserved residues shaded in grey.

BRC repeat 1

BRC repeat 14

BRC repeat 15

Figure 3.19 – Alignments of the BRC repeats from various *T. brucei* strains. 427 - Lister 427;
795 - EATRO 795; 927 - TREU 927; 222 – *T. brucei rhodesiense*; Eliane - *T. brucei gambiense.*
BRC repeat indicates the predicted sequence obtained from the genome sequence. Sm and lg
indicate the smaller and larger alleles, respectively.

```
        M   V   F   S   Q   K   S   K   G   N   V   C   D   V   C   S   H   V   N   K
1       ATGGTCTTTTCCCAGAAGTCTAAGGGCAATGTTTGTGACGTGTGCTCCCACGTCAATAAG
        ------------------------------------------------------------
        TACCAGAAAAGGGTCTTCAGATTCCCGTTACAAACACTGCACACGAGGGTGCAGTTATTC

        V   D   Q   M   R   C   D   K   C   S   H   I   I   G   R   R   S   Y   S   G
61      GTGGACCAAATGAGATGCGATAAGTGTAGCCATATCATTGGGAGGCGATCTTATTCTGGG
        ------------------------------------------------------------
        CACCTGGTTTACTCTACGCTATTCACATCGGTATAGTAACCCTCCGCTAGAATAAGACCC

        S   A   R   N   S   S   K   L   S   T   P   R   K   S   N   H   E   A   N   Q
121     AGCGCCAGAAATTCGTCGAAACTGTCAACTCCACGGAAAAGTAATCATGAAGCAAATCAG
        ------------------------------------------------------------
        TCGCGGTCTTTAAGCAGCTTTGACAGTTGAGGTGCCTTTTCATTAGTACTTCGTTTAGTC

        S   D   E   H   E   A   A   K   G   A   E   N   E   S   S   V   E   R   T   A
181     AGCGATGAACATGAGGCTGCCAAAGGTGCAGAAAATGAGAGTTCTGTAGAACGTACCGCG
        ------------------------------------------------------------
        TCGCTACTTGTACTCCGACGGTTTCCACGTCTTTTACTCTCAAGACATCTTGCATGGCGC

        T   G   V   A   T   L   F   S   T   A   A   G   K   T   V   S   V   S   E   S
241     ACAGGTGTCGCCACGCTTTTTTCAACCGCAGCTGGTAAAACCGTAAGCGTTTCTGAGTCA
        ---------------------------------------------------------------
        TGTCCACAGCGGTGCGAAAAAAGTTGGCGTCGACCATTTTGGCATTCGCAAAGACTCAGT

            S   L   R   A   A   R   R   K   L   G   Q   E   L   C   A   D   G   S   T   L
301     TCTCTGCGGGCTGCCAGGAGGAAATTGGGACAGGAGTTGTGCGCCGATGGAAGCACGTTG
        ------------------------------------------------------------
        AGAGACGCCCGACGGTCCTCCTTTAACCCTGTCCTCAACACGCGGCTACCTTCGTGCAAC

            T   E   P   P   L   Q   E   S   G   P   G   V   A   T   L   F   S   T   A   A
361     ACGGAACCACCATTACAAGAGAGCGGGCCGGGTGTCGCCACGCTTTTTTCAACCGCAGCT
        ------------------------------------------------------------
        TGCCTTGGTGGTAATGTTCTCTCGCCCGGCCCACAGCGGTGCGAAAAAAGTTGGCGTCGA

            G   K   T   V   S   V   S   E   S   S   L   R   A   A   R   M   K   L   G   Q
421     GGTAAGACCGTAAGCGTTTCTGAGTCATCTCTGCGGGCCGCTAGAATGAAATTGGGACAG
        ------------------------------------------------------------
        CCATTCTGGCATTCGCAAAGACTCAGTAGAGACGCCCGGCGATCTTACTTTAACCCTGTC

            E   L   C   A   D   D   E   A   T   V   E   N   T   A   Q   S   E   S   V   G
481     GAGTTGTGCGCCGATGATGAGGCGACGGTGGAAAATACCGCGCAAAGCGAAAGTGTGGGA
        ------------------------------------------------------------
        CTCAACACGCGGCTACTACTCCGCTGCCACCTTTTATGGCGCGTTTCGCTTTCACACCCT

        V   P   P   P   S   T   P   V   A   G   R   R   A   K   G   F   R   A   A   H
541     GTTCCTCCGCCCTCCACCCCTGTGGCTGGGCGAAGGGCGAAGGGTTTCCGAGCCGCCCAT
        ------------------------------------------------------------
        CAAGGAGGCGGGAGGTGGGGACACCGACCCGCTTCCCGCTTCCCAAAGGCTCGGCGGGTA

        A   E   C   R   G   G   V   S   D   P   M   M
601     GCGGAGTGTCGGGGAGGCGTTTCCGACCCAATGATGA
        -------------------------------------
        CGCCTCACAGCCCCTCCGCAAAGGCTGGGTTACTACT
```

Figure 3.20 – Sequence of the BRC repeat region of *BRCA2* from *T. congolense.* The predicted BRC repeats for *T. congolense* are highlighted in red.

```
        M   K   Q   R   Q   V   G   E   K   S   P   G   A   F   H   G   R   G   S   E
1       ATGAAGCAGCGGCAAGTAGGTGAAAAGAGTCCCGGAGCCTTCCACGGACGAGGAAGTGAA
        ------------------------------------------------------------
        TACTTCGTCGCCGTTCATCCACTTTTCTCAGGGCCTCGGAAGGTGCCTGCTCCTTCACTT


        E   L   Q   V   S   Y   S   E   H   N   L   N   Q   R   K   R   A   X   A   S
61      GAGTTGCAGGTAAGTTACTCTGAGCACAACCTTAACCAGCGTAAGCGAGCGNGTGCCTCC
        ------------------------------------------------------------
        CTCAACGTCCATTCAATGAGACTCGTGTTGGAATTGGTCGCATTCGCTCGCNCACGGAGG


        V   C   D   L   T   E   T   S   G   G   S   T   E   A   T   I   A   Q   G   D
121     GTTTGTGATCTCACTGAAACGAGCGGAGGAAGCACTGAAGCAACCATAGCTCAAGGAGAT
        ------------------------------------------------------------
        CAAACACTAGAGTGACTTTGCTCGCCTCCTTCGTGACTTCGTTGGTATCGAGTTCCTCTA


        G   Q   A   R   K   M   T   M   F   S   T   A   A   G   T   K   L   S   V   S
181     GGGCAAGCAAGAAAAATGACCATGTTCTCAACAGCTGCTGGGACAAAGCTTAGTGTTTCC
        ---------------------------------------------------------
        CCCGTTCGTTCTTTTTACTGGTACAAGAGTTGTCGACGACCCTGTTTCGAATCACAAAGG


        T   D   S   L   E   K   A   K   K   K   L   E   D   I   E   W   R   E   V
241     ACTGACTCGCTTGAGAAGGCTAAGAAAAAGTTGGAGGATATTGAGTGGCGGGAAGAAGTG
        ------------------------------------------------------------
        TGACTGAGCGAACTCTTCCGATTCTTTTTCAACCTCCTATAACTCACCGCCCTTCTTCAC


        Q   N   N   E   A   P   L   K   Q   T   A   L   Q   S   C   A   S   S   V   P
301     CAAAACAATGAAGCGCCTCTAAAACAAACTGCTCTCCAATCTTGTGCCTCGTCTGTGCCT
        ------------------------------------------------------------
        GTTTTGTTACTTCGCGGAGATTTTGTTTGACGAGAGGTTAGAACACGGAGCAGACACGGA


        V   N   S   D   V   K   T   S   R   V   E   T   H   I   R   A   N   N   V   P
361     GTGAACTCTGATGTAAAGACATCCAGAGTGGAAACACATATCCGGGCAAACAACGTTCCA
        ------------------------------------------------------------
        CACTTGAGACTACATTTCTGTAGGTCTCACCTTTGTGTATAGGCCCGTTTGTTGCAAGGT


        S   S   M   S   A   R   S   S   I   S   D   Q   R   N   A   S   R   L   D   T
421     TCCAGCATGAGTGCTCGATCTTCGATAAGTGATCAGCGCAATGCTAGCAGACTGGATACT
        ------------------------------------------------------------
        AGGTCGTACTCACGAGCTAGAAGCTATTCACTAGTCGCGTTACGATCGTCTGACCTATGA


        S   K   G   S   S   S   P   S   L   T   P   S   T   S   R   P   Q   R   V   L
481     TCAAAGGGCAGTTCCTCACCATCACTAACGCCTTCCACTAGTAGGCCACAGCGGGTGTTG
        ------------------------------------------------------------
        AGTTTCCCGTCAAGGAGTGGTAGTGATTGCGGAAGGTGATCATCCGGTGTCGCCCACAAC


        W   Y   R   T   L   N   H   I   T   L   P   R   E   W   N   R   G
541     TGGTACCGTACGCTAAACCACATTACCCTGCCAAGGGAATGGAACAGAGGC
        ---------------------------------------------------
        ACCATGGCATGCGATTTGGTGTAATGGGACGGTTCCCTTACCTTGTCTCCG
```

Figure 3.21 – Sequence of the BRC repeat region of *BRCA2* from *T. vivax* ILDAT2.1. The predicted BRC repeat for *T. vivax* is highlighted in red.

3.8.2 Analysis of copy number of T. brucei BRCA2

In order to determine the number of copies of *BRCA2* in the *T. brucei* genome, Southern analysis was performed. Genomic DNA from Lister 427 and ILTat1.2 bloodstream stage cells was restriction digested using *Pst*I, *Eco*RI, *Eco*RV, *Hind*III, *Hinf*I and *Apo*I. The restriction digestions were then separated by gel electrophoresis on a 0.8 % agarose gel and transferred to a nylon membrane by Southern blotting (see section 2.4.1). PCR primers (*BRCA2 probe 5'* and *BRCA2 probe 3'*) were designed from the sequence encoding the region C terminal to the BRC repeats and used to amplify a 378 bp product from Lister 427 genomic DNA. This PCR product was subsequently used as a probe for the Southern analysis. The sizes of restriction fragments predicted from the genome sequence in this Southern analysis are presented in figure 3.22.

The Southern blot presented in figure 3.23 revealed the *BRCA2* gene to be present in single copy in the *T. brucei* genome, with two allelic variants for the strain Lister 427 and one allelic variant for ILTat 1.2, correlating with the results from PCR-amplification of the BRC repeat region for TOPO cloning (section 2.9.2) and from the MVR mapping (section 2.7.2). The sizes of the DNA fragments detected in this hybridisation are in all cases smaller than the predicted sizes displayed in figure 3.22. This result is again consistent with the reduced number of BRC repeats in these strains, (from the data presented in sections 3.9.1.1 and 3.9.1.2) relative to the genome prediction.

Southern analysis was similarly performed on genomic DNA from *T. brucei gambiense* (strain STIB 386) (figure 3.24). Two bands were detected for this strain, most clearly visible in the *Pst*I digest. This result suggested the presence of two allelic forms of *BRCA2* for *T. brucei gambiense,* strain STIB 386, which correlates with the result of the MVR mapping, which suggested two allelic variants for *T. brucei gambiense* (strain Eliane). The DNA fragment sizes were again smaller than the predicted sizes, which can be assumed to be due to a reduction in the number of BRC repeats, as was found for *T. brucei gambiense* (strain Eliane) (sections 3.9.1.1 and 3.9.1.2).
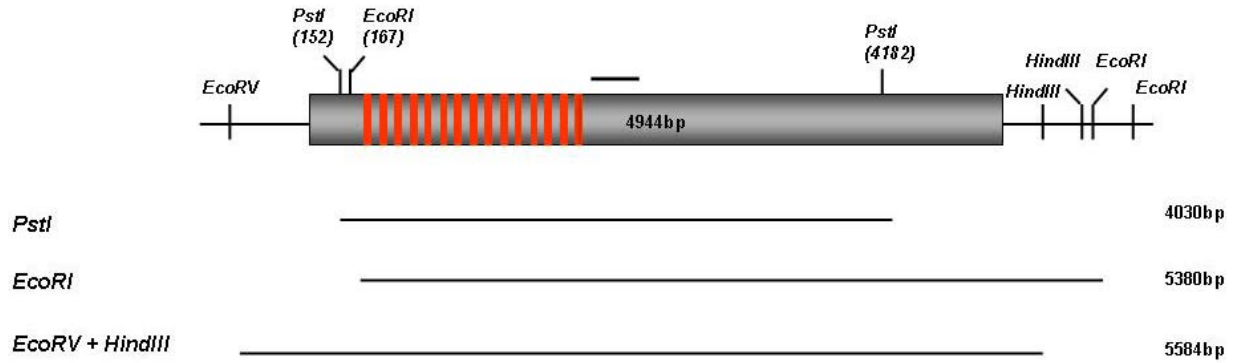
Figure 3.22 – Representation of the predicted restriction map of the *T. brucei BRCA2* locus.
The predicted ORF of *T. brucei BRCA2* is displayed as a grey box. The restriction enzymes used to analyse the copy number are indicated, and the predicted resulting fragment sizes are shown. The black line represents the region of *BRCA2* used as a probe for Southern analysis.
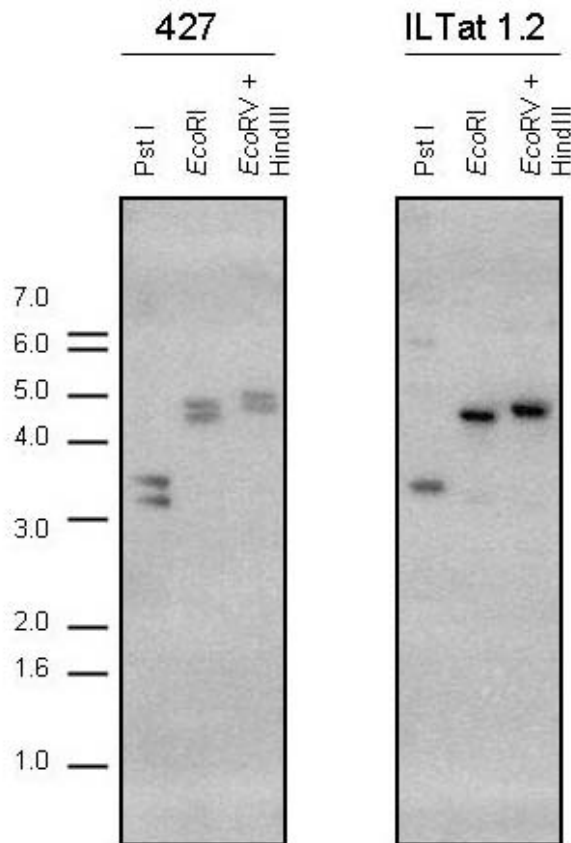
Figure 3.23 – Southern analysis of the copy number of *T. brucei BRCA2*. Genomic DNA from Lister 427 and ILTat 1.2 was digested with a range of restriction enzymes (indicated) and probed with a 378 bp region of BRCA2.
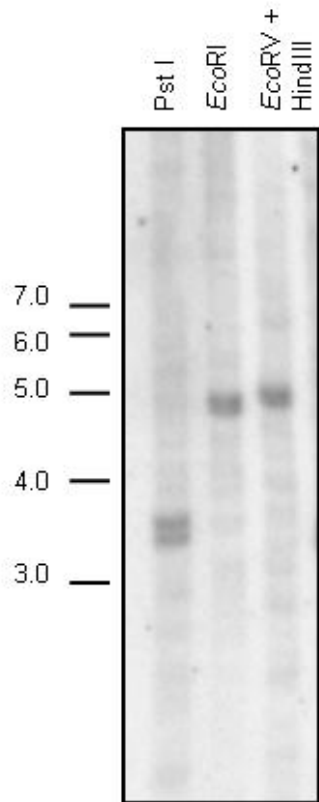
Figure 3.24 – Southern analysis of the copy number of *BRCA2* in *T. brucei. gambiense*.
Genomic DNA from STIB 386 was digested with a range of restriction enzymes (indicated) and probed with a 378 bp region of *BRCA2* (figure 3.21).

3.8.3 Analysis of T. brucei BRCA2 expression in different life cycle stages and strains or subspecies.

In order to determine if *BRCA2* is transcribed in *T. brucei*, and to ask if the expression levels are the same in both the bloodstream and procyclic stages of the life cycle, northern blots were performed on total RNA isolated from *T. brucei*. In procyclic form cells, a number of different strains and subspecies were also examined. To do this, total RNA was extracted (RNeasy Mini Kit, Qiagen) from 25 mls of bloodstream stage culture grown to a density of 2×10^6 cells.ml^{-1} and 10 mls of procyclic culture grown to a density of 1×10^7 cells.ml^{-1}. The RNA was quantified by spectrophotometry (Beckman DU650 spectrophotometer) before 10 μg and 20 μg samples were separated by electrophoresis on a denaturing formaldehyde gel. The RNA was transferred to a nylon membrane by capillary blotting and blots probed with the 378 bp fragment of the *T. brucei BRCA2* ORF. Following autoradiography, the blots were stripped by submerging them in boiling 0.1 % SDS, and subsequently re-probed with a 452 bp fragment of the RNA polymerase I ORF. The hybridising bands generated in each lane were assumed to be mature mRNA, based on their size and were quantified using the software ImageQuant (Adobe). The results of these analyses are shown in figure 3.25 and table 3.5.

The northern blot displayed in figure 3.25 demonstrates that *BRCA2* mRNA is detectable in both bloodstream stage and procyclic form cells, and in all strains and subspecies examined. The quantitative analysis shown in table 3.5 suggests that *BRCA2* mRNA may be present at a slightly higher level in bloodstream Lister 427 (158.7 %) than in procyclic Lister 427 (77.0 %), when compared to the levels of *PolI* mRNA. This would fit with the hypothesis that BRCA2 is required for VSG switching and is therefore transcribed to a higher level in the bloodstream stage. It is also interesting to note that BRCA2 mRNA may be transcribed at higher levels in different strains and subspecies, with TREU 927 transcribing the highest level of *BRCA2* mRNA in the procyclic form cells. However, since the experiment was only performed once, the differences are relatively small, and we do not know about relative protein expression levels, we cannot make any definitive conclusions.
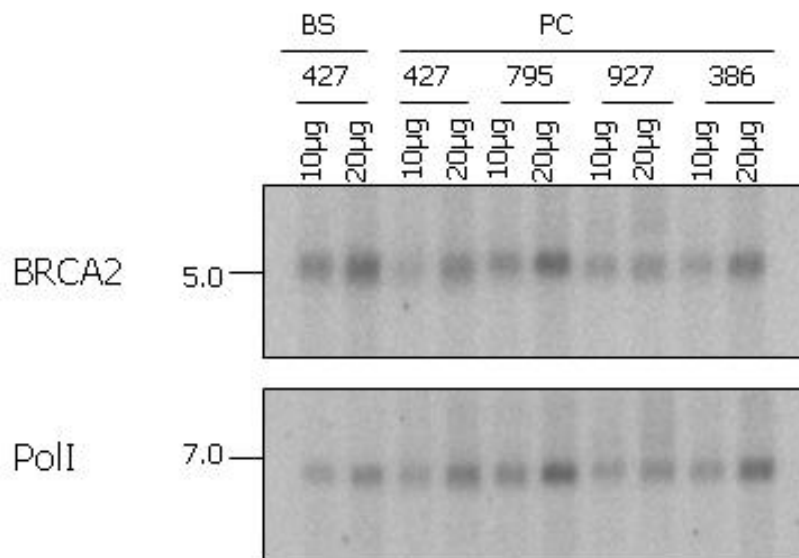
Figure 3.25 – Northern analysis of *T. brucei BRCA2* in different cell lines. Northern blots of total RNA was probed with a region of the open reading frame of *BRCA2*, then stripped and re-probed with RNA polymerase I. The quantity of total RNA loaded in each lane is indicated and size markers are shown (kb). 427, 795 and 927 correspond to the *T. brucei brucei* cell line Lister 427 and TREU 927 respectively; 386 corresponds to the *T. brucei gambiense* cell line STIB 386. BS indicates bloodstream stage, whilst PC indicates procyclic form.

	% proportion
427 BS	158.7
427 PC	77.0
795 PC	83.6
927 PC	116.2
386 PC	93.6

Table 3.5 – Quantitative analysis of BRCA2 mRNA abundance detected by northern analysis. The percentages shown represent the abundance of BRCA2 mRNA compared to the abundance of PolI mRNA in the 20 µg samples shown in figure 3.24. 427, 795 and 927 correspond to the *T. brucei brucei* cell line Lister 427 and TREU 927 respectively; 386 corresponds to the *T. brucei gambiense* cell line STIB 386. BS indicates bloodstream stage, whilst PC indicates procyclic form.

3.9 DSS1

The alpha helical and OB1 domains of *H. sapiens* BRCA2 have been shown to interact with DSS1, a protein which is absent or mutated in split-hand split-foot syndrome (Marston *et al.*, 1999;Crackower *et al.*, 1996). DSS1 has subsequently been shown to be critical for efficient function of BRCA2 during homologous recombination in both mammals and *U. maydis* (Marston *et al.*, 1999;Kojic *et al.*, 2003;Gudmundsdottir *et al.*, 2004). It was decided to investigate if homologues of DSS1 existed within the kinetoplastida, as this would enhance our knowledge of how the BRCA2 homologue operates within this order.

3.9.1 Identification of DSS1 in the trypanosomatids

Putative trypanosomatid *DSS1* genes were identified through BLAST searches of the *T. brucei, T. congolense, T. vivax, T. cruzi and L. major* genome databases using Gene DB (Sanger) (http://www.genedb.org/). Initially a BLASTp search was performed using *H. sapiens* DSS1 (NP_006295) as the query protein sequence against the *T. brucei* database. This revealed a hypothetical gene (Tb03.28C22.546), situated on chromosome 3, encoding a protein of 137 amino acids. BLASTp searches were then performed using the putative *T. brucei* DSS1 homologue as the query protein sequence against the *T. congolense, T. cruzi, T. vivax, and L. major* databases. The *T. congolense* database revealed a hypothetical gene (congo1342c06.q1k_0), encoding a protein of 138 amino acids. The *T. cruzi* database revealed a hypothetical gene (Tc00.1047053509999.70), encoding a protein of 144 amino acids. The *T. vivax* database revealed a hypothetical gene (tviv1332g04.p1k_2), situated on chromosome 3, encoding a protein of 125 amino acids and the *L. major* database revealed a hypothetical gene (LmjF29.1290), situated on chromosome 29, encoding a protein of 118 amino acids. Accession numbers for the polypeptides used in this analysis are located in the appendix.

3.9.2 Alignments

A global multiple alignment of the putative trypanosomatid DSS1 polypeptides with DSS1 orthologues from *H. sapiens, A. thaliana, S. cerevisiae, S. pombe* and *U. maydis* was produced using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003). This was then visualised using the Boxshade server (http://www.ch.embnet.org/software/BOX_form.html), as shown in figure 3.26. The alignment shows that a high level of conservation is observed throughout the DSS1 polypeptides from eukaryotes. Within this broad conservation, most of the residues that

contact BRCA2 are conserved (Yang *et al.*, 2002), most likely indicating functional conservation.

3.9.3 Pair-wise comparisons

The level of sequence homology between DSS1 from eukaryotes appears to be quite high. This result was confirmed by determining the levels of sequence identity and similarity of the putative DSS1 polypeptides from these eukaryotes by pair-wise comparisons. This was performed using AlignX (Vector NTI), and the percentage sequence identities calculated (see table 3.6). A graphical representation of pair-wise alignments between *T. brucei* DSS1 and orthologues from *T. congolense*, *T. cruzi, T. vivax, L. major, H. sapiens, A. thaliana, S. cerevisiae, S. pombe* and *U. maydis* is shown in figure 3.27. The pair-wise comparisons show that the putative *T. brucei* DSS1 polypeptide shares the highest level of sequence identity with the trypanosomatid orthologues, ranging from 43.1 % with the *L. major* homologue to 69.6 % with the *T. congolense* homologue. The level of sequence identity with DSS1 from other eukaryotes is much lower, with only 10.2 % compared to the *H. sapiens* DSS1.

```
S.pombe        1  -------------------MS----------------------RAALPSLENLED-DDEF
U.maydis       1  -------------------MSATSSNSQSIKPEAQSSKPTEDNKSSLPSLGALDE-DDEF
H.sapiens      1  -------------------MS--------------------EKKQPVDLGLLEE-DDEF
S.cervisiae    1  -------------------MS-----TDVAAAQAQSKIDLTKKKNEEINKKSLEE-DDEF
A.thaliana     1  -------------------MA---------------AEPKAATAEVVKVDLFED-DDEF
Tbrucei        1  MSNREPMSPLPTQPTSVPSVASLKAVLTPQVVLQTAQASLELPKADEVWSNWYKESEDRF
Tcruzi         1  MSNQRRTSPSPSSVATTRDTTNP----TGQEILQTARLSLELTRADEVWSAWYKETEDQF
Lmajor         1  ---------MAATAASTTEVKPS--------ILQTTDALLAVPSSETVWSAWYKESEDKF
                                                            *  *  ****  ****


S.pombe       19  EDFAT-ENWPM---------KDTELD----------------------TGDDTLWENN
U.maydis      41  EEFDE-QDWND---------AETDLSHLTAATSNALGAGVGLTMGASSASTGDHLWQDS
H.sapiens     20  EEFPA-EDWAG---------LDEDED-------------------------AHVWEDN
S.cervisiae   36  EDFPI-DTWAN---------GETIKSN--------------------AVTQTNIWEEN
A.thaliana    25  EEFEINEDWLE---------KEEVKEV------------------------SQQWEDD
Tbrucei       61  LDFISF-DATEGGGN-----GGVVRSG------SSHRTLSARRISGQFEEEVMKSWEED
Tcruzi        57  VDFQSFGDNNEGGLNVTATAGSNIMSGS-AQFSSSHRTLSARRVSGQFEEEVMKSWDED
Lmajor        44  LEFKSFELAPATGGP-----GNMSRAG-------SVRTLSMRQISGQFEEEVAKSWEED
                                                                        *****


S.pombe       45  W DDEDIGDDDFSVQLQAELKKKGVAAN-----
U.maydis      90  W DDDTV-EDDFSKALRAELDKQSSAPQAMST-
H.sapiens     43  W DDDNV-EDDFSNQLRAELEKHGYKMETS---
S.cervisiae   64  W DDVEV-DDDFINELKAELDRYKRENQ-----
A.thaliana    50  W DDDDV-NDDFSRQLRKELEN-GTDKK-----
Tbrucei      108  W EDEDV-EDTFD-AVMGRIGRYEASRAASSQK
Tcruzi       115  W EDEDV-EDTFD-AIMGQIGRYEASRAASVLK
Lmajor        91  W EDEDV-EDTFE-HIIGQISQLHATKAASK--
                    *  *         **    ****  **
```

Figure 3.26 – Global multiple alignment of the putative *T. brucei* DSS1 polypeptide with a range of DSS1 orthologues. Multiple sequence alignment of the putative *T. brucei* DSS1 polypeptide with homologues of DSS1 from other eukaryotes: *S. pombe, U. maydis, H. sapiens, S. cerevisiae, A. thaliana, T. cruzi* and *L. major.* Sequences were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and shaded using the BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html): residues that are identical in at least 50 % of the proteins are shaded in black and similarly conserved residues shaded in grey. ∗ indicates the BRCA2-contacting residues of DSS1 (Yang *et al.*, 2002).

	T.b	T.cong	T.cruz	T.v	L.m	H.s	A.t	S.c	S.p	U.m
T.b	**100**	69.6	57.4	63.8	43.1	10.2	11.7	18.0	10.2	16.8
	100	79.0	64.9	73.2	56.2	18.2	20.4	30.2	13.9	29.9
T.cong		**100**	49.7	55.4	39.1	9.4	10.9	11.5	9.4	16.7
		100	58.4	64.7	52.9	16.7	18.8	22.3	13.0	27.5
T.cruz			**100**	53.5	43.8	8.3	12.5	15.2	8.3	14.6
			100	65.3	54.9	16.7	18.8	24.1	12.5	29.2
T.v				**100**	46.4	10.4	10.4	17.6	9.6	15.4
				100	64.8	19.2	18.4	24.0	12.8	26.2
L.m					**100**	12.7	16.9	16.1	11.0	17.2
					100	20.3	23.7	30.5	15.3	35.2
H.s						**100**	42.9	31.9	40.5	27.7
						100	49.4	46.2	51.4	31.9
A.t							**100**	24.7	31.2	21.8
							100	37.1	42.9	31.9
S.c								**100**	31.1	26.1
								100	42.2	35.3
S.p									**100**	24.2
									100	35.0
U.m										**100**
										100

Table 3.6 – Pair-wise comparison of the putative *T. brucei* DSS1 polypeptide with a range of DSS1 homologues. The full length putative *T. brucei* DSS1 (T.b) polypeptide was compared with DSS1 homologues from *T. congolense* (T.cong)*, T. cruzi* (T.cruz)*, T. vivax* (T.v)*, L. major* (L.m)*, H. sapiens* (H.s)*, A. thaliana* (A.t)*, S. cerevisiae* (S.c)*, S. pombe* (S.p) and *U. maydis* (U.m). Pair-wise alignments were performed using AlignX (Vector NTI) and the percentage identities and similarities calculated. The percentage identities are displayed in bold.
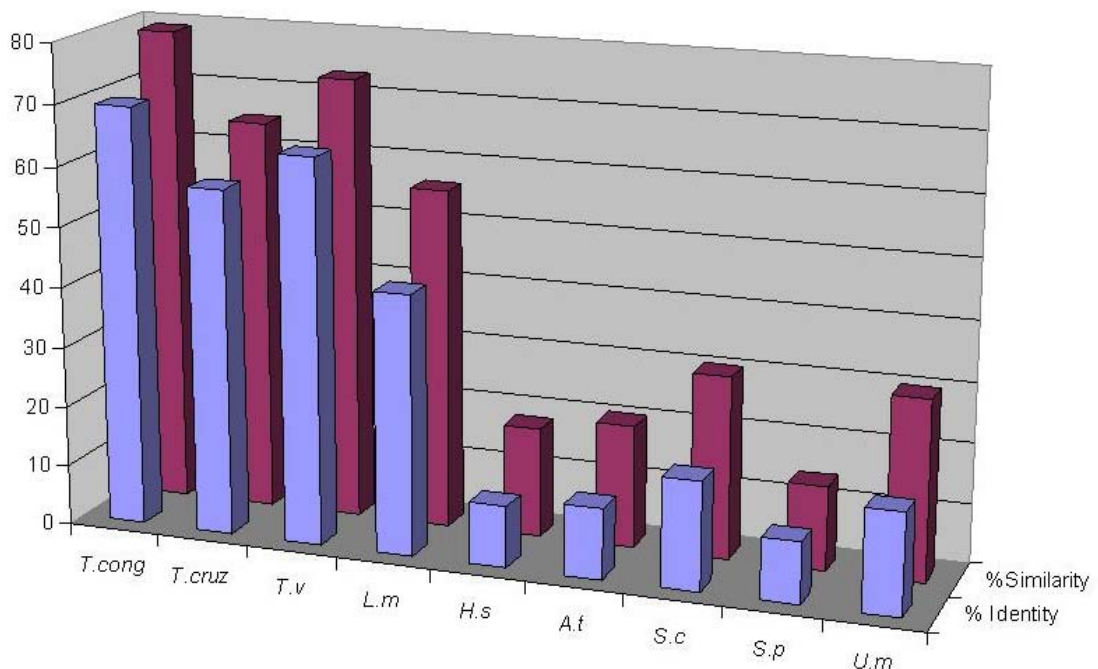
Figure 3.27 – Graph displaying the % similarity and identity between *T. brucei* DSS1 and DSS1 from other organisms. Pair-wise alignments were performed as described in Table 3.6 to compare the putative *T. brucei* DSS1 polypeptide sequence with DSS1 orthologues. Percentage identity is shown in blue and percentage similarity in maroon.

3.10 Summary

As a result of searching the trypanosomatid genome databases, BRCA2 homologues have been identified in *T. brucei, T. congolense, T. cruzi, T. vivax* and *L. major.* BRCA2 has also been identified in a range of other protists, such as *T. gondii, P. falciparum, E. histolytica, G. lamblia* and *T. vaginalis.* These results therefore display that BRCA2 is widely conserved, from protists to higher eukaryotes.

The initial characterisation of *BRCA2* found the gene to be present in single copy in all *T. brucei* strains investigated, with it predicted to be present in single copy also in other trypanosome species.

Homology with BRCA2 from other eukaryotes was limited essentially to the DBD and BRC repeats region. It appears that the DBD in *T. brucei* BRCA2 contains all 5 conserved domains, unlike BRCA2 from some other eukaryotes, and not lacking OB3 as suggested by Lo *et al.*, 2003. The protein DSS1 is also predicted to be present in the *T. brucei* genome, therefore indicating that the process of binding/regulation is likely to be the same as in higher eukaryotes. Little evidence for C terminal homology was present, apart from a putative CDK binding motive. This therefore leaves the question open whether RAD51 could bind bimodally in *T. brucei*, and this will need to be answered biochemically.

The most striking difference in *T. brucei* BRCA2 from BRCA2 in other organisms is the BRC repeat region. The *T. b. brucei* BRCA2 appears highly unusual due to the large number of BRC repeats (see figure 3.28). *T. b. gambiense* and *T. b. rhodesiense* also have this BRC repeat expansion, but less pronounced. Other trypanosome strains, however, appear much more like other protists, in that they contain only a few BRC repeats. Another notable exception is the similarity of the BRC repeats at the nucleotide level; all BRC repeats in *T. brucei* BRCA2 were discovered to be virtually identical apart from the most C terminal repeat, which appears to be a truncated version. Finally, the BRC repeats are present in a tandem array within the protein, which is unlike BRCA2 from other organisms, where the BRC repeats appear to be randomly distributed. The basis for the large number of BRC repeats within *T. brucei* BRCA2 appears to be a recent expansion, but quite why is yet unknown.

It is interesting to note that the large BRC repeat number is not limited to trypanosomes, as this phenomenon also appears in the apicomplexans, with *P. falciparum* and *C. hominis* having 8 BRC repeats each.

T. brucei BRCA2 is expressed as mature mRNA in both the bloodstream stage and procyclic form cells. We therefore wanted to follow this up by examining its function in bloodstream stage cells genetically to discover if has a role similar to other eukaryotes and indeed if it has a role in VSG switching.
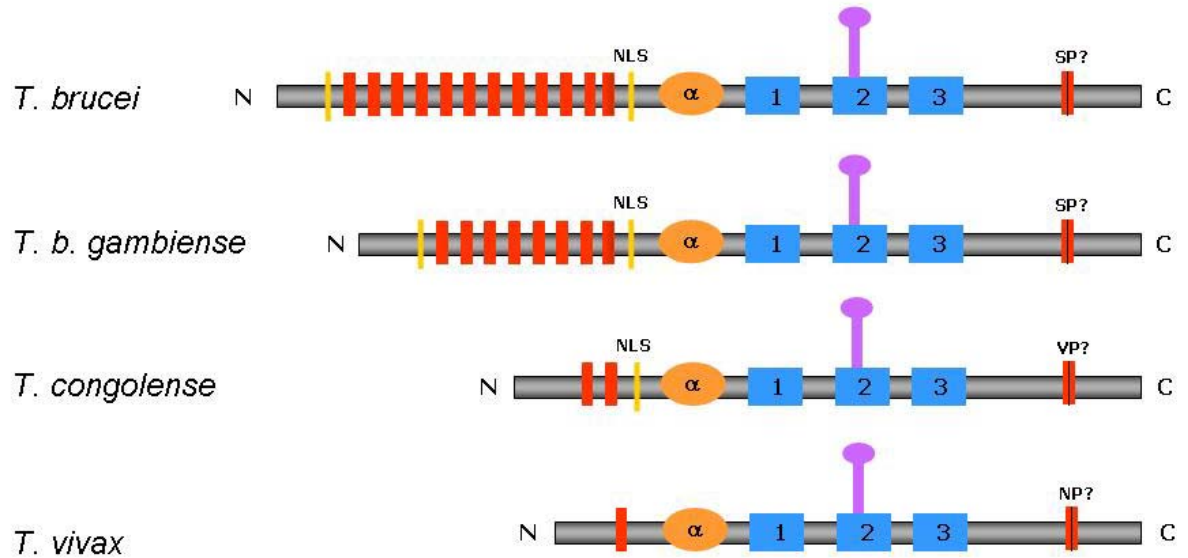
Figure 3.28 – Representation of the BRCA2 polypeptides from the trypanosomatids investigated. The figure represents the predicted domains of BRCA2 for the trypanosomatids investigated. Red bars – BRC repeats; orange oval – alpha helical domain; blue squares – OB domains; purple bar – tower domain; yellow bars with NLS – nuclear localisation signals; red bar with SP?, VP? and NP? – possible CDK phosphorylation domain. The predicted number of BRC repeats are represented.

CHAPTER 4

Analysis of the role of BRCA2 in DNA repair, recombination and antigenic variation

4.1 Introduction

To date, two genes have been identified in *T. brucei* that function in VSG switching: *RAD51* and *RAD51-3*. The results observed for *rad51-3-/-* mutants (Proudfoot and McCulloch, 2005) were highly reminiscent of those obtained for *rad51-/-* mutants (McCulloch and Barry, 1999), with both mutants displaying an impaired growth phenotype, sensitivity to a DNA damaging agent, an impaired ability to perform homologous recombination and a VSG switching defect. Although VSG switching was reduced, the trypanosomes were still able to perform homologous recombination and to switch their VSG coat, suggesting the presence of one or more pathways that can compensate in the absence of these proteins.

Homologous recombination is a complex process which involves contribution from many proteins. BRCA2 has recently been uncovered as being central to this process, (Venkitaraman, 2002;Davies *et al.*, 2001), with its absence *in vivo* compromising the ability of RAD51, the core strand exchange enzyme, to target sites of DNA damage, thereby curtailing the progression of homologous recombination.

A BRCA2 homologue has been identified in *T. brucei* and has been shown to have a highly unusual organisation, due to the large number of BRC repeats. Another notable exception is that the BRC repeat organisation forms a tandem array of repeats that are virtually identical in sequence, which is unlike BRCA2 from other organisms, where the BRC repeats appear to be randomly distributed. Homology with BRCA2 from other eukaryotes suggests that *T. brucei* BRCA2 contains all 5 conserved DBD domains, unlike BRCA2 from some other eukaryotes. A DSS1 homologue was also identified in *T. brucei,* suggesting that the process of binding/regulation is likely to be the same as in higher eukaryotes.

This chapter aims to describe the generation of *brca2-/-* mutants in bloodstream stage *T. brucei* and to analyse the role of BRCA2 in DNA damage repair, recombination and VSG switching.

4.2 Generation of gene disruption mutants in the cell lines 427 and 3174.2 in T. brucei

4.2.1 Generation of BRCA2 knockout constructs

To examine the function of BRCA2, this chapter aimed to make homozygous mutants of the *BRCA2* gene in the bloodstream stage of the cell lines Lister 427, strain MITat 1.2a, and 3174.2 (McCulloch *et al.*, 1997;Rudenko *et al.*, 1996). 3174.2 is a bloodstream derivative of Lister 427 (Melville *et al.*, 2000), which allows for the analysis of VSG switching frequency and mechanisms. The strategy used was to create 'classical knockouts', where the entire ORF is removed. In this method 5' and 3' flanking regions of the ORF were PCR-amplified, cloned into pBluescript SK and used as flanking sequence to enable homologous recombination following transformation. Between the flanks was cloned one of two antibiotic resistance markers (blasticidin and puromycin), allowing the selection of constructs that have integrated into the genome and the disruption of both alleles by replacing the ORF. A second method was also attempted, which created constructs equivalent to those for 'classical knockouts', but didn't require the cloning steps. This strategy is depicted in figure 4.1 and uses PCR to create the knockout constructs. One oligonucleotide primer was designed to represent the 5' flank of *BRCA2* (*BRCA2* KO 5'): this contained 20 bases of sequence that was homologous to the αβ Tub region of the plasmids pCP101 or pCP121 (C. Proudfoot, gift) and 100 bases homologous to the sequence upstream of the *BRCA2* ORF start codon. An equivalent primer for the 3' flank (*BRCA2* KO 3') contained 20 bases of sequence that was complementary to the ACT IR regions of the plasmids pCP101 or pCP121 and 100 bases of sequence that was complementary to the sequence downstream of the *BRCA2* ORF stop codon. PCR-amplification using these primers, and pCP101 or pCP121 as template, generated the DNA fragments *ΔBRCA2::BSDa* and *ΔBRCA2::PURa,* respectively, which should delete the ORF of *BRCA2* using 100 bp of flanking sequence to integrate the construct by homologous recombination. The PCR generated DNA fragments of the expected sizes (1240 bp and1540 bp respectively), this was PCR purified (section 2.7.1.1) and approximately 5µg used for transformations.

For the more cloning-based approach to generate the knockout constructs, oligonucleotide primers were designed for 5' (primers *BRCA2* KO5' *Xho*I and *BRCA2* KO5' *Bam_Nru*) and 3' (primers *BRCA2* KO3' *Nru*_RV and *BRCA2* KO3' *Xba*I) flanking regions of the *BRCA2* ORF. The PCR-amplified 5' flank contained 390 bp and was immediately upstream of the *BRCA2* ORF start codon. The PCR-amplified 3' flank contained 411 bp

and was immediately downstream of the *BRCA2* ORF stop codon. The sequences were
PCR-amplified using a high fidelity polymerase (Herculase, Stratagene) from Lister 427
genomic DNA. The resulting products were subsequently cloned into pBluescript SK
using the restriction sites contained within the primers (*Xho*I, *Nru*I and *Xba*I). Blasticidin
and puromycin resistance cassettes, with flanking processing signals derived from tubulin
and actin intergenic sequences to allow RNA trans-splicing and polyadenylation
(Vanhamme and Pays, 1995), were PCR-amplified (primers *Nru*Itub and *Nru*Iact) from
pCP101 and pCP121 (C. Proudfoot, gift) and cloned into the *Nru*I restriction site
introduced between the *BRCA2* 5' and 3' flanks. This generated the constructs
ΔBRCA2::BSD and *ΔBRCA2::PUR* as shown in figure 4.2. For transformation the
constructs were excised from pBluescript SK by restriction digestion with *Xho*I and *Xba*I,
and the digested DNA was then phenol: chloroform extracted and ethanol precipitated.
Approximately 5 µg of digested DNA was used in each transformation. The advantage to
this approach, relative to the BSDa and PURa constructs, is that longer stretches of
flanking sequence are provided for integrating the construct by homologous
recombination.
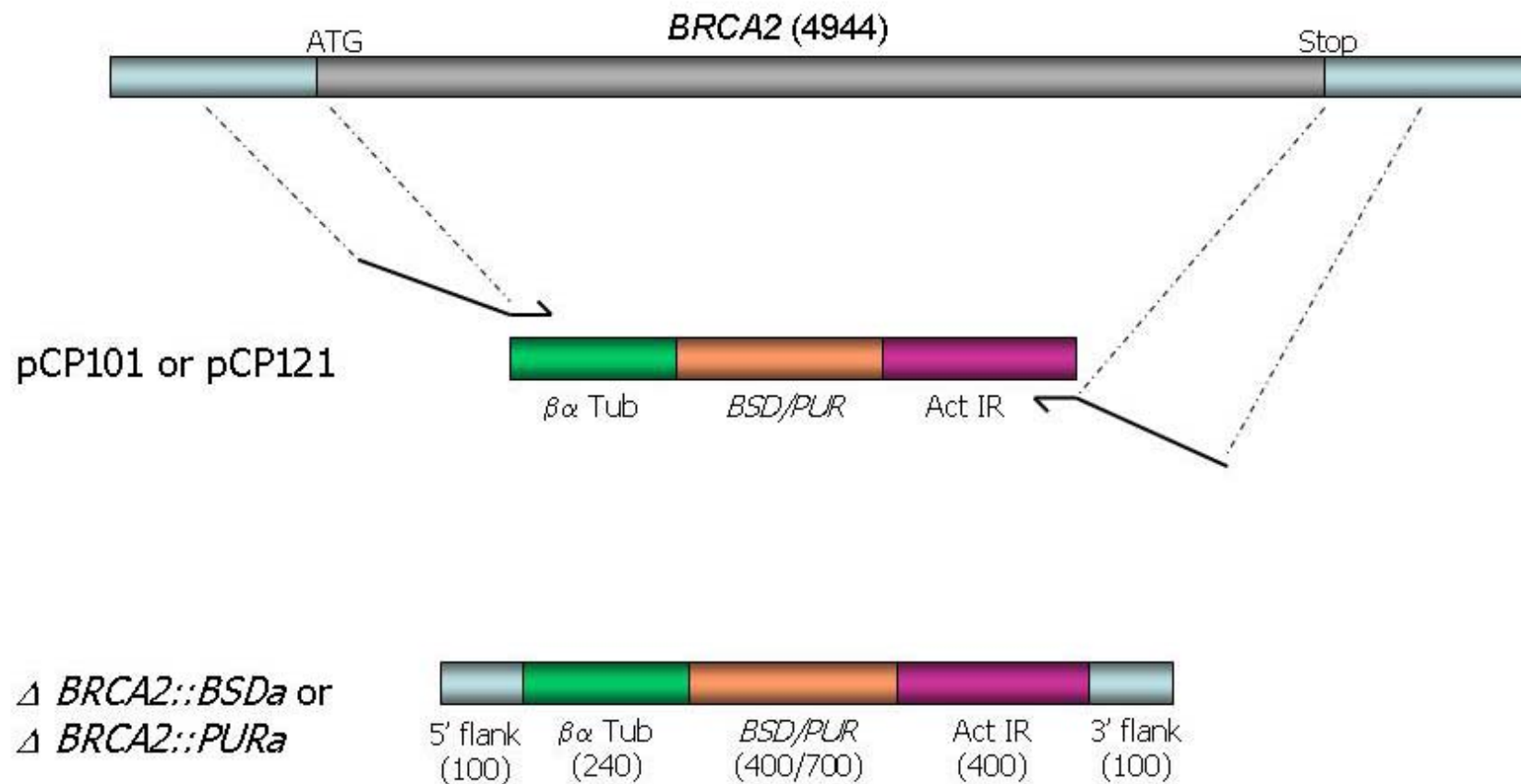
Figure 4.1 – Strategy for obtaining gene disruption constructs by PCR. PCR primers were designed that contained 20b of sequence that recognises the αβ Tub and ACT intergenic (IR) regions of the plasmids pCP101 or pCP121, and 100b of sequence from the 5' and 3' flanks of *BRCA2*. 5' flank and 3' flank correspond to the homologous regions upstream and downstream of the *BRCA2* ORF. αβ Tub: αβ tubulin intergenic region (processing signal). ACT IR: Actin intergenic region (processing signal. BSD: blasticidin resistance gene ORF. PUR: puromycin resistance gene ORF. The PCR products that are generated are shown at the bottom.
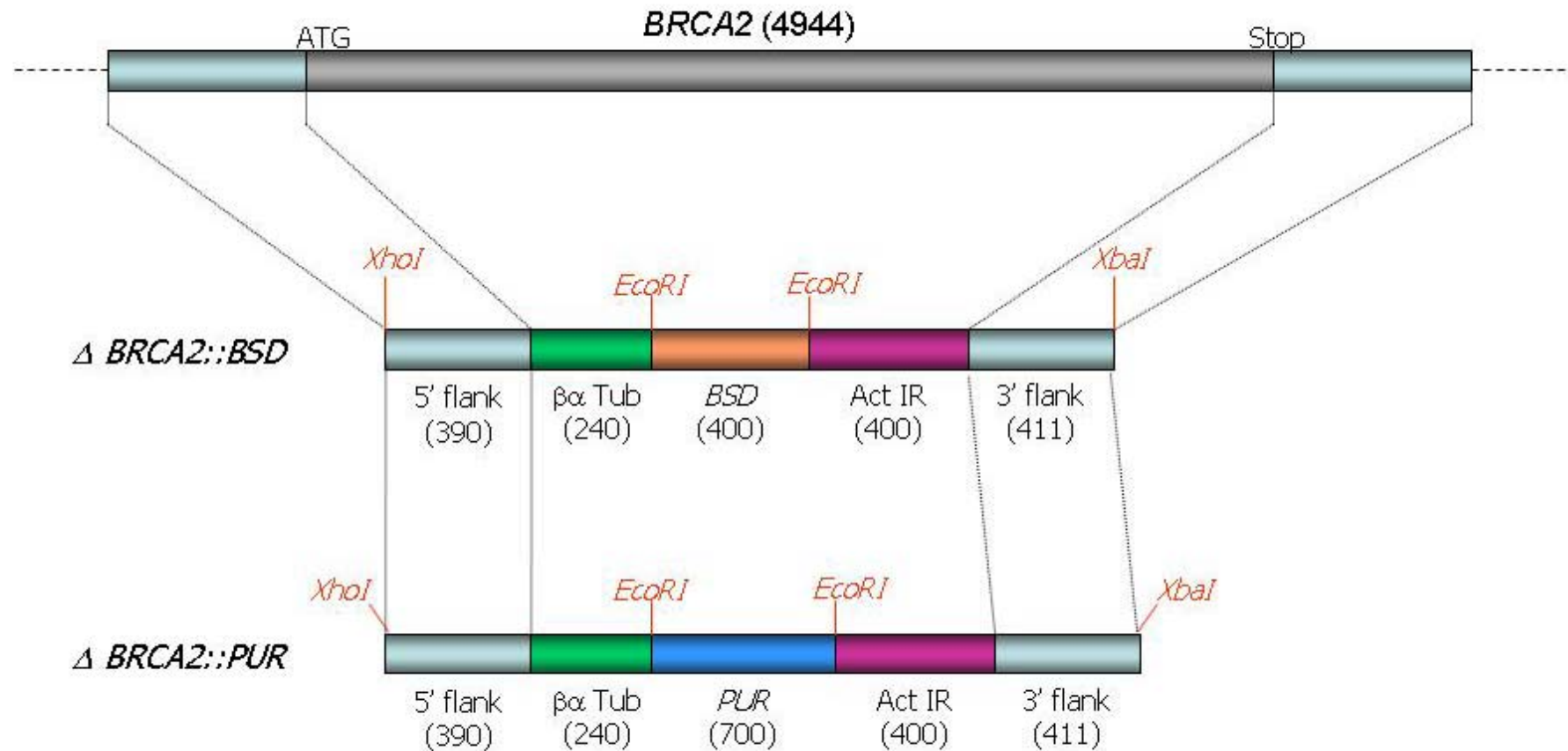
Figure 4.2 – *BRCA2* gene deletion constructs. Restriction maps of the constructs used for the deletion of *BRCA2* are shown, relative to the *BRCA2* ORF (top). Sizes of the individual components are shown in base pairs. Constructs were cloned into the pBC SK plasmid. 5' flank and 3' flank correspond to regions upstream and downstream of the *BRCA2* ORF. αβ Tub: αβ tubulin intergenic region (processing signal). ACT IR: Actin intergenic region (processing signal. BSD: blasticidin resistance gene ORF. PUR: puromycin resistance gene ORF.

4.2.2 Generation of BRCA2 mutants in the Lister 427 cell line

Two separate transformations were carried out to generate two independent *BRCA2* heterozygous (+/-) cell lines using the *ΔBRCA2::BSDa* construct. To do this, Lister 427 cells were transformed using the protocol described in section 2.1.3 and antibiotic resistant transformants were selected by placing cells on 5 µg.ml^{-1} blasticidin. The generation of heterozygous mutants was confirmed by Southern analyses, performed on *Sac*II and *Hin*dIII digested genomic DNA from seven blasticidin resistant clones and probing with the *BRCA2* 5' flank. Two independent *BRCA2+/-* clones were chosen and subsequently transformed with the *ΔBRCA2::PURa* construct in order to generate two independent homozygous (*brca2-/-*) mutants. Antibiotic resistant transformants were selected by placing cells on 5 µg.ml^{-1} blasticidin and 0.5 µg.ml^{-1} puromycin. No antibiotic resistant cells were obtained from this transformation, so it was decided to transform the two independent heterozygous mutants with the *ΔBRCA2::PUR* construct. Antibiotic resistant transformants were selected as above and the generation of homozygous mutants was confirmed by Southern analyses, performed on *Sac*II and *Hin*dIII digested genomic DNA from six blasticidin and puromycin resistant clones and probed with the *BRCA2* 5' flank.

4.2.3 Generation of BRCA2 mutants in the 3174.2 cell line

Transformations were initially carried out on 3174.2 cell lines using the *ΔBRCA2::BSDa* construct following the protocol described in section 2.1.3. Antibiotic resistant transformants were selected for by placing the cells on 5 µg.ml^{-1} BSD. Despite a number of blasticidin resistant clones being obtained, none of these were found to be *brca2+/-* mutants by Southern analyses (data not shown). Most likely the plasmids integrated by non-BRCA2 sequences, such as the tubulin or actin intergenic regions (IRs), but this was not confirmed by further analysis.

Following the failure of the *ΔBRCA2::BSDa* construct to generate *BRCA2+/-* cells in 3174.2, transformations were carried out using the *ΔBRCA2::BSD* construct. Antibiotic resistant transformants were selected for as above and the generation of heterozygous mutants was confirmed by Southern analyses, performed on *Sac*II and *Stu*I digested genomic DNA from thirteen blasticidin resistant clones and probed with the *BRCA2* 5' flank. Two independent *BRCA2+/-* mutants were chosen and subsequently transformed with the *ΔBRCA2::PUR* construct. Here, antibiotic resistant transformants were selected for placing cells on 5 µg.ml^{-1} blasticidin and 0.5 µg.ml^{-1} puromycin. *brca2-/-* mutants were identified by Southern analyses, performed on *Sac*II and *Stu*I digested genomic DNA from five blasticidin and puromycin resistant clones and probed with the *BRCA2* 5' flank.
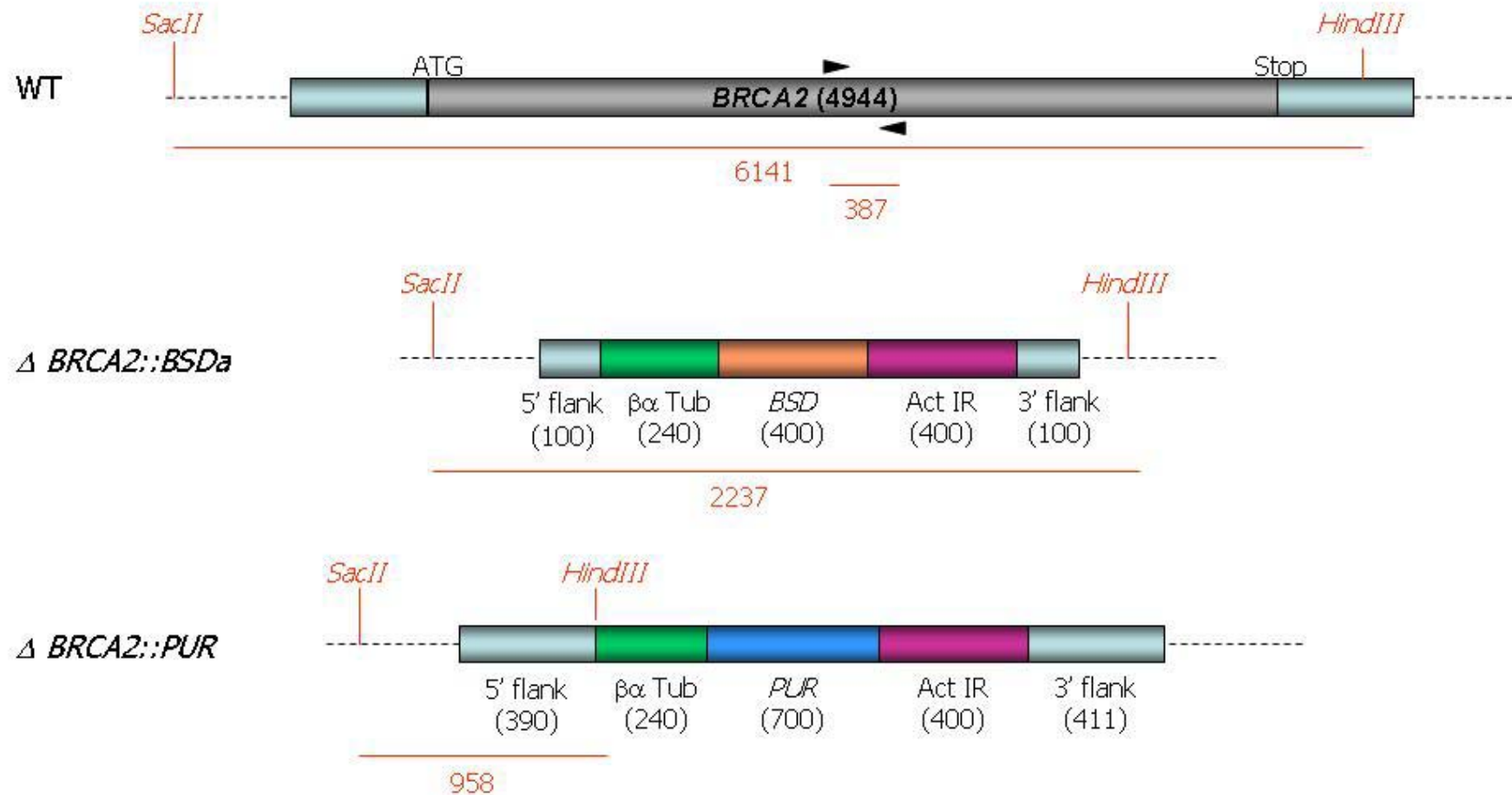
Figure 4.3 – Generation and screening of *brca2* knockout mutants in Lister 427 cells. The first allele of *BRCA2* in Lister 427 cells was deleted using *Δ BRCA2::BSDa,* a construct generated by the PCR strategy (figure 4.1). The second allele of *BRCA2* was deleted using *Δ BRCA2::PUR,* a construct generated by cloning (figure 4.2). The restriction sites and predicted size fragments used to confirm the correct integration of the constructs by Southern analyses are shown. The position of the primers used in RT-PCR are indicated by black triangles.
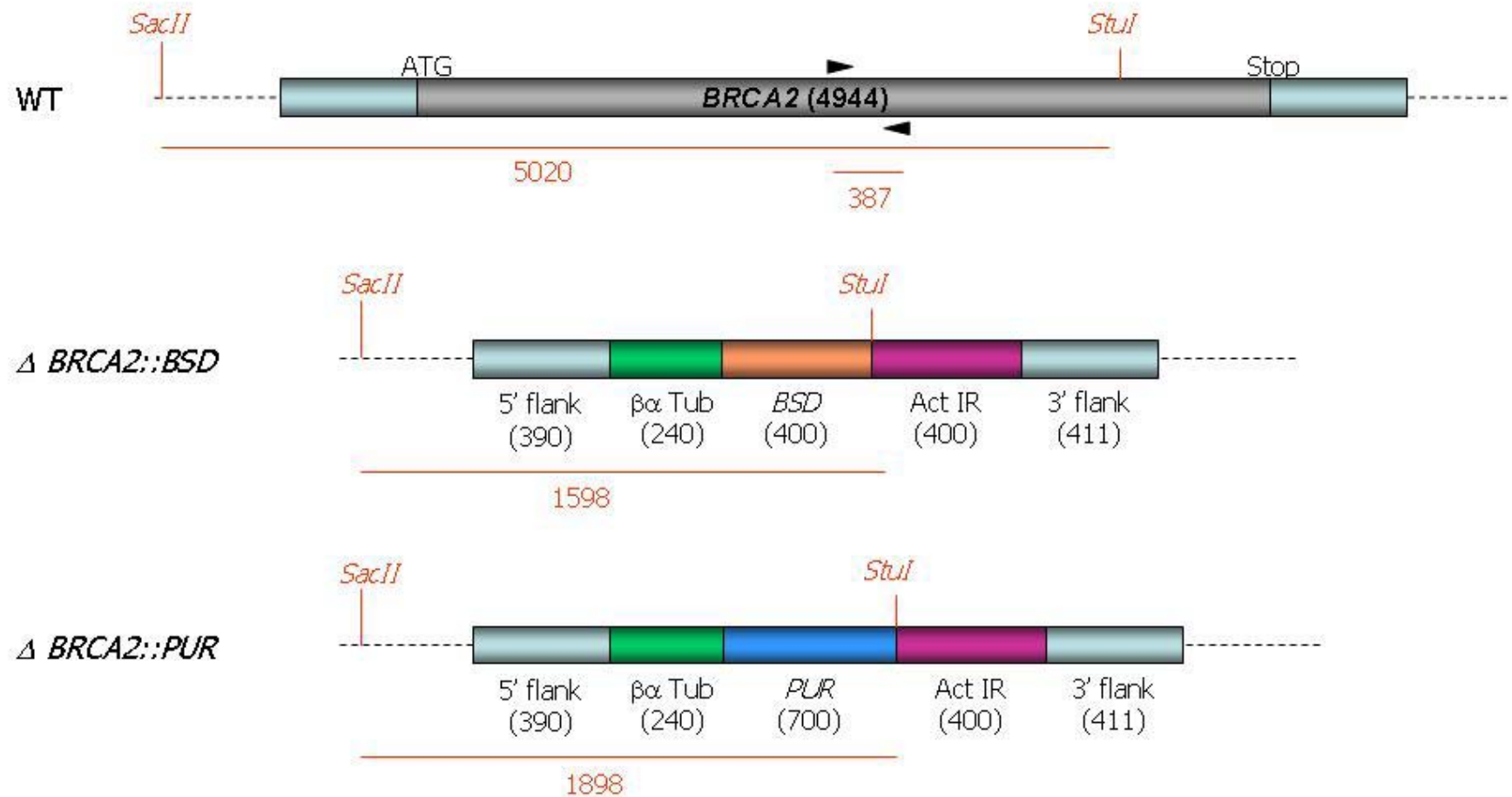
Figure 4.4 – Generation and screening of *brca2* knockout mutants in 3174.2 cells. The first allele of *BRCA2* in 3174.2 cells was deleted using *Δ BRCA2::BSD* and the second allele of *BRCA2* was knocked out using *Δ BRCA2::PUR,* both constructs generated by cloning (figure 4.2). The restriction sites and predicted size fragments used to confirm the correct integration of the constructs by Southern analyses are shown. The position of the primers used in RT-PCR are indicated by black triangles.

4.2.4 Confirmation of BRCA2 mutants by Southern analysis

To confirm the generation of two independent *BRCA2+/-* and *brca2-/-* mutants in each *T. brucei* cell line, Southern analysis was carried out on genomic DNA and compared with the wild type parent DNA. Approximately 5 μg of genomic DNA from each cell line was restriction digested overnight before being electrophoresed on a 0.8 % agarose gel and Southern blotted. The blots were probed with the 5' flank of the construct, upstream of the *BRCA2* ORF. The restriction enzymes used and expected size fragments are displayed in figures 4.3 and 4.4.

The Southern blots in figure 4.5 demonstrate that the intact *BRCA2* gene exists as two allelic variants for the wild type cell lines, distinguishable as different sized DNA fragments, as was seen in previous analysis (section 3.8.2). The *BRCA2+/-* mutants retain one allele of the intact gene and have one allele disrupted and replaced by the blasticidin construct. In all cases, it appears that the larger *BRCA2* allele was targeted first, though the significance of this is unknown. The blot shows that the *brca2-/-* cells no longer possess intact *BRCA2*, instead both alleles of the gene are deleted, one replaced by the blasticidin construct and the other replaced by the puromycin construct.

4.2.5 Confirmation of BRCA2 mutants by Reverse Transcriptase-PCR

To support the results of the Southern analyses, RT-PCR was carried out. Total RNA was prepared from the WT, *BRCA2+/-* and *brca2-/-* cells described above and 1 μg from each cell line 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). RT-PCR was carried out using primers specific for part of the *BRCA2* ORF (shown in figures 4.3 and 4.4), to detect the presence or absence of intact *BRCA2* RNA. For each cDNA prepared, a reaction was carried out without any reverse transcriptase to control for any genomic DNA that may survive the DNase I treatment. A specific product of the expected size was generated in the WT and *BRCA2+/-* cell lines. Disruption of both alleles of *BRCA2* in the *brca2-/-* mutants resulted in no PCR product being generated. Control reactions with RNA Polymerase I-specific primers showed that the cDNA in these samples was intact. This confirms that intact *BRCA2* mRNA is not present in the *brca2-/-* mutants for the cell lines Lister 427 and 3174.2 (figure 4.6).
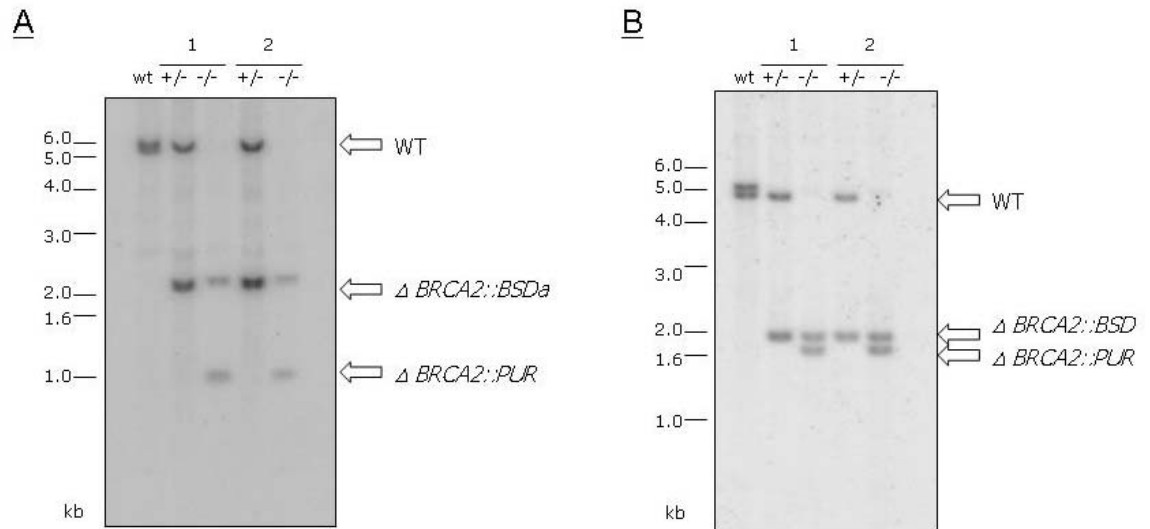
Figure 4.5 – Confirmation of the generation of *BRCA2* mutants by Southern analysis. (A) Lister 427 cell lines were digested with *Sac*II and *Hin*dIII and (B) 3174.2 cell lines were digested with *Sac*II and *Stu*I. 5 μg of genomic DNA of each cell line was restriction digested for 12 hours before being run out on a 0.8% agarose gel. The DNA was Southern blotted before being probed with the 5' upstream flank of the *BRCA2* gene. The two independent heterozygous mutants are indicated by +/- 1 and 2, and the homozygous mutants derived from these are indicated by -/- 1 and 2. WT refers to genomic DNA from untransformed cell lines.
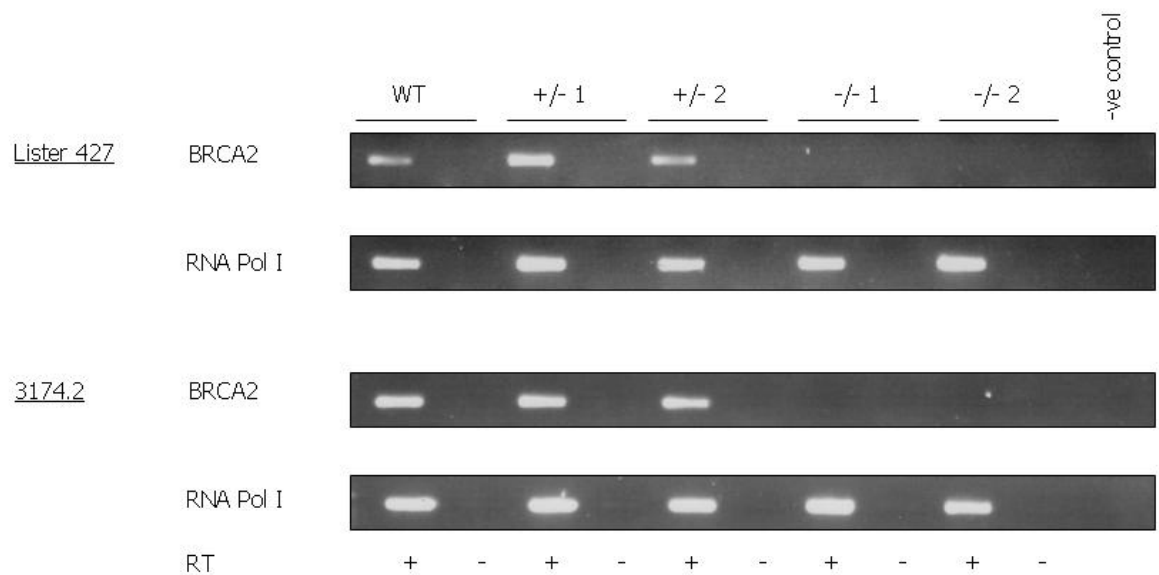
Figure 4.6 – Confirmation of the generation of *BRCA2* mutants by RT-PCR. RT-PCR was carried out on cDNA generated from total RNA from wild type (WT) cells, heterozygous mutants (+/-) and homozygous mutants (-/-). RNA polymerase I specific primers were used to control for the generation of intact cDNA. Primers specific for *BRCA2* were used to examine the expression that gene. The negative control contains no cDNA substrate. RT + denotes cDNA generated with reverse transcriptase, RT – denotes control reactions that were treated equivalently but no RT was added to the reactions.

4.3 Phenotypic analysis of BRCA2 mutants

4.3.1 Analysis of in vitro growth

To begin to analyse the role of BRCA2 in *T. brucei* the population doubling time of all cell lines was analysed to determine if mutation of *BRCA2* caused any gross growth defect. It has previously been observed that *rad51-/-* mutants (McCulloch and Barry, 1999) have a significantly increased population doubling time relative to wild type cells, as do *mre11-/-* mutants (Robinson *et al.*, 2002), *rad51-3-/-* and *rad51-5-/-* mutants (Proudfoot and McCulloch, 2005). Other genes putatively involved in DNA repair reactions, such as *KU70* (Conway *et al.*, 2002a), *DMC1* (Proudfoot and McCulloch, 2006), *MSH2* and *MLH1*(Bell and McCulloch, 2003) do not display growth defects. In addition, changes in population doubling times need to be quantified to allow any such defect to be taken into account when performing further assays, such as recombination efficiency and VSG switching (see later).

In vitro growth analysis was carried out on the *BRCA2* heterozygous and homozygous cell lines in Lister 427, and compared with the wild type cells. 5 ml cultures were inoculated at a cell density of 5 x 10^4 cells.ml^{-1} and cell concentrations counted using a haemocytometer (Bright-line, Sigma) at 24, 48, 72 and 96 hours subsequently. Three repetitions of each growth experiment, for all cell lines, were carried out and the results plotted on a semi-logarithmic scale (figure 4.7). The population doubling times for all cells lines were calculated for these data and are presented in table 4.1.
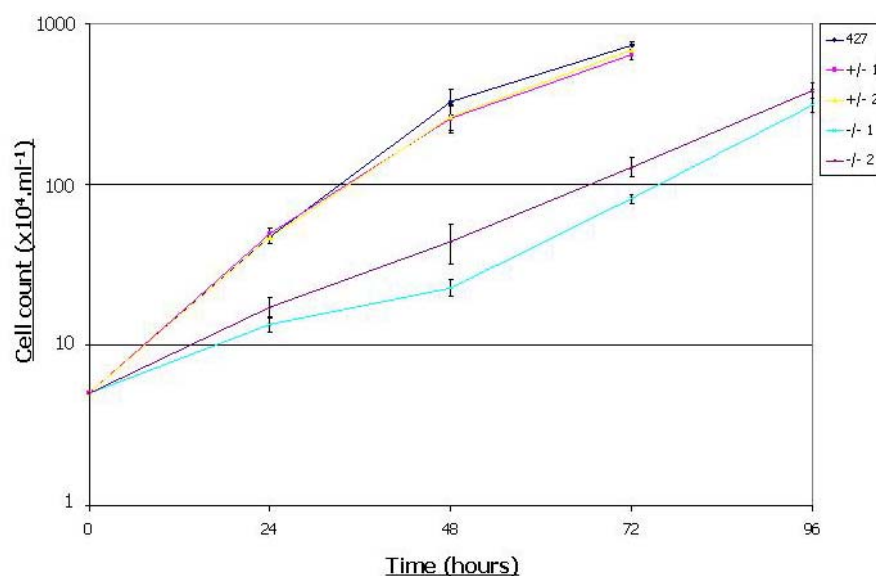
Figure 4.7 – Analysis of *in vitro* growth of *BRCA2* mutants. 5 ml cultures were set up at 5 x 10^4 cells.ml^{-1} and cell densities counted 24, 48, 72 and 96 hours subsequently. Standard errors are indicated for the counts using data from three repetitions. WT: wild type; +/-: heterozygote; -/-: homozygote.

Cell line	427	+/- 1	+/- 2	-/- 1	-/- 2
Doubling time	8.19+/-0.4	8.72+/-0.5	8.63+/-0.4	16.30+/- 0.41	15.50+/-0.34

Table 4.1 – *in vitro* population doubling times for *BRCA2* mutants. The mean doubling time for each of the independent heterozygous (+/-) and homozygous (-/-) mutants is displayed in hours. The table also displays the mean doubling times for wild type Lister 427. Standard errors are indicated.

From the growth curves shown in figure 4.7 and the population doubling times shown in table 4.1, it is apparent that the disruption of one allele had no effect on growth. However, disruption of both alleles caused the cells to grow at a much slower rate, with the population doubling time increasing by approximately a factor of two compared to wild type cells. This result is confirmed by the statistical tests displayed in table 4.2, which revealed that there was no statistical difference between the population doubling times of wild type cells and either heterozygous mutant ($p > 0.05$). A significant difference was found between wild type or heterozygous cell lines and the homozygous mutants ($p < 0.05$). Similar growth defects have previously been observed in *rad51-/-* mutants, which double in approximately 11 hours (McCulloch and Barry, 1999), and in *rad51-3-/-* and *rad51-5-/-* mutants, which double in approximately 15 hours and 13 hours respectively (Proudfoot and McCulloch, 2005).

	+/- 1	+/- 2	-/- 1	-/-2
WT	0.3803	0.1511	0.0002	0.0002
+/- 1		0.7845	0.0001	0.0004
+/- 2			0.0000	0.0001
-/- 1				0.1159

Table 4.2 – Statistical analysis of the population doubling times of the *BRCA2* mutants. P values are shown for two sample T-tests comparing population doubling times of wild type cells, *BRCA2* heterozygous mutants (+/-) and *brca2* homozygous mutants (-/-). Areas shaded in yellow indicate a significant difference.

4.3.2 Analysis of in vivo growth

In vivo growth was analysed to determine whether the growth defect observed *in vitro* was also observed during growth in mice, and to determine if the mutants remain infective. This is of particular importance when examining the rate of VSG switching in *brca2-/-* mutants, as the assay relies upon infections. *In vivo* growth analysis was carried out on the *brca2* heterozygous and homozygous cell lines in 3174.2, and compared with the wild type cells. To do this, ICR mice were infected with 1×10^6 *T. brucei* cells, previously grown in culture, *via* intraperitoneal injection. The density of trypanosomes was determined every 24 hours, up to a maximum of 120 hours. Small amounts of blood were taken from the tail of each mouse and placed in a heparin-coated capillary tube (Hawksley). 1 μl of blood was diluted in 99 μl of 0.85 % ammonium chloride, which lyses the red blood cells, allowing

the trypanosomes to be counted *via* a haemocytometer (Bright-line, Sigma). The results were plotted on a semi logarithmic scale (figure 4.8) and the doubling times calculated (table 4.3). It is important to note that for each cell line only a single mouse infection was performed, in the interests of reducing animal use.
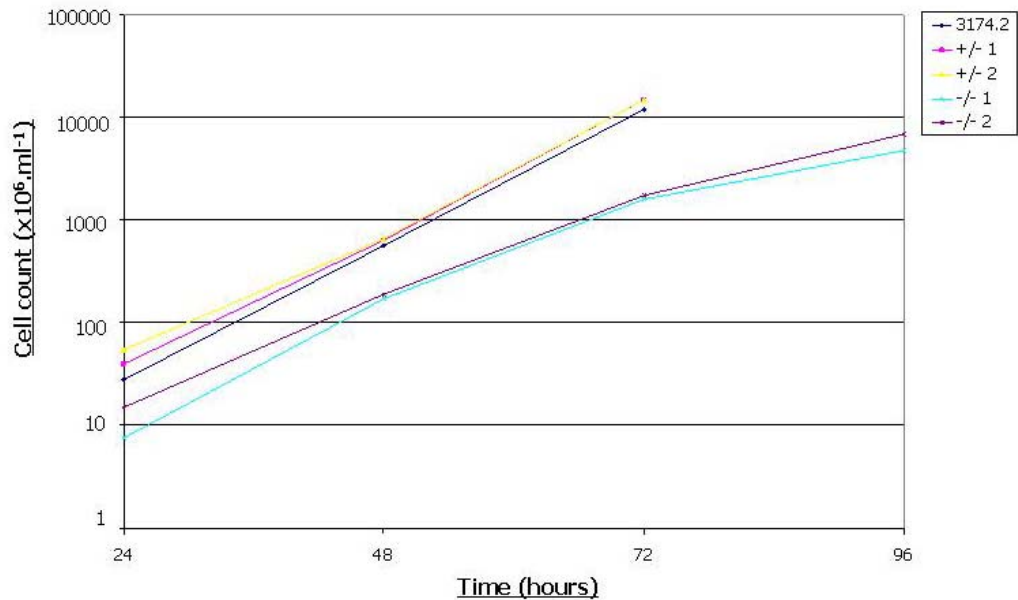
Figure 4.8 – Analysis of the growth of *BRCA2* mutants *in vivo*. Mice were infected with 1×10^6 trypanosomes, previously grown in culture, *via* peritoneal injection. The density of trypanosomes was recorded every 24 hours up to a maximum of 120 hours. One experiment was carried out for each of the heterozygous (+/-) and homozygous (-/-) mutants and for the wild type cells (WT), and the results plotted on semi-logarithmic scale graphs.

Cell line	3174.2	+/- 1	+/- 2	-/- 1	-/- 2
Doubling time	4.98	4.87	4.87	7.32	7.03

Table 4.3 – *in vivo* population doubling times for *BRCA2* mutants. The doubling time for each of the independent heterozygous (+/-) and homozygous (-/-) mutants is displayed in hours. The table also displays the doubling time for wild type 3174.2 cells.

From the growth curves shown in figure 4.8 and the population doubling times shown in table 4.3, it was again apparent that the disruption of one *BRCA2* allele had no effect on growth, but the disruption of both alleles caused the cells to grow at a slower rate. The population doubling times for all cell lines appears to be quicker *in vivo* than *in vitro*, with wild type and heterozygous cell lines doubling in approximately 5 hours compared to approximately 8 hours *in vitro*. Although the *brca2-/-* mutants have an impaired growth phenotype *in vivo*, this appeared to be less severe than *in vitro*, with the population doubling time increased by a factor of 1.4 compared to wild type cells.

4.3.3 Analysis of the cell cycle

Since cells deficient in *BRCA2* have an impaired growth rate, we wished to investigate the reason behind this. Increased population doubling times could be due to a cell cycle stall, cells taking longer to complete all stages of the cell cycle, including cytokinesis, or due to increased cell death in the population. It is particularly easy to define cell cycle stages in Kinetoplastids due to the presence of their mitochondrial DNA (kinetoplast), which is structured into an observable entity following staining and replicates and divides at distinct times in the cell cycle relative to that of the nuclear DNA (McKean, 2003) (see figure 4.9). Cells in the G1 and S phases of the cell cycle contain 1 nucleus and 1 kinetoplast (1N 1K). Kinetoplast division precedes nuclear division, which results in cells containing 1 nucleus and 2 kinetoplasts (1N 2K) that are placed cells in the G2 phase of the cell cycle. Following this, the nucleus divides and generates cells with 2 nuclei and 2 kinetoplasts (2N 2K). Such cells are in the M, mitotic, phase of the cell cycle. Cytokinesis subsequently generates two 1N 1K cells in G1 phase. Staining for DNA in Kinetoplastids therefore provides a picture of the cell cycle stage of individual cells in the population.
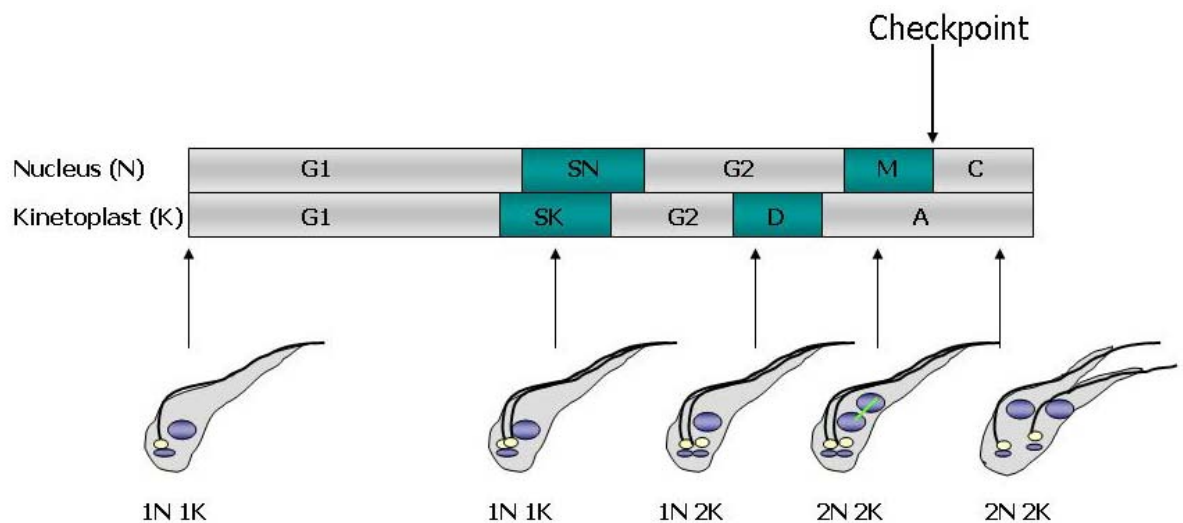
Figure 4.9 – The cell cycle of bloodstream form *T. brucei*. The diagram shows the differences between replication and division of the nucleus and kinetoplast during the cell cycle. During the G1 and S phases of the cell cycle *T. brucei* contains 1 nucleus and 1 kinetoplast (1N 1K). Synthesis of the kinetoplast DNA (SK) commences shortly before the synthesis of the nuclear DNA (SN). Kinetoplast division (D) occurs before the nucleus, resulting in cells having 1 nucleus and 2 kinetoplasts (1N 2K). Nuclear mitosis (M), which leads to cells having 2 nuclei and 2 kinetoplasts (2N 2K), occurs prior to cytokinesis (C), which generates two progeny containing 1N 1K, which will restart the cell cycle. G1 and G2 represent cell cycle growth phases. The apportioning (A) phase of the cell cycle is when the basal bodies (white circles) move apart. Figure adapted from McKean (2003).

To examine the DNA content of the *BRCA2* mutants, the cells were grown in culture to a density of 1 x 10^6 cells.ml^{-1}. 1ml of culture was then centrifuged, washed with PBS (Phosphate buffered saline) and resuspended in 1 ml of PBS. 10 µl samples were spotted

onto microscope slides (C.A.Hendley Ltd) and allowed to air dry. The trypanosomes were then fixed by soaking in methanol for 5 minutes, before being allowed to air dry. A drop of vectashield with 4', 6-Diamidino-2-phenyindole (DAPI) (Vector Laboratories Inc.) was added to the slides, a cover slip placed over the slides and sealed with clear nail varnish. The slides were visualised in phase contrast, to determine the cell outline, and under UV to visualise the DAPI. The number of cells in each cell cycle stage was counted (see table 4.4) and the percentages graphed (see figure 4.10). Cell counts were conducted blind by two independent researchers. To compare any phenotypes that might be seen in the *BRCA2* mutants with other *T. brucei* repair factors, *rad51-/-* cells were also examined in this way.

Cell line	Cell cycle stage				Total
	1N 1K	1N 2K	2N 2K	Other	
Wild type	143	28	21	0	192
	109	20	12	4	145
	152	36	14	3	205
	127	26	7	6	166
	109	23	6	3	141
Total	640	133	60	16	849
+/-1	153	27	19	0	199
	97	31	17	5	150
	153	32	9	4	198
	134	21	13	5	173
Total	537	111	58	14	720
+/-2	158	32	17	0	207
	117	15	13	2	147
	144	31	11	1	187
	121	26	8	4	159
Total	540	104	49	7	700
-/-1	100	42	17	17	176
	74	23	9	18	124
	130	33	22	19	204
	54	12	5	15	86
	129	30	12	13	184
Total	487	140	65	82	774
-/-2	120	45	24	17	206
	83	18	6	18	125
	119	35	24	28	206
	65	16	9	18	108
	137	30	16	19	202
Total	524	144	79	100	847
rad51-/-	190	19	13	3	225
	145	17	8	5	175
	102	22	9	5	138
	115	21	7	5	148
Total	552	79	37	18	686

Table 4.4 – DAPI analysis of the cell cycle of *BRCA2* mutants. The DNA content of the *BRCA2* heterozygous (+/-) and homozygous (-/-) mutant cell lines were visualised by DAPI and compared with wild type Lister 427 cells and *rad51-/-* cells. The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications (others) were counted in 4 separate experiments and tabulated in separate rows.
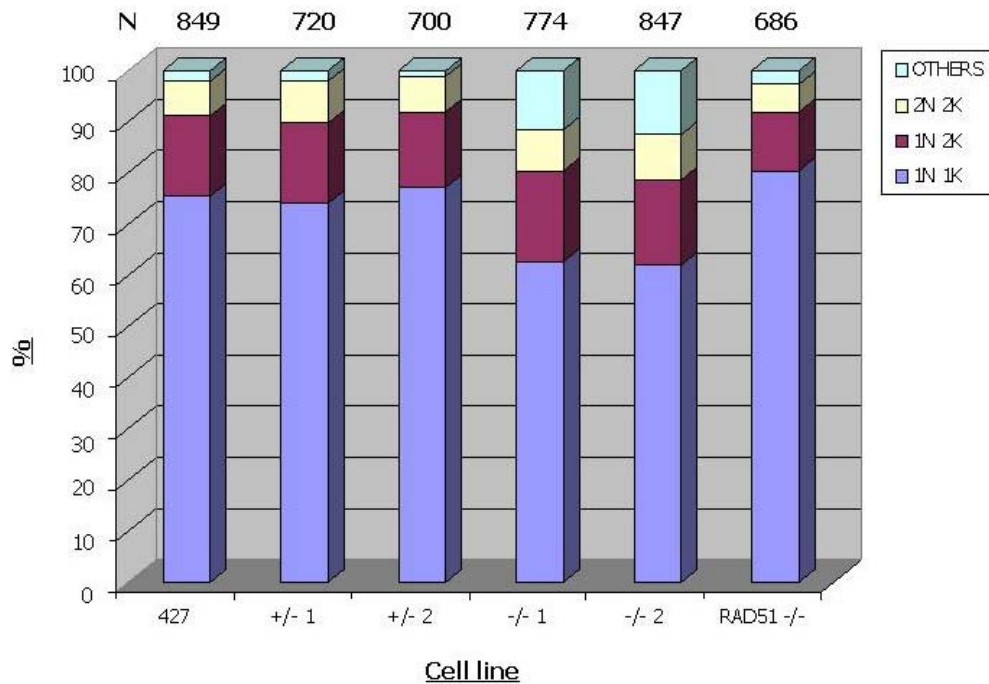
Figure 4.10 – DAPI analysis of the *BRCA2* mutants. The DNA content of the *BRCA2* heterozygous (+/-) and homozygous (-/-) mutant cell lines were visualised by DAPI and compared with wild type Lister 427 cells and *rad51-/-* cells. The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications (others) were counted and represented by their mean count as a percentage of the total cells counted (N).

From this cell cycle analysis, it was clear that mutation of one *BRCA2* allele did not affect the relative distribution of cells into different stages. This result was expected since there was no growth phenotype observed in the heterozygous cell lines. However, mutation of both *BRCA2* alleles caused a reduction in the number of cells with 1N 1K DNA content: from 75 % in wild type cells to 61-62 % in homozygous cells. Most of this reduction appeared to be accounted for by an increase in cells we have described as 'others', which rose from 2 % in wild type cells to 10-12 % in homozygous cells; the 'others' are those that do not fall into any of the expected phases, and include cells with a greater than expected number of nuclei or kinetoplasts, or an absence of one or both. These cells arise due to incorrect segregation of the nuclei and kinetoplasts during cell division, or due to problems during DNA replication. Examples of some of these cells are shown in figure 4.11, and include zoids (cells with 1 kinetoplast and no nucleus) and 'monsters' (cells in which, the number of nuclei and kinetoplasts could not be determined). The DNA content of the 'others' found in the *brca2-/-* cells are presented in figure 4.12.

Chi squared analysis was performed on these data comparing the relative distribution of cells into cell cycle stages. The results are shown in table 4.5 and display that the cell cycle distribution of both independent *brca2-/-* mutants are significantly different from the distribution of cells in wild type and *BRCA2+/-* cells, with chi-squared values of 40.95 to 120.97 (at P = 0.0001) for the *brca2-/-* cells relative their *BRCA2+/-* precedents.

The increase in aberrant cell types, which were only occasionally seen in wild type or *BRCA2+/-* cells could either result from the inappropriate timing of cytokinesis, difficulties associated with DNA replication prior to cytokinesis, or a severe DNA repair/recombination defect that fails to complete the repair of endogenous levels of DNA damage prior to cell division.

It is interesting to note that in *rad51-/-* mutants, the number of cells in each cell cycle stage did not differ from those found in wild type cells. The distribution of cells in both of the independent *brca2-/-* mutants were found to be significantly different from *rad51-/-* mutants, with chi squared values of 14.36 at P = 0.0025 and 16.18 at P = 0.0010. This would therefore appear to indicate either that the cell cycle abnormalities in *brca2-/-* cells arise from a BRCA2 specific function, distinct from its role in RAD51 recombination, or a significantly more severe DNA repair defect than found in *rad51-/-* cells. These findings are further supported by research indicating that in *rad51-3-/-* and *rad51-5-/-* mutants, the number of cells in each cell cycle stage do also not differ from those found in wild type cells (Proudfoot and McCulloch, 2005), suggesting that any such abnormalities are not generally true of factors that mediate RAD51 action.

The data also display that the impaired growth rate in *brca2-/-* mutants was not due to a cell cycle stall, since the number of 1N 2K and 2N 2K cells was essentially equivalent to that of wild type cells. A cell cycle stall in any stage should reduce the number of these cells and cause an accumulation of the cells in the preceding stage (McKean, 2003). The most likely explanation for the increased growth rate, in common with *rad51-/-* mutants is increased cell death.

	+/- 1	+/- 2	-/- 1	-/-2	rad51-/-
WT	0.1530 0.9848	0.4960 0.9199	43.1260 0.0001	55.7100 0.0001	2.1330 0.5452
+/- 1		0.7030 0.8725	40.9540 0.0001	52.7460 0.0001	2.5740 0.4620
+/- 2			95.7590 0.0001	120.967 0.0001	3.8890 0.2737
-/- 1				0.3230 0.9556	14.3610 0.0025
-/-2					16.1800 0.0010

Table 4.5 – Statistical analysis of the cell cycle data for *BRCA2* mutants. Chi squared analysis the of cell cycle data for wild type cells, *BRCA2* heterozygous mutants (+/-), *brca2* homozygous mutants (-/-) and *rad51-/-* mutants. The numbers indicated in bold represent the Chi squared value, whilst the numbers below represent the P value at which it was calculated. Areas shaded in yellow indicate a significant difference.
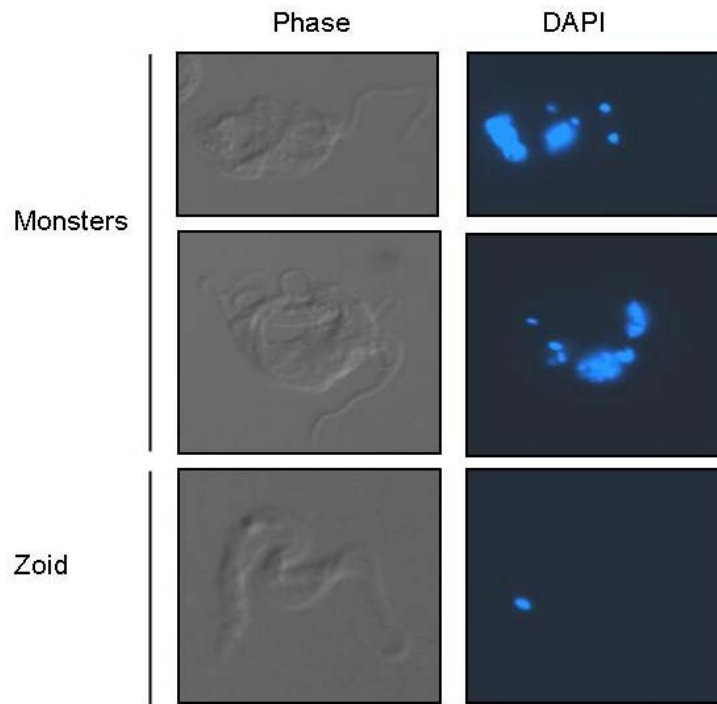
Figure 4.11 – Examples of 'other' cells in *brca2-/-* mutants. Each cell is shown in phase contrast (phase) and after staining with DAPI. The 'monsters' cells shown above are highly enlarged, and contain an elevated amount of nuclear and kDNA. An example of a cell with no nucleus, containing a single kinetoplast is also shown.
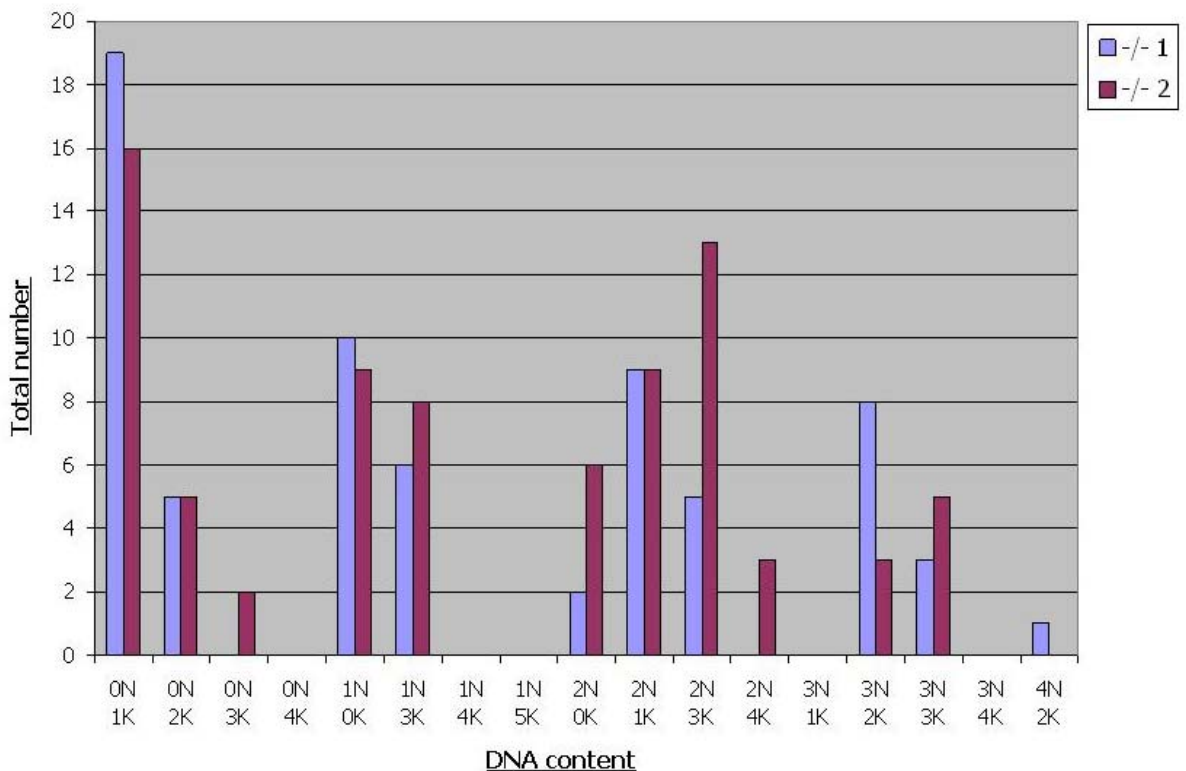
Figure 4.12 – DNA content of 'others' in *brca2-/-* mutants. Total numbers of cells that do not fit into the expected classifications (others) are represented for the *brca2-/-* mutants. The DNA content is displayed as the number of nuclei (N) and the number of kinetoplasts (K).

To investigate further whether the phenotype of *brca2-/-* cells is the consequence of a severe DNA repair deficiency, the DNA content of all the cell lines was analysed after DNA damage. To do this, cells were grown to a density of 1×10^6 cells.ml^{-1}, before

adding 1.0 µg.ml^{-1} of phleomycin and growing for a further 18 hours. The *brca2-/-* cells were also similarly treated with 0.25 µg.ml^{-1} of phleomycin, as they were predicted to be more sensitive to DNA damaging agents. After 18 hours of damage the cells were prepared for DAPI staining as before. The number of cells in each cell cycle stage was again counted (see table 4.6) and the percentages graphed (see figure 4.13).

Cell line	Cell cycle stage				Total
	1N 1K	1N 2K	2N 2K	Other	
Wild type 1.0µg.ml^{-1}	123	23	9	12	167
	92	14	6	3	115
	146	27	7	11	191
Total	361	64	22	26	473
+/-1 1.0µg.ml^{-1}	122	26	13	7	168
	163	22	7	9	201
Total	285	48	20	16	369
+/-2 1.0µg.ml^{-1}	133	24	9	7	173
	157	28	7	17	209
Total	290	52	16	24	382
-/-1 0.25µg.ml^{-1}	84	42	10	35	171
	105	48	14	31	198
Total	189	90	24	66	369
-/-1 1.0µg.ml^{-1}	53	45	6	33	137
	113	67	13	28	221
Total	166	112	19	61	358
-/-2 0.25µg.ml^{-1}	64	51	19	42	176
	102	54	12	35	203
Total	166	105	31	77	379
-/-2 1.0µg.ml^{-1}	82	63	11	41	197
	99	58	11	37	205
Total	181	121	22	78	402
rad51-/- 0.25µg.ml^{-1}	46	26	9	24	**105**
	98	48	15	36	**197**
Total	144	74	24	60	302
rad51-/- 1.0µg.ml^{-1}	65	37	3	35	**140**
	68	28	5	24	**125**
Total	133	65	8	59	265

Table 4.6 – DAPI analysis of the *BRCA2* mutants after DNA damage. The DNA content of *BRCA2* heterozygous (+/-) and homozygous (-/-) mutant cell lines were visualised by DAPI and compared with wild type Lister 427 cells and *rad51-/-* mutants, after cells had been damaged by phleomycin. The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications (others) were counted in separate experiments, for each cell line and are tabulated. Wild type and heterozygous cells were grown in media with 1.0 µg.ml^{-1} of phleomycin, whilst homozygous cells were grown in both 0.25 µg.ml^{-1} and 1.0 µg.ml^{-1} of phleomycin. Data are tabulated for 2 separate experiments in all cases.
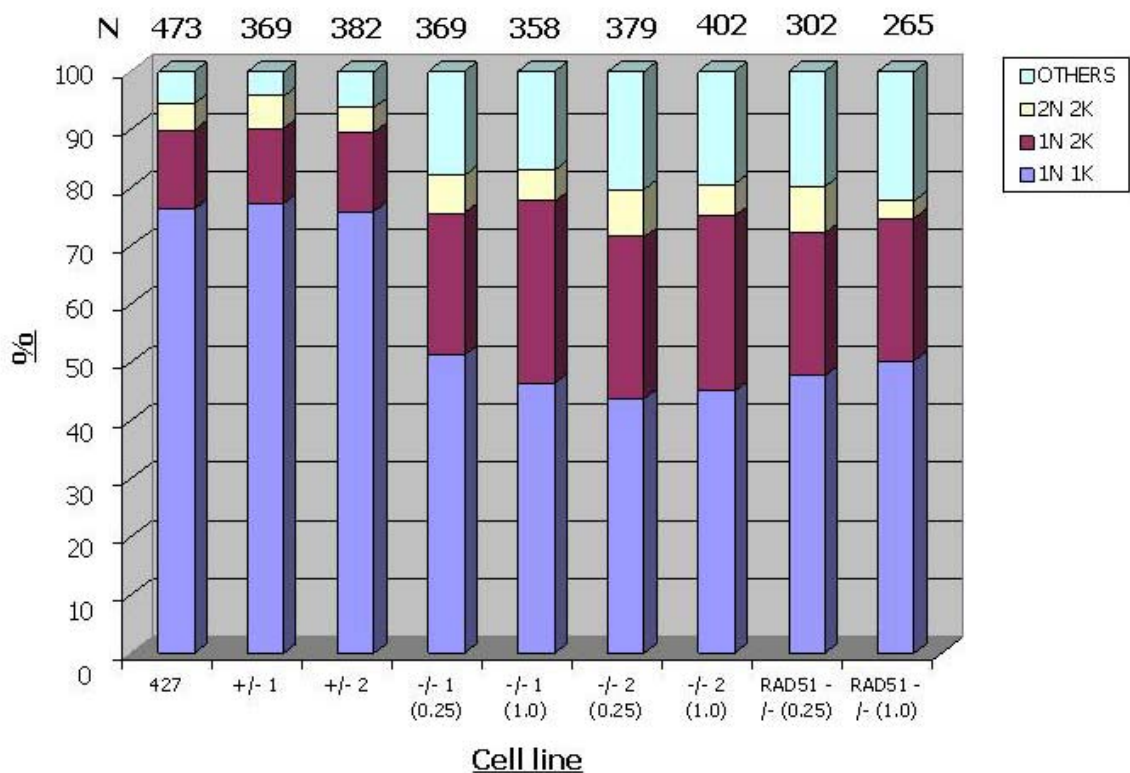
Figure 4.13 – DAPI analysis of the *BRCA2* mutants after DNA damage. The DNA content of *BRCA2* heterozygous (+/-) and homozygous (-/-) mutant cell lines were visualised by DAPI and compared with wild type Lister 427 cells and *rad51-/-* mutants, after cells had been damaged by phleomycin. The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications cells (others) were counted and represented by their mean count as a percentage of the total cells counted. Wild type and heterozygous cells were grown in media with 1.0 $\mu g.ml^{-1}$ of phleomycin, whilst homozygous cells were grown in media with 0.25 $\mu g.ml^{-1}$ and 1.0 $\mu g.ml^{-1}$ of phleomycin. N = number of cells counted.

From this cell cycle analysis, it was found, perhaps surprisingly, that induction of DNA damage by phleomycin in either wild type or *BRCA2+/-* cells did not alter the relative abundance of cells in different cell cycle stages. This indicates, contrary to findings in other organisms (Nakada *et al.*, 2003), that phleomycin does not induce any cell cycle stall, despite the fact that damage is generated, as evidenced by the generation of RAD51 repair foci (see section 4.3.6). In the *brca2-/-* mutants the presence of DNA damage caused dramatic cell cycle effects. A reduction in the number of cells with 1N 1K content was apparent, reducing from 76 % in wild type cells to 43-51 %. An increase in the number of 1N 2K cells was also seen, rising from 13 % in wild type cells to 24-31 %. Finally, there was also an increase in 'others', rising from 5 % in wild type cells to 17-20 % in homozygous cells. This was approximately double the number of these cells found in *brca2-/-* cells prior to damage. Only the number of 2N 2K cells appeared unaltered. Chi squared analysis was performed on these data comparing the relative distribution of cells into cell cycle stages. The results are shown in table 4.7 and display that the cell cycle distribution of both independent *brca2-/-* mutants are significantly different from the

distribution of cells in wild type and *BRCA2+/-* mutants, with chi squared values ranging from 39.31 to 91.3 at P = 0.0001.

A very similar phenotype was also observed in *rad51-/-* cells, with 47-50 % 1N 1K cells, 24-25 % 1N 2K cells and 19-22 % of 'others'. Chi-squared analysis showed the *rad51-/-* mutants were not significantly different from *brca2-/-* mutants in their distribution of cell types, but were significantly different from the wild type and *BRCA2* heterozygous cell line, with chi squared values of 52 to 94.72 at P = 0.0001. This suggests that the changes in relative numbers of cells at different stages of the cell cycle in the presence of DNA damage are a consequence of an absence of recombination and DNA repair that is common to BRCA2 and RAD51.

	+/- 1	+/- 2	-/- 1 0.25	-/- 1 1.0	-/-2 0.25	-/-2 1.0	rad51-/- 0.25	rad51-/- 1.0
WT	0.403 0.9396	0.159 0.9840	45.619 0.0001	59.346 0.0001	71.300 0.0001	68.395 0.0001	59.529 0.0001	689.533 0.0001
+/- 1		1.196 0.7539	61.234 0.0001	75.154 0.0001	91.302 0.0001	88.134 0.0001	78.207 0.0001	94.723 0.0001
+/- 2			39.312 0.0001	53.178 0.0001	63.364 0.0001	60.349 0.0001	52.001 0.0001	58.471 0.0001
-/- 1 0.25				2.664 0.4464	2.288 0.5147	2.376 0.4982	0.788 0.8524	2.952 0.3991
-/- 1 1.0					2.735 0.4344	0.415 0.9371	3.289 0.3492	4.358 0.2253
-/-2 0.25						1.182 0.7573	0.730 0.8662	4.735 0.1922
-/-2 1.0							2.334 0.5061	3.141 0.3704
rad51-/- 0.25								3.477 0.3238

Table 4.7– Statistical analysis of the cell cycle data for *BRCA2* mutants after DNA damage. Chi squared analysis of the cell cycle data after phleomycin induced DNA damage for wild type cells, *BRCA2* heterozygous mutants (+/-), homozygous mutants (-/-) and *rad51-/-* mutants. 0.25 and 1.0 indicate 0.25µg.ml^{-1} and 1.0 µg.ml^{-1} of phelomycin respectively. Wild type and heterozygous cell lines were only treated with 1.0 µg.ml^{-1} of phelomycin. The numbers indicated in bold represent the Chi squared value, whilst the numbers below represent the P value at which it was calculated. Areas shaded in yellow indicate a significant difference.

The increase in 1N 2K cells suggest that when *T. brucei* recombination proteins are mutated, and the cells are subjected to DNA damage, they become impaired in their ability to complete nuclear division, presumably because they struggle to repair the damage to their DNA and to replicate past lesions. The increase in 'other' cells, which was of equivalent magnitude in both *brca2-/-* and *rad51-/-* mutants following damage, suggests that these arise as a result of cytokinesis before the completion of repair and replication of nuclear DNA. This is consistent with lack of alteration of 2N 2K numbers, suggesting that the cells enter and leave M phase as normal in these conditions. However, because an

increase in 'others', albeit of a lesser magnitude was found only in *brca2-/-* mutants prior to damage, it was decided to quantify these observations in more detail by characterising the relative amount of nuclear and kDNA. These results are displayed in figure 4.14.
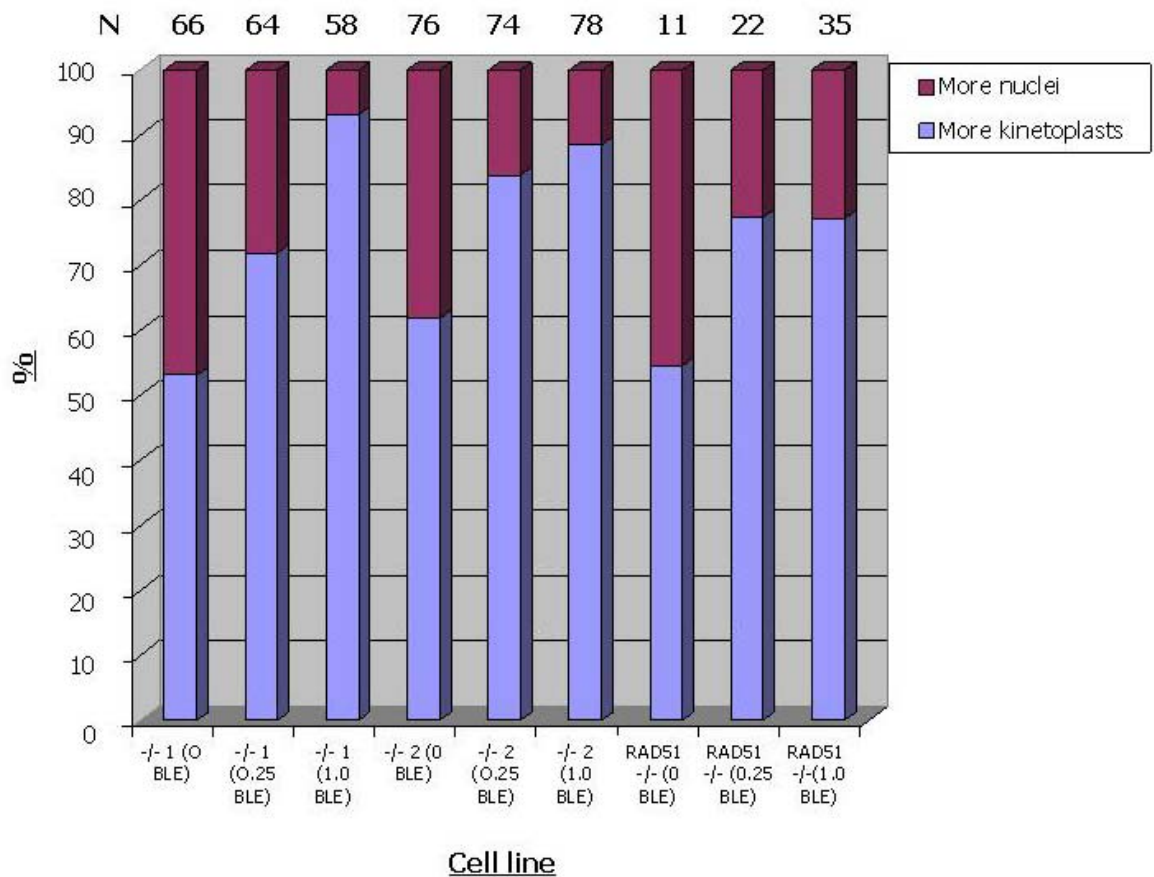
Figure 4.14 – Analysis of 'others' in *brca2-/-* and *rad51-/-* mutants before and after damage. The numbers of non-standard cell types in *brca2 -/-* and *rad51-/-* mutants are shown prior to damage and following 0.25 μg.ml^{-1} and 1.0 μg.ml^{-1} of phleomycin (BLE). The numbers of 'other' cells are split into 2 categories: cells containing more kinetoplast, or more nuclei, than normal. Cells with standard ratios of nuclei and kinetoplast are not shown. N = number of cells counted.

These data demonstrate that when *brca2-/-* cells were subjected to damage, those with an increased quantity of kinetoplast DNA dominated the 'other' cells. This phenotype became more exaggerated as phleomycin increased from 0.25 μg.ml^{-1} to 1.0 μg.ml^{-1}. The same phenotype was also observed in *rad51-/-* cells, confirming that this is likely to be due to an impairment in the repair of damage because of mutation to the recombination machinery. This suggests that when such mutant cells are subjected to DNA damage, the nuclear DNA is affected more strongly than the kinetoplast DNA. This is consistent with a role for each protein in nuclear repair, and suggests that damage to the kinetoplast DNA is repaired by a different route, or possibly monitored by different checkpoints. It seems unlikely that the kDNA would be unaffected by phleomycin treatment. Irrespective of this, cytokinesis appears to occur before nuclear DNA replication is completed, resulting in daughters with increased kDNA content. Importantly, the pattern of DNA content in

'other' cells in both *brca2-/-* and *rad51-/-* mutants before damage is quite different, with no strong bias towards increased nuclear or kinetoplast DNA. This infers that the generation of such mutants in the absence of induced damage is not primarily a result of loss of recombination factors. Since the number of these 'others' is increased in *brca2-/-* mutants, this infers that this phenotype is not DNA damage related, but is most likely due to a replication or cytokinesis dis-function that affect both the nuclear and kinetoplast DNA, and are exacerbated in *BRCA2* mutants.

Another observation offers further support for these conclusions. During cell counting, it was noticed that *brca2-/-* mutants, though they do not alter the number of 2N 2K cells, seemed to have a larger number of cells with 2N 2K content in which the nuclear DNA had not completed segregation. Examples of this are shown in figure 4.15. To quantify this in more detail, *brca2-/-, BRCA2+/-* and wild type cells were prepared for DAPI staining, and the number of 2N 2K cells with incompletely or completely separated nuclei counted blind by two researchers (see figure 4.16).
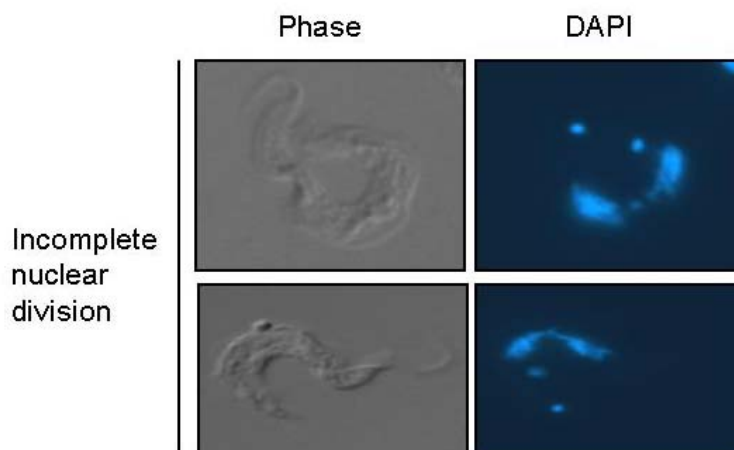
Figure 4.15 – Examples of 2N 2K cells with incomplete nuclear division. The cells shown to have incomplete nuclear division are not classes as 'other' cells, but 2N 2K cells and demonstrate a cell going through nuclear division before cytokinesis.

In wild type and *BRCA2+/-* cells, the majority of 2N 2K cells had completed nuclear division, with only 12 % having visibly connected nuclei. In contrast, mutation of both *BRCA2* alleles caused the percentage of 2N 2K cells that had not completed nuclear division to rise to 34 %. This infers that *brca2-/-* mutants take a greater amount of time to undergo nuclear division, which could be due to the *brca2-/-* mutants taking longer to complete DNA repair, therefore delaying mitosis, or could be a result of impaired replication or segregation.
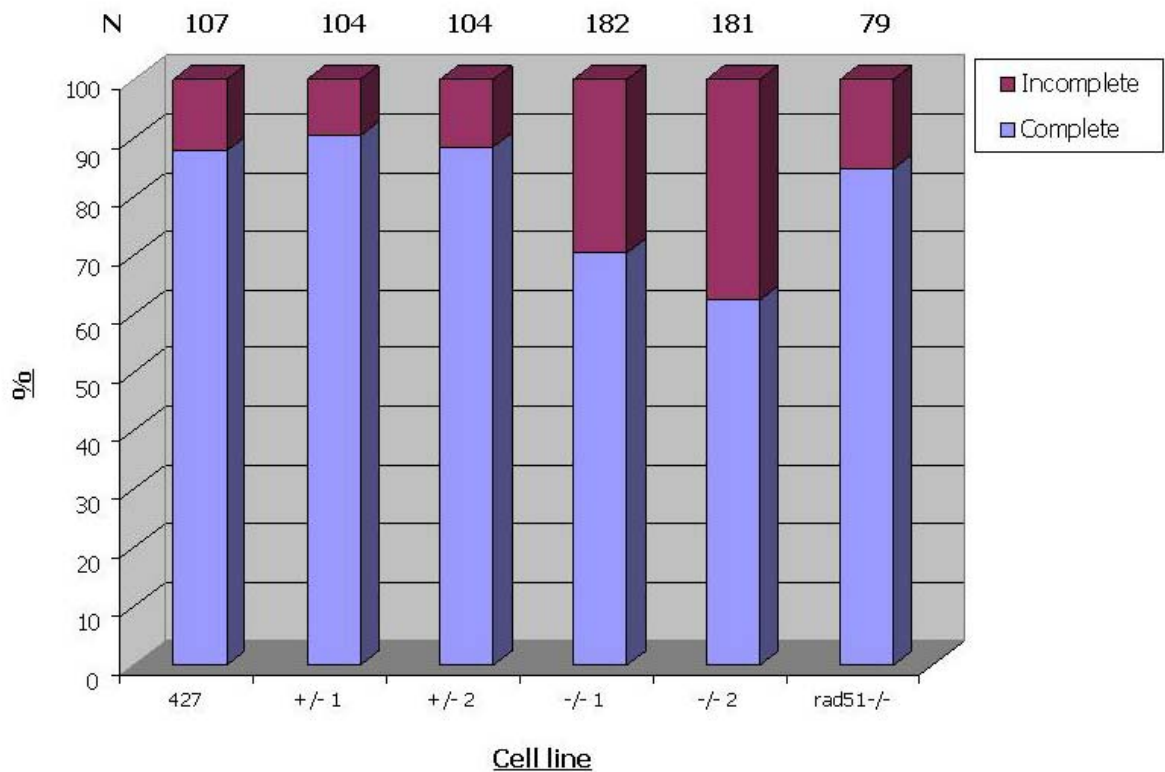
Figure 4.16 – Analysis of the number of 2N 2K cells that have completed nuclear division.
2N 2K cells in *BRCA2* heterozygous (+/-) and homozygous (-/-) mutant cell lines were visualised by DAPI and compared with wild type Lister 427 cells. The cells were analysed for the number that had completed nuclear division (complete), and those that were still dividing the nucleus (incomplete).

Examining all of the cell cycle results reveals that *brca2-/-* mutants have a putative delay in nuclear division and accumulate cells with aberrant DNA content, phenotypes that are not seen in other *T. brucei* recombination mutants, most notably *rad51-/-* cells. This appears consistent with the possibility that BRCA2 has a role beyond simply regulation of RAD51-catalysed recombination, in either the regulation or execution of cell division. We hypothesise that mutation of *BRCA2* in *T. brucei* causes impaired replication of the nucleus, but without a cell cycle stall, leading to the accumulation of chromosomal aberrations.

4.3.4 Analysis of DNA damage sensitivity

In figure 4.13 (section 4.3.3), it was shown that *brca2-/-* cells respond to phleomycin damage differently from WT or *BRCA2+/-* cells, inferring that BRCA2 acts in DNA repair. To analyse this role in more detail, DNA damage assays were carried out that allow quantification of the effect on the cells of damage by two agents: methyl methane sulphonate (MMS) and phleomycin. MMS is a methylation agent that is capable of modifying DNA at both guanine (generating 7-methylguanine residues) and adenine (generating 3-methyladenine residues), resulting in lethal and/or mutagenic lesions (Sedgwick, 2004);(Beranek, 1990). Phleomycin is a glycopeptide antibiotic of the

bleomycin family, which binds to and intercalates into DNA, destroying the integrity of the double helix (Giloni *et al.*, 1981), directly causing double strand breaks.

A clonal survival assay was initially used to examine the sensitivity of the *brca2-/-* mutants to MMS, as this assay had previously demonstrated that *rad51-/-* (McCulloch and Barry, 1999)*, rad51-3-/-* and *rad51-5-/-* mutants (Proudfoot and McCulloch, 2005) displayed a greater level of sensitivity to MMS compared to wild type cells. The clonal survival assay was performed by growing cultures to a density of 1×10^6 cells.ml^{-1} and plating out one cell per well over five 96 well plates, containing MMS concentrations of 0, 0.0001, 0.0002, 0.0003 or 0.0004 %. Four repetitions for each strain were carried out and the number of wells containing a viable parasite population after up to 20 days of growth was counted. The number of wells growing on the plate without MMS was taken as being 100% and the number of wells growing through on the MMS containing plates calculated relative to this, thereby removing any errors due to plating efficiency and growth rate differences between the *brca2-/-* cells and the others. The results of this assay are displayed in figure 4.17.
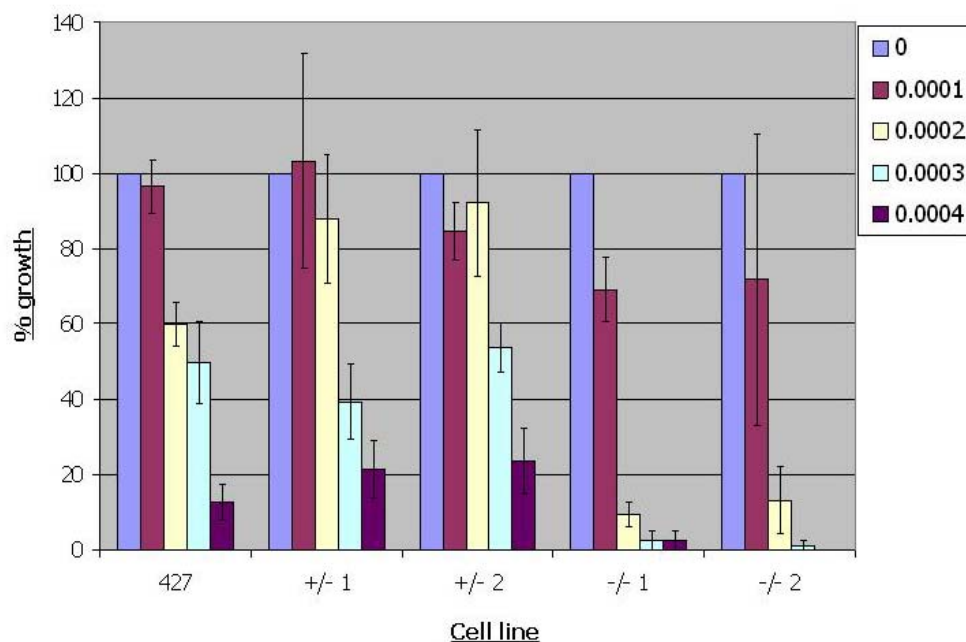
Figure 4.17 – Analysis of DNA damage sensitivity in the *BRCA2* mutants. Each strain was plated at one cell per well in five 96 well plates, each containing a different concentration of MMS: 0, 0.0001, 0.0002, 0.0003 and 0.0004 %. Four repetitions were carried out for each cell line. The mean number of cells to grow through for each cell line at each concentration was calculated and represented as a percentage of the number that had grown through on the 0 % plate for that cell line. Standard errors are indicated and the data is presented for the two independent heterozygous mutants (+/- 1, 2), the two independent homozygous mutants (-/- 1, 2) and the wild type Lister 427 cell line.

These data demonstrate that mutation of one *BRCA2* allele does not affect sensitivity to MMS. However, mutation of both *BRCA2* alleles causes an increased sensitivity to MMS, with very little or no growth occurring at 0.0003 % MMS and above. In contrast, the wild

type and heterozygous cells showed survival rates between 39 – 54 % at 0.0003 % MMS, and 12 – 24 % at 0.0004 % MMS. At 0.0001 % and 0.0002 % MMS the wild type and heterozygous cell lines displayed between 60 – 100 % survival, whilst the homozygous mutants displayed between 69 – 71 % and 9 – 13 % survival, respectively. Similar levels of sensitivity to MMS have been shown in *rad51-/-* mutants (McCulloch and Barry, 1999), and *rad51-3-/-* and *rad51-5-/-* mutants (Proudfoot and McCulloch, 2005), consistent with BRCA2 acting together with RAD51 and the RAD51 related proteins in DNA damage repair, though this does not address this directly.

In order to quantify the extent of *brca2-/-* mutants sensitivity to MMS and phleomycin, a distinct assay was used in which the metabolic capacity of the cells was measured *via* reduction of the compound Alamar Blue (Resazurin, Sigma). This assay allows the results of the clonal survival assay to be evaluated independently, and also allows IC50s to be calculated (Raz, B. *et al.*, 1997;Onyango, J. D. *et al.*, 2000). Reduction of Alamar Blue was examined by growing cultures to a density of 2×10^5 cells.ml^{-1} and placing 100 µl into 11 wells, each with doubly diluting concentrations of drug (either MMS or phleomycin). The final, 12th well acted as a control without drug. After 48 hours of growth, 20 µl of Alamar Blue was added. The plates were left for a further 24 hours for the cells to metabolise the resazurin, which is blue and non-fluorescent. In actively metabolising cells, resazurin is reduced to resorufin, which is pink and fluorescent (O'Brien *et al.*, 2000). This fluorescence was then measured on a Perkin Elmer LS55 Luminometer at 539 nm excitation and 590 nm emission. Three repetitions were performed and the IC50s calculated and their means plotted graphically (see figures 4.18 and 4.19).

The graph shown in figure 4.18 confirms the result from the clonal survival assay; mutation of one *BRCA2* allele does not affect sensitivity to MMS, but mutation of both *BRCA2* alleles causes an increased sensitivity to MMS. Wild type and heterozygous cell lines displayed mean IC50s to MMS of 0.0015 to 0.0019 %, whilst the homozygous mutants displayed mean IC50s of 0.00058 to 0.00064 %. *rad51-/-* mutants had very similar levels of sensitivity to the *brca2-/-* mutants, with a mean IC50 of 0.0007 % MMS. The statistical analyses of these data confirm these results (see table 4.8); no statistical difference was observed between wild type cells and *BRCA2+/-* mutants (p>0.05), but a statistically significant difference was found between wild type or heterozygous cells and *brca2-/-* mutants and *rad51* homozygous mutants (p<0.05). No significant difference in IC50 was found between the 2 *brca2-/-* cell lines and *rad51-/-* cells.

	+/- 1	+/- 2	-/- 1	-/-2	rad51-/-
WT	0.0881	0.0763	0.0008	0.0025	0.0023
+/- 1		0.4248	0.0016	0.0016	0.0005
+/- 2			0.0056	0.0067	0.0019
-/- 1				0.0959	0.4244
-/-2					0.6910

Table 4.8 – Statistical analysis of the Alamar Blue results for MMS. P values are shown for two sample T-tests comparing the IC50s for MMS sensitivity of wild type cells, *BRCA2* heterozygous mutants (+/-), *brca2* homozygous mutants (-/-) and *rad51-/-* mutants. Areas shaded in yellow indicate a significant difference.

When phleomycin was used to damage the cells, a similar phenotype to MMS sensitivity was observed (see figure 4.19); mutation of both *BRCA2* alleles caused an increased sensitivity to phleomycin. For this drug, wild type cells displayed a mean IC50 of 0.095 μM, whilst homozygous mutants displayed mean IC50s of 0.013 to 0.018 μM. The statistical analysis shown in table 4.9 confirms this; a statistically significant difference was seen between wild type cells and *brca2-/-* mutants ($p<0.05$). *rad51-/-* mutants again generated a similar IC50 to the *brca2-/-* mutants, with an IC50 of 0.019 μM, which was not significantly different from the *brca2-/-* mutants but was significantly different compared to wild type cells. In contrast to MMS, an unusual phenotype was observed when one *BRCA2* allele was mutated: a slight increase in sensitivity to phleomycin was observed, with mean IC50s of 0.06 to 0.067 μM, which were confirmed as being significantly different from both wild type and *brca2-/-* cells in paired T-tests. This phenotype is somewhat reminiscent of *mre11-/-* mutants, which displayed sensitivity to phleomycin but not to MMS (Robinson *et al.*, 2002). However, the BRCA2 haploinsufficiency observed here was not observed for MRE11. Nevertheless, mutation of a single allele of *BRCA2* presumably displays sensitivity to phleomycin and not to MMS due to the different modes of action of the DNA damaging agents, which is also reflected in MRE11 function. Phleomycin is known to directly cause DNA double strand breaks and is likely to be repaired via the homologous recombination pathway, whist MMS causes lesions in the DNA that only lead to DNA double strand breaks by further processing (Choy and Kron, 2002;Ui *et al.*, 2005). MMS-induced lesions can also be repaired *via* base excision repair (BER) (Lindahl and Wood, 1999), explaining why recombination mutants might be less sensitive to this form of damaging agent. It is thought that DNA incisions of MMS lesion by BER, when they are close enough, can cause DSBs. Equally, single strand breaks made by BER could be converted to DSBs during DNA replication. Unfortunately, previous research on *RAD51*, *RAD51-3* and *RAD51-5*, has only investigated sensitivity to MMS, and in the experiments shown in this section, only the *rad51* homozygous mutant was investigated, so it is impossible to know at this stage if mutation of a single allele in these

repair genes would display a similar phenotype to *brca2+/-* mutants with sensitivity to phleomycin and not to MMS.

	+/- 1	+/- 2	-/- 1	-/-2	rad51-/-
WT	0.0139	0.0023	0.0076	0.0065	0.0001
+/- 1		0.2476	0.0073	0.0041	0.0003
+/- 2			0.0151	0.0109	0.0008
-/- 1				0.0989	0.7867
-/-2					0.1623

Table 4.9 – Statistical analysis of the Alamar Blue results for phleomycin. P values are shown for two sample T-tests comparing the IC50s for phleomycin sensitivity of wild type cells, *BRCA2* heterozygous mutants (+/-), *brca2* homozygous mutants (-/-) and *rad51-/-* mutants. Areas shaded in yellow indicate a significant difference.

The *rad51-/-* mutants displayed statistically indistinguishable IC50s to the *brca2-/-* mutants for both MMS and phleomycin, which appears consistent with the hypothesis that BRCA2 has a similar role in DNA damage repair to RAD51, and that they may act together.
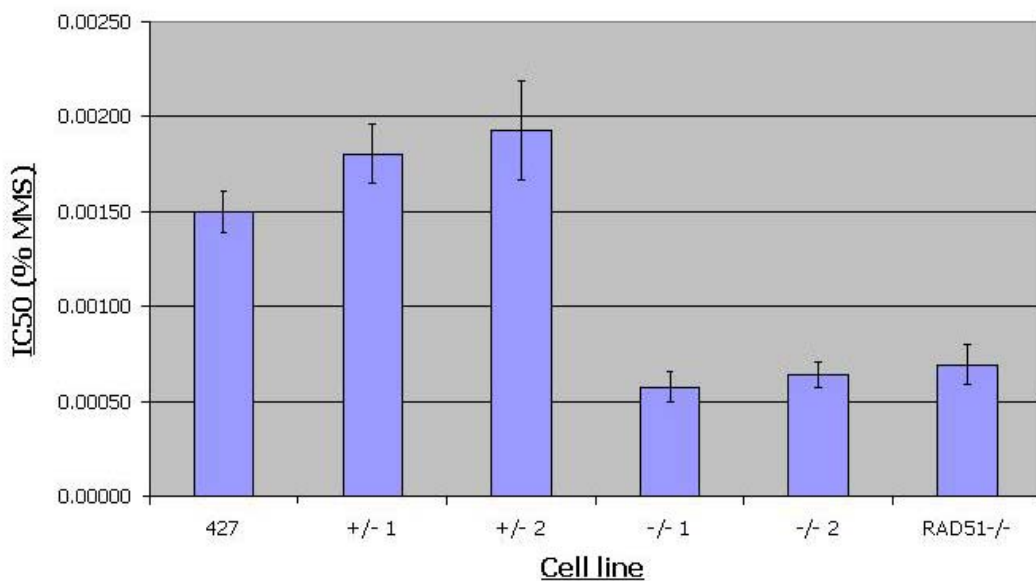
Figure 4.18 – IC50s of *T. brucei BRCA2* mutants exposed to MMS. Wild type, *BRCA2+/-*, *brca2-/-* and *rad51-/-* cell lines were placed in serially decreasing amounts of MMS and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments; bars indicate standard error.
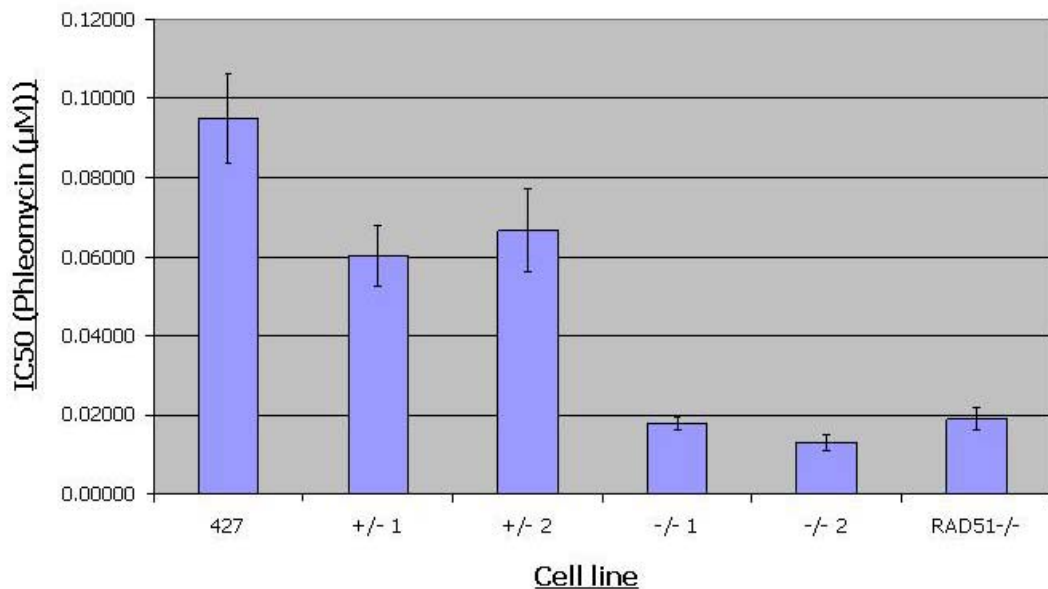
Figure 4.19 – IC50s of *T. brucei BRCA2* mutants exposed to phleomycin. Wild type, *BRCA2+/-*, *brca2-/-* and *rad51-/-* cell lines were placed in serially decreasing amounts of phleomycin and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments; bars indicate standard error.

4.3.5 Analysis of homologous recombination

To examine the function of BRCA2 in *T. brucei* recombination, a transformation assay was used. This involves the electroporation of a DNA construct containing an antibiotic resistance marker (in this case hygromycin) flanked by tubulin intergenic sequences into the cell lines. The construct is integrated into the tubulin array by homologous recombination, so that the number of transformants obtained should relate to the efficiency of recombination. In wild type cells, transformation of such constructs appears to always integrate using homologous recombination rather than other pathways, such as NHEJ (Conway *et al.*, 2002a;Conway *et al.*, 2002c).
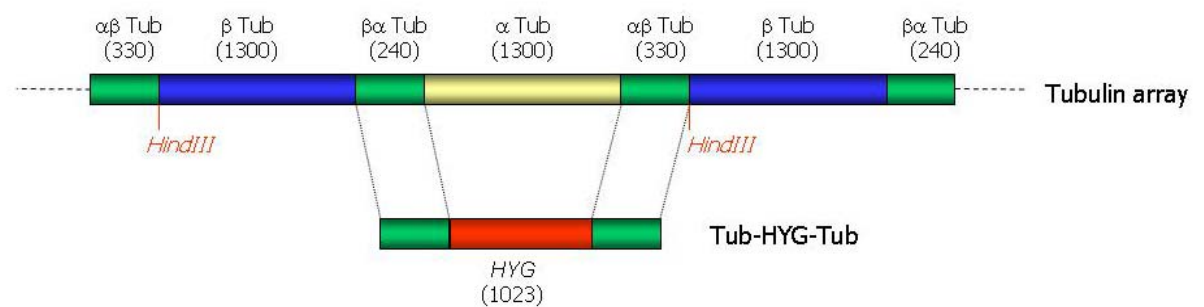
Figure 4.20 – Integration of the construct used in the recombination efficiency assay. The construct (Tub-HYG-Tub) contains tubulin (tub) intergenic sequences flanking an antibiotic resistance marker, and integrates into the tubulin array by homologous recombination, replacing an α tubulin sequence with the antibiotic marker. The size of β and α tubulin ORFs, the hygromycin phosphotransferase (HYG) ORF and intergenic sequences are shown (in bp).

The transformation constructs were excised from the plasmid Tub-HYG-Tub (R. McCulloch, gift) by *Xho*I and *Xba*I restriction digestion. These digestions were subsequently phenol: chloroform extracted and ethanol precipitated, before resuspending in sterile dH$_2$O. In each transformation, 5 x 10^7 cells were electroporated with 5 μg of construct DNA. The transformed cells were recovered for three generations before being plated out in selective media containing 5 μg.ml^{-1} of hygromycin. 5 x 10^6 cells were plated out over 24 wells for the wild type and heterozygous mutant cells, whilst 2 x 10^7 cells were plated out over 48 wells for the homozygous mutant cells, where fewer transformants were expected. The number of wells containing antibiotic resistant transformants were counted after 14 days and were expressed as the number of transformants per 10^6 cells plated out (see figure 4.21). For each cell line, the experiment was repeated on 3 independent occasions.
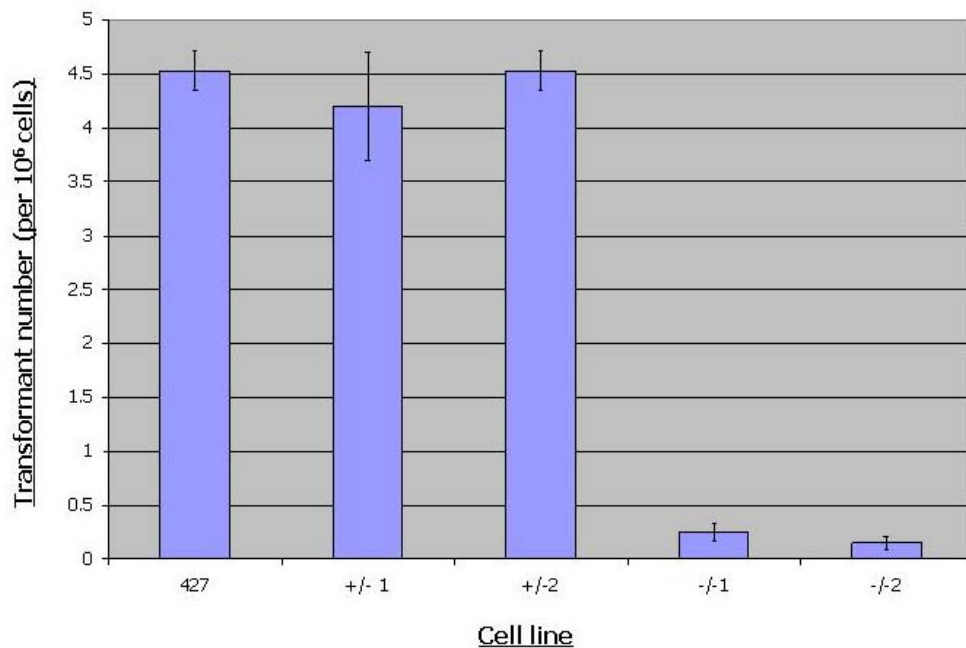
Figure 4.21 – Recombination efficiency in *BRCA2* mutants. Values are mean numbers of transformants obtained per 10^6 cells transformed; error bars are shown from 3 repetitions. The data are presented for two independent heterozygous mutants (+/-), two independent homozygous mutants (-/-) and wild type Lister 427 cells (427).
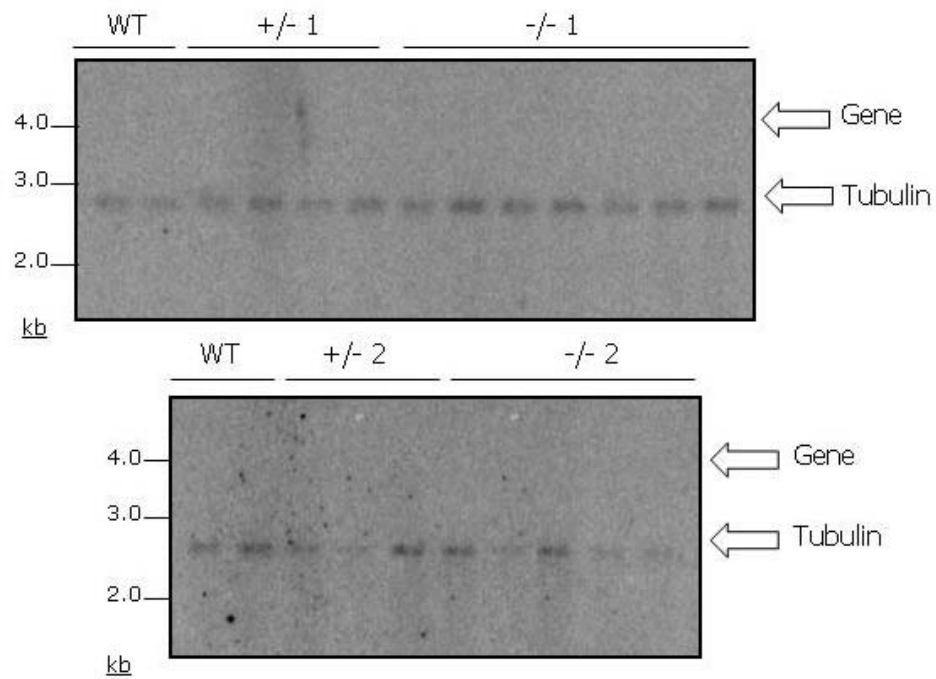
Figure 4.22 – Analysis of construct integration in *BRCA2* mutants. Two Southern blots are shown of Tub-HYG-Tub transformants digested with *Hin*dIII and probed with the *HYG* ORF. Transformants are shown for wild type (WT) cells, 2 independent *BRCA2* heterozygous mutants (+/-) and 2 independent *brca2* homozygous mutants (-/-). The expected fragment sizes of Tub-HYG-Tub integrated into the tubulin array (Tubulin) and the disrupted *BRCA2* gene (Gene) are indicated.

The results shown in figure 4.21 demonstrate that mutation of one *BRCA2* allele did not affect transformation efficiency. However, mutation of both *BRCA2* alleles caused a severe reduction in the ability of cells to generate transformants, with the transformation efficiencies reducing from an average of 4.5 x 10^{-6} in wild type cells to 0.2 x 10^{-6} in *brca2-/-* cells. This result is confirmed by the statistical analysis shown in table 4.10. No significant difference was seen between wild type cells and heterozygous mutants (P>0.05), whilst a significant difference was seen between wild type or heterozygous cells and homozygous mutants (P<0.05). A similar phenotype was observed for *rad51-/-* (Conway *et al.*, 2002c), *rad51-3-/-* and *rad51-5-/-* mutants (Proudfoot and McCulloch, 2005) and suggests that BRCA2 is important in the process of homologous recombination in *T. brucei*, since it is highly unlikely that any of these proteins affect the capacity of the cells to take DNA into the cytoplasm or nucleus.

	+/- 1	+/- 2	-/- 1	-/-2
WT	0.4444	1.0000	0.0037	0.0025
+/- 1		0.4226	0.0206	0.0184
+/- 2			0.0032	0.0028
-/- 1				0.3206

Table 4.10 – Statistical analysis of the recombination efficiency of *BRCA2* mutants. P values are shown for two sample T-tests comparing recombination efficiencies of wild type (WT) cells, *brca2* heterozygous mutants (+/-) and homozygous mutants (-/-). Areas shaded in yellow indicate a significant difference.

A number of transformants from each of the cell lines were examined by Southern analysis to determine the locus of the integrated construct. The construct should integrate into the tubulin array by homologous recombination (see figure 4.20). However, the mutated copy of the *BRCA2* gene contains the β-α intergenic sequence from the tubulin array acting as a processing signal (figure 4.1, section 4.2), identical to one flank of the *Tub-HYG-Tub* construct. It is therefore possible that this might mediate one-ended homologous recombination in the *BRCA2+/-* and *brca2-/-* cells. Wild type cells, in contrast, should only allow integration of the construct into the tubulin array by homologous recombination. 5 µg of genomic DNA from each cell line was restriction digested with *Hin*dIII, separated by electrophoresis on a 0.8 % agarose gel and transferred to a nylon membrane by Southern blotting. The blots were then probed with the hygromycin open reading frame (figure 4.22), demonstrating that all of the transformants analysed had integrated the constructs using two-ended homologous recombination, since all of the transformants had integrated the construct into the tubulin array. It is interesting to note that despite the greatly reduced transformation efficiency of the *brca2-/-* mutants, they could still perform homologous recombination, and no evidence for non-homologous mechanisms was revealed. This result is slightly different to that observed for *rad51-/-* mutants, where low levels of aberrant integrations were observed (Conway *et al.*, 2002c), and were mediated by sequence microhomology. It is more reminiscent of the exclusive use of homologous recombination events observed for *rad51-3-/-* and *rad51-5-/-* mutants (Proudfoot and McCulloch, 2005), though it cannot be excluded that such aberrant events would be revealed by screening greater numbers of transformants.

4.3.6 Analysis of RAD51 focus formation

In order to begin to examine the function of *T. brucei* BRCA2 in the regulation of RAD51 action, we wanted to examine the ability of *brca2-/-* mutants to form RAD51 foci after DNA damage. RAD51 foci have been shown to form in the nucleus of eukaryotes following DNA damage and at S phase (Tarsounas *et al.*, 2003;Tarsounas *et al.*, 2004). This has also been demonstrated to occur in *T. brucei* following damage (Proudfoot and McCulloch, 2005). Since BRCA2 has been shown to sequester RAD51 (Pellegrini and

Venkitaraman, 2004;Kojic *et al.*, 2005;Martin *et al.*, 2005;Tarsounas *et al.*, 2003) until it is needed for repair, it is hypothesised that in the absence of BRCA2 *in vivo*, cells will be unable to form RAD51 foci.

5 ml bloodstream stage cultures were diluted to a cell density of 1×10^6 cells.ml^{-1} before treatment with phleomycin. All cell lines were treated with 0 μg.ml^{-1}and 1.0 μg.ml^{-1} of phleomycin and grown for a further 18 hours. The *brca2-/-* mutants were also treated with an additional lower concentration of phleomycin (0.25 μg.ml^{-1}) due their sensitivity to this DNA damaging agent (see section 4.3.4). Following growth with phleomycin, 3 ml of the cultures were centrifuged, and treated following the immunofluorescence protocol described in section 2.11.2. The primary antibody used for these experiments was anti-RAD51 antiserum (rabbit polyclonal antiserum (Diagnostics Scotland) generated in response to His tag purified, *E. coli* expressed recombinant *T. brucei* RAD51 (supplied by K. Norrby)), which was diluted 1:500 in 3 % FBS/PBS. The secondary antibody used was Alexa 488 conjugated goat-derived anti-rabbit IgG (Molecular Probes, Invitrogen), and was diluted 1:1000 in 3 % FBS/PBS.

Fluorescence microscopic analysis was performed using an Axioskop 2 microscope (Zeiss) using DIC, UV and FITC filters. To quantify any effects displayed by the *brca2-/-* mutants, approximately 200 cells were scored in each cell line, for the number of sub-nuclear RAD51 foci that were visible (table 4.11). Images of representative cells following phleomycin treatment are shown in figure 4.24.

		Number of foci (%)						
	BLE	0	1	2	3	4	5	6 or more
WT	0.0	96.4	3.6	0.0	0.0	0.0	0.0	0.0
	1.0	24.8	22.6	18.8	16.5	13.5	2.3	1.5
+/- 1	0.0	94.2	5.8	0.0	0.0	0.0	0.0	0.0
	1.0	25.2	25.9	20.9	16.5	7.9	2.2	1.4
+/- 2	0.0	96.0	4.0	0.0	0.0	0.0	0.0	0.0
	1.0	19.9	24.8	19.9	19.1	10.6	3.5	2.1
-/- 1	0.0	97.9	2.1	0.0	0.0	0.0	0.0	0.0
	0.25	98.4	1.6	0.0	0.0	0.0	0.0	0.0
	1.0	99.4	0.6	0.0	0.0	0.0	0.0	0.0
-/- 2	0.0	98.2	1.3	0.4	0.0	0.0	0.0	0.0
	0.25	96.9	2.0	0.5	0.5	0.0	0.0	0.0
	1.0	98.6	0.7	0.7	0.0	0.0	0.0	0.0

Table 4.11 – RAD51 foci formation in wild type cells and *BRCA2* mutants. The percentages of cells showing foci at given concentrations of phleomycin (BLE) are shown. Phleomycin concentrations are shown in μg.ml^{-1}. Boxes without shading contain no foci, boxes shaded in light yellow contain foci and boxes shaded in bright yellow contain the highest percentage of foci.

From these results it can be seen that in wild type cells, without the presence of damage, RAD51 foci are rarely seen. However, once damage is induced, the number of cells with foci present is greater than those without (only ~ 25 % had no foci). Loss of a single allele of *BRCA2* had no effect on the cells' ability to form RAD51 foci; as for wild type, foci were rarely seen without damage and appeared in most cells after damage (19-25 % had none). Deletion of both *BRCA2* alleles, however, resulted in a greatly reduced ability to form RAD51 foci, with the majority of cells containing no foci, at any drug concentration. In fact, it is not clear that any induction of RAD51 foci occurred, as the very small percentage of cells in which foci were detected appears no different from untreated cells. These results demonstrate that mutation of BRCA2 impairs the visible concentration of RAD51 in foci following DNA damage.

To ensure that these results do not simply result from decreased RAD51 levels in the *brca2-/-* mutants, western analysis was carried out on total protein extracted from all *BRCA2* cell lines, before and after phleomycin-induced damage. Equivalent amounts of protein from cell extracts were separated on 10 % SDS-PAGE gels and probed with polyclonal anti-RAD51 antiserum and detected with HRP-coupled anti-rabbit IgG. Figure 4.23 demonstrates that RAD51 is still clearly expressed in *brca2-/-* cells, and there is no evidence for an increase in RAD51 levels after DNA damage.
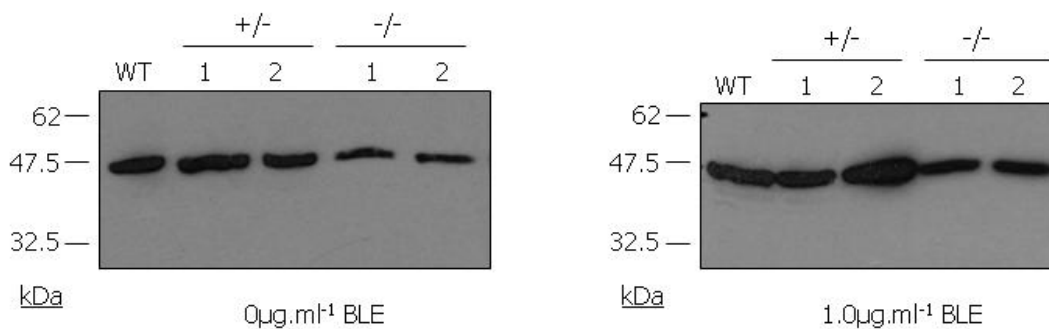
Figure 4.23 – Western blots of RAD51 in *BRCA2* cell lines. The western blots display total protein extracts from wild type (WT), BRCA2 heterozygous (+/-) and brca2 homozygous (-/-) cells probed with anti-RAD51 antiserum (RAD51). The blot on the left displays protein extracts prepared without damage ($0\mu g.ml^{-1}$ BLE), whilst the blot on the right displays protein extracts prepared with damage ($1.0\mu g.ml^{-1}$ BLE). Size markers are indicated. The endogenous copy of RAD51 is visible at 47kDa.
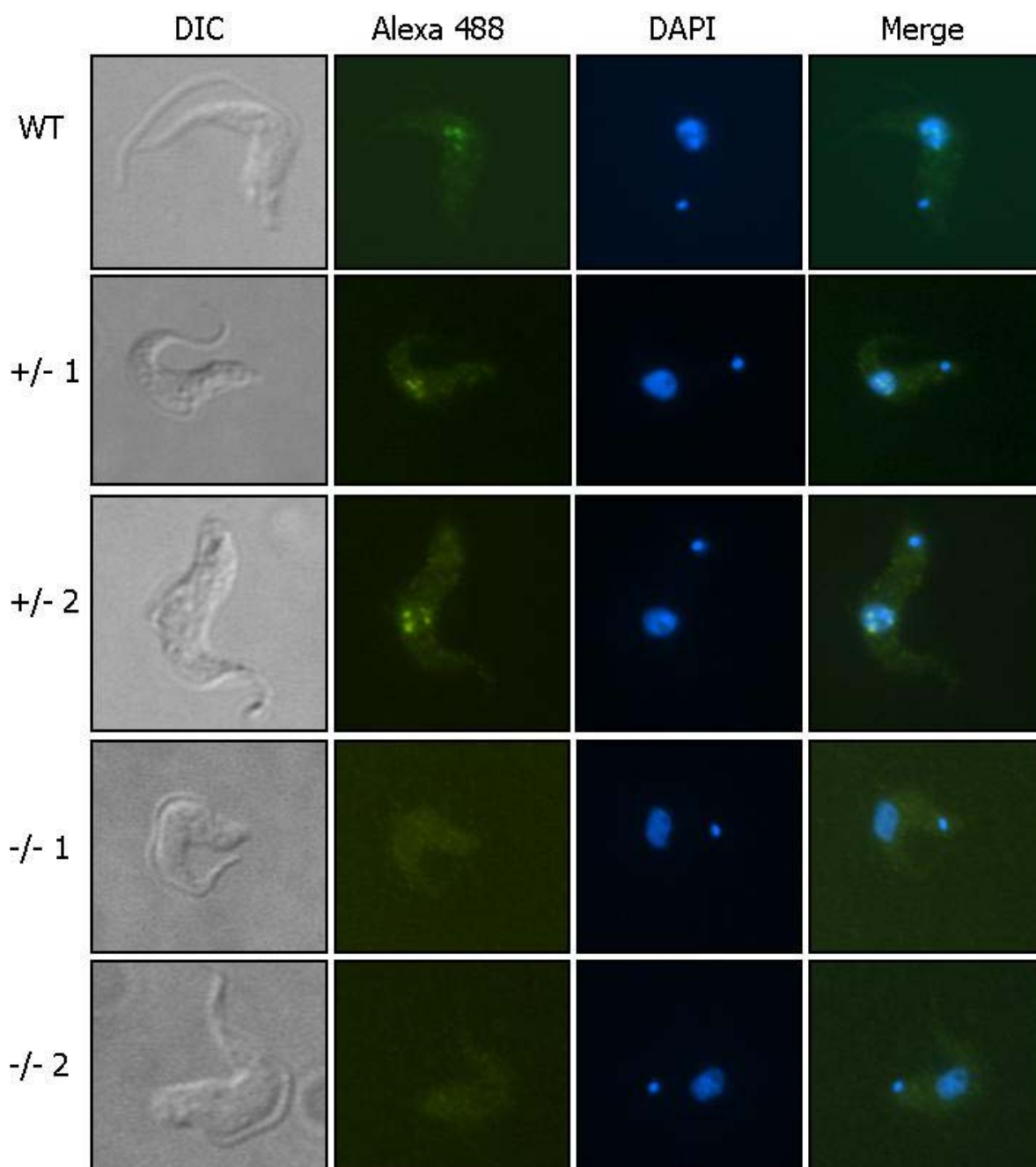
Figure 4.24 – RAD51 immunolocalisation in wild type cells and *BRCA2* mutants.
Representative images of T. brucei cells following growth in 1.0µg.ml^{-1} phleomycin for 18 hours are shown. Each cell is shown in differential interface contrast (DIC), after staining with DAPI and after hybridisation with anti-RAD51 antiserum and secondary hybridisation with Alexa Fluor 488 conjugate (Alexa 488). Merged images of DAPI and Alexa 488 cells are also shown. Wild type Lister 427 (WT) cells, heterozygous *BRCA2* mutants (+/-) and homozygous mutants (-/-) are shown.

4.3.7 Analysis of VSG switching frequency

A VSG switching assay was carried out in order to determine if the absence of BRCA2 has any effect on antigenic variation. Since *brca2-/-* mutants have been shown to have an impaired ability to perform homologous recombination (section 4.3.5), and a large amount of VSG switching is thought to occur by homologous recombination, it was hypothesised that cells lacking BRCA2 would be impaired in their ability to switch their VSG coat. However, the measurement of recombination efficiency employed an *in vitro* assay, which may not reflect conditions *in vivo,* which we used to assay VSG switching.

Analysis of VSG switching uses an assay that has been described before (McCulloch and Barry, 1999;Robinson *et al.*, 2002;Bell and McCulloch, 2003;Proudfoot and McCulloch, 2005;Proudfoot and McCulloch, 2006) and employs the cell line 3174.2, which contains a modified active VSG expression site containing antibiotic resistance markers for hygromycin and G418 (see figure 4.25). These antibiotic resistance markers not only allow VSG switching frequency to be determined, but also allow VSG switching mechanisms to be determined. The experimental procedures for this assay are described in section 2.1.8. In essence, growth *in vitro* on hygromycin and G418 ensured that all cells expressed the modified expression site and therefore the VSG 221 protein. *T. brucei* cells grown on the antibiotics were used to infect mice, which were then cured of *T. brucei*, generating mice that are immune to VSG 221. The immunised mice were then used to cure unswitched *T. brucei* cells from wild type and *BRCA2* mutant cell populations.
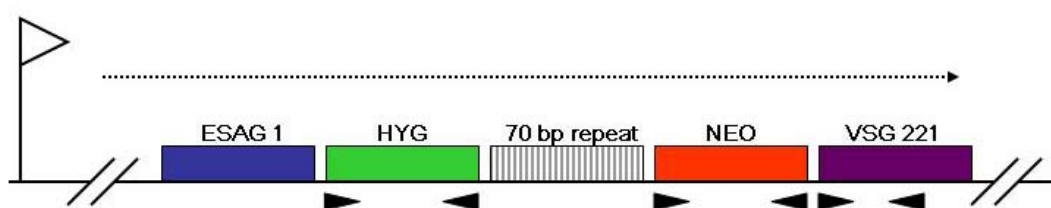
Figure 4.25 – The expression site of the trypanosome cell line 3174.2. This figure shows the telomeric region of the actively transcribed expression site in 3174.2 bloodstream form cells. The expression site has been modified to contain the resistance markers for hygromycin (*HYG*), between ESAG 1 and the 70 bp repeats, and G418 (*NEO*), between the 70 bp repeats and the *VSG 221* gene. The dashed arrow represents transcription of the site, whilst the black triangles represent oligonucleotide primer sites.

To generate switched variants, all the cell lines were grown on hygromycin and G418, before removing them from antibiotic selection and allowing them to grow for nine generations. For the wild type and heterozygous cell lines 4×10^7 cells were injected into immune mice, whereas for the homozygous mutants, where a switching defect was expected, 8×10^7 cells were injected into immune mice in an attempt to increase the

number of switching events per mouse, thereby ensuring higher accuracy. In each case, multiple, independently grown switched populations were injected into the mice, and in all cases 24 hours after injection the surviving trypanosomes were recovered from the mice and plated out over 96 well plates. The number of wells that showed growth after a maximum of 4 weeks were counted and used to calculate the VSG switch frequency, taking the reduced growth rate of the *brca2-/-* cells into account. VSG switching frequency was calculated for each cell line, in each mouse, by multiplying the number of wells showing growth by 2.5 in order to calculate the number of switched variants in the total blood volume, since 2 x 0.4 ml of blood was used to isolate switched variants and the total blood volume of a mouse is assumed to be 2 ml. This was then divided by the number of generation times that had occurred during the 24 hour period following injection into the mouse. Finally, the number of cells that were injected were taken into account, thereby producing the number of switched variants per 10^7 cell injected. The results of this analysis are shown in table 4.13 and figure 4.26. This analysis shows that *brca2-/-* mutants were greatly impaired, relative to the *BRCA2* heterozygous and WT cells, in their ability to switch their VSG coat. WT cells underwent VSG switching at approximately 8 events per 10^{-7} cells, and the *BRCA2+/-* cells switched at approximately 11-12 events per 10^{-7} cells. In contrast, VSG switched variants arose at only approximately 1-2 events per 10^{-7} cells in the *brca2-/-* mutants. Statistical analysis was performed on these results (table 4.12), demonstrating that significant differences were found between the heterozygous mutants and homozygous mutants ($p<0.05$). In contrast, no significant difference was noted between wild type cells and the *brca2* homozygous mutants; this is most likely due to the high variability in the wild type experiment. The extent of this phenotype, and the fact that residual VSG switching still occurs, is highly reminiscent of observations in *rad51-/-* cells (McCulloch and Barry, 1999) and highlights the importance of homologous recombination in the VSG switching mechanism.

	+/- 1	+/- 2	-/- 1	-/-2
WT	0.4143	0.2025	0.1828	0.1220
+/- 1		0.8405	0.0064	0.0002
+/- 2			0.0192	0.0092
-/- 1				0.3955

Table 4.12 – Statistical analysis of the VSG switching frequencies in the *brca2* mutants. P values are shown for two sample T-tests comparing VSG switching frequencies of wild type cells, *brca2* heterozygous mutants (+/-) and homozygous mutants (-/-). Areas shaded in yellow indicate a significant difference (P<0.05).

Cell line	No. of wells	Switch frequency	Average	St Dev	SE
	(out of 192)	(per 10^7 cells)			
Wild type	76	5.94			
	52	4.06			
	184	14.38	**8.13**	5.49	3.18
+/- 1	142	11.09			
	149	11.64	**11.37**	0.39	0.27
	D				
+/- 2	142	11.09			
	125	9.77			
	181	14.14	**11.67**	2.24	1.30
-/- 1	31	3.42			
	1	0.11			
	15	1.66	**1.73**	1.66	0.96
-/- 2	1	0.11			
	4	0.44			
	10	1.10	**0.55**	0.51	0.29
	0				
	0				

Table 4.13 – Determining the VSG switching frequencies of *BRCA2* mutants. The numbers of wells (from 2 96 well dishes) containing growing *T. brucei* populations are shown for each of the switching assays carried out on the wild type 3174.2 cells, heterozygous (+/-) and homozygous (-/-) mutants. The switch frequencies, calculated from these values, and standard deviation (St Dev) and standard error (SE) are shown. D indicates death of a mouse, and 0 indicates no growth. The experiments with no growth were not included in calculating the mean switch frequency.
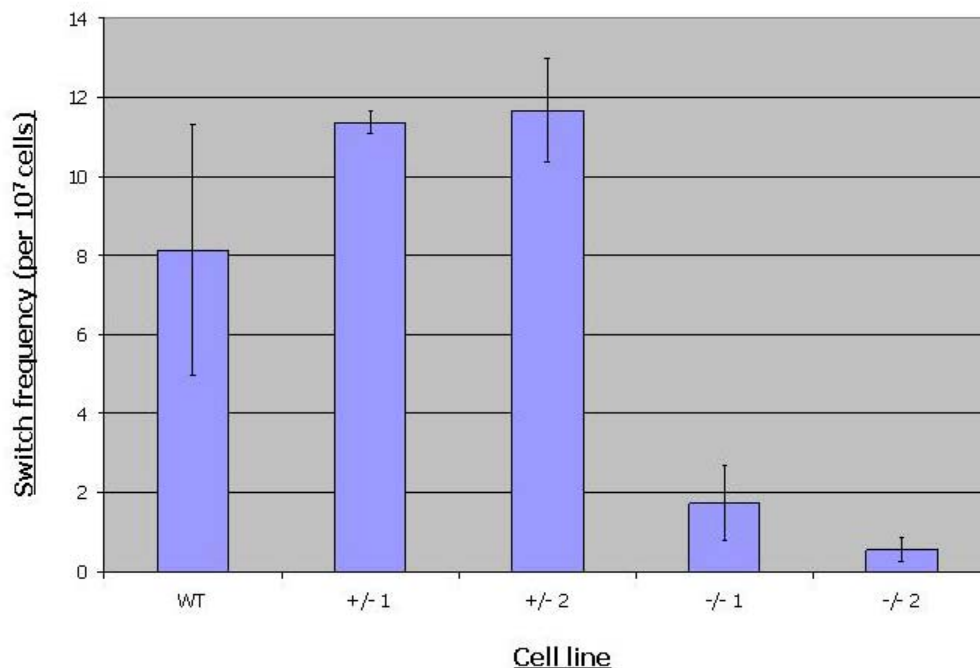
Figure 4.26 – VSG switching frequencies in *BRCA2* mutants. Values shown are the average switching frequencies for 3174.2 wild type cells (WT) and the 2 independent (1 and 2) *BRCA2* heterozygous (+/-) and homozygous (-/-) mutants. Data are from at least 2 experiments and standard error is indicated by bars.

4.3.8 Analysis of VSG switching mechanism

The 3174.2 cell line allows three different types of switching mechanism to be determined in each clonal switch variant recovered from the 96 well plates. These are *in situ* transcriptional switching, VSG gene conversion switching and expression site gene conversion switching (see figure 4.27). Cells that have performed an *in situ* switch transcriptionally inactivate the marked *VSG221* ES and activate another expression site, which results in the cells being sensitive to hygromycin and G418 (Hyg^S G418^S), as the antibiotic resistance genes are no longer expressed. PCR products for each of the antibiotic resistance genes and *VSG221* can, however, still be obtained (Hyg$^+$, Neo$^+$, 221$^+$), since no loss of sequence occurred. Cells that have performed a VSG gene conversion replace sequence from the 70-bp repeats to the 3' end of *VSG221* with VSG sequence from another expression site from the mini-chromosomes or from silent array VSGs, resulting in cells which are resistant to hygromycin, but sensitive to G418 (Hyg^R G418^S). In these switches, PCR products can only be obtained for the hygromycin marker (Hyg$^+$, Neo$^-$, 221$^-$). Cells that have performed an expression site gene conversion replace all the unique markers in the expression site, since they convert VSG sequence by long range gene conversion from another ES upstream of the 70 bp repeats, resulting in cells which are sensitive to hygromycin and G418 (Hyg^S G418^S), and from which no PCR products can be obtained (Hyg$^-$, Neo$^-$, 221$^-$). In fact, such switches may simply have deleted much of the ES sequence, and have then activated transcription from another ES, as such events have been described (Cross *et al.*, 1998; Rudenko *et al.*, 1998), though appear to be less frequent than gene conversions (McCulloch *et al.*, 1997).

For the wild type cells and each of the 2 *BRCA2+/-* cell lines, ten switched variants from each cell line and from each mouse were analysed to determine the switch mechanism. For the *brca2-/-* cell lines, all switched variants that were generated were analysed. The antibiotic resistance of each switched variant was scored using 5.0 μg.ml^{-1} hygromycin or 2.5 μg.ml^{-1} neomycin. PCR-amplification of the antibiotic resistance markers and the *VSG221* gene were used to determine the presence or absence of each of the genes in the expression site. Intact genomic DNA was confirmed in all cases using primers recognising RNA polymerase I as a positive control PCR. Taking the results from the drug sensitivities and PCR production together, the switching mechanisms could be determined, and are shown in table 4.14 and figure 4.28.

This analysis demonstrated that mutation of one *BRCA2* allele had little effect on the relative ratio of switching mechanism utilised. Indeed, the absence of both *BRCA2* alleles

also appeared to have no significant effect. It might have been predicted that an absence of BRCA2 would cause the *in situ* transcriptional switching mechanism to be primarily utilised, since an absence of BRCA2 leads to an impaired ability to perform homologous recombination and there is no evidence that recombination is involved in *in situ* switching. For one of the *brca2-/-* mutants (*-/-1*), such a situation may be seen. However, this result was not replicated for the other homozygous mutant, and since a similar pattern was seen in *BRCA2+/-1*, it seems likely that this simply reflects variation in sampling of the small numbers of variants examined. One conclusion that can be drawn from these data, however, is that homologous recombination events (ES and VSG gene conversion) were still seen to occur for both *brca2-/-* mutants. This correlates with the results of the recombination efficiency assay (section 4.3.5), whereby homologous recombination-mediated integration of *tubHYGtub* was observed in the absence of BRCA2.

From these results we can therefore conclude that VSG switching primarily occurs by homologous recombination, since defects in this pathway (through mutation of BRCA2, RAD51 or RAD51-3) substantially decreases the ability of *T. brucei* to switch its VSG coat. An important observation is to note that, although reduced, homologous recombination can still occur when proteins involved in this pathway are removed. This leads us to hypothesise that there is a RAD51, BRCA2 and RAD51-3 - independent recombination present at low levels in the cell, and that this may, in fact, be a single back-up mechanism. It seems likely, therefore, that this pathway or pathways that allows the cell to perform recombination in the absence of much of the homologous recombination machinery, also acts in VSG switching.

Figure 4.27 – Types of switching mechanism utilised by *T. brucei*.
Three different switching mechanisms can be distinguished in the 3174.2 trypanosome strain. During an *in situ* transcriptional switch, the *VSG221* ES is inactivated and an inactive expression site containing an unknown VSG (X) is transcriptionally activated. Cells that have undergone this switching mechanism are sensitive to hygromycin (Hyg^S) and G418 (G418^S), and each unique gene in the ES can be amplified by PCR (Hyg+ Neo+ 221+). In an expression site gene conversion event, all the markers are replaced by sequence from another expression site containing an unknown VSG. Cells that have undergone this switching mechanism are sensitive to hygromycin (Hyg^S) and G418 (G418^S), and none of the cassettes can be amplified by PCR (Hyg- Neo- 221-). In a VSG gene conversion event, the sequence from the 70 bp repeats is replaced by sequence from another expression site containing an unknown VSG. Cells that have undergone this switching mechanism are resistant to hygromycin (Hyg^S) and sensitive to G418 (G418^S), the hygromycin cassette can be amplified by PCR, whilst neomycin and VSG 221 cannot (Hyg- Neo- 221-).

	ES GC	VSG GC	In situ	Other	Total
WT	7	2	1	0	**10**
	7	1	2	0	**10**
	0	0	8	2	**10**
Total	**14**	**3**	**11**	**2**	30
+/- 1	3	0	5	2	**10**
	2	1	7	0	**10**
Total	**5**	**1**	**12**	**2**	20
+/- 2	6	0	4	0	**10**
	7	0	3	0	**10**
	6	1	3	0	**10**
Total	**19**	**1**	**10**	**0**	30
-/- 1	4	1	22	0	**27**
	0	0	1	0	**1**
	2	2	5	0	**9**
Total	**6**	**3**	**28**	**0**	37
-/- 2	1	5	4	0	**10**
	0	0	1	0	**1**
	4	0	0	0	**4**
Total	**5**	**5**	**5**	**0**	15

Table 4.14 – VSG switching mechanisms in *BRCA2* mutants. The type of switching mechanism used in a number of putatively clonal switched variants from wild type cells and *BRCA*2 heterozygous (+/-) or homozygous (-/-) mutant cell lines are indicated. In situ: in situ transcriptional switch; ES GC: expression site gene conversion; VSG GC: VSG gene conversion; Other: unknown mechanism.
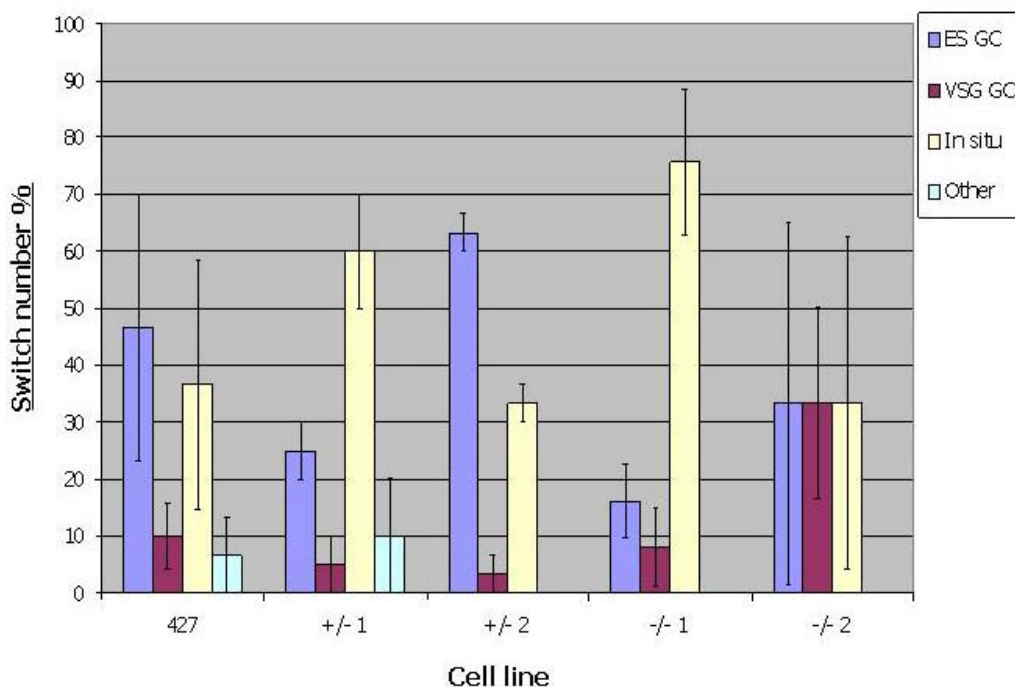
Figure 4.28 – Analysis of switching mechanism in the *brca2* mutants. Values indicate the mean type of switching mechanism used in a number of putatively clonal switched variants, for 3174.2 wild type cells (WT) and the 2 independent (1 and 2) *BRCA2* heterozygous (+/-) and *brca2* homozygous (-/-) mutants are shown. Standard errors are indicated. In situ: in situ transcriptional switch; ES GC: expression site gene conversion; VSG GC: VSG gene conversion; Other: unknown mechanism.

4.3.9 Analysis of genomic stability

BRCA2 has been shown to have a critical role in the maintenance of genome integrity in both mammals and *U. maydis* (Moynahan *et al.*, 2001;Kraakman-van der Zwet *et al.*, 2002;Tutt *et al.*, 2001;Kojic *et al.*, 2002;Patel *et al.*, 1998;Kojic *et al.*, 2002). To examine if this is also true in *T. brucei* BRCA2, the molecular karyotypes of the *brca2-/-* mutants were examined after prolonged growth and compared with WT and *BRCA2+/-* cells. All cell lines were cloned, by plating out one cell per well over 96 well plates, grown *in vitro* for approximately 290 generations, re-cloned and genomic DNA prepared from a number of clones for Southern analysis and pulsed field gel electrophoresis (PFGE). Genomic DNA was prepared from four clones for WT cells (1-4), 3 for *BRCA2+/-1* (5-7), 6 for *brca2-/-1* (8-13), 3 for *BRCA2+/-1* (14-16) and 6 for *brca2-/-1* (17-22)

Genomic instability was initially investigated by examining the *VSG121* gene family, which in wild type cells consists of five genes, one of which is telomeric and the others are likely to be present in subtelomeric arrays (Robinson *et al.*, 2002). Genomic DNA from clones of all cell lines was digested with *Xmn*I, separated by electrophoresis on a 0.8 % agarose gel and transferred to a nylon membrane by Southern blotting. The blots were then probed with the *VSG121* ORF, the results of which are shown in figure 4.29. As expected, wild type and *BRCA2+/-* clones produced five distinct hybridising fragments, four of which were of constant size, as they are present in unaltered loci, and one telomeric copy, which lies in an inactive BES (Liu *et al.*, 1985) and varies in size depending on differences in telomere tract length. In all but one of the *brca2-/-* clones, at least one non-telomeric *VSG121* gene copy had been lost, with clone 18 appearing to have lost two copies. This indicates the presence of chromosomal rearrangements in *brca2-/-* cells, which result in the loss of non-essential genetic material. It is interesting to note that the telomeric *VSG121* copy was never lost, suggesting that such chromosomal rearrangements infrequently affect telomeres.
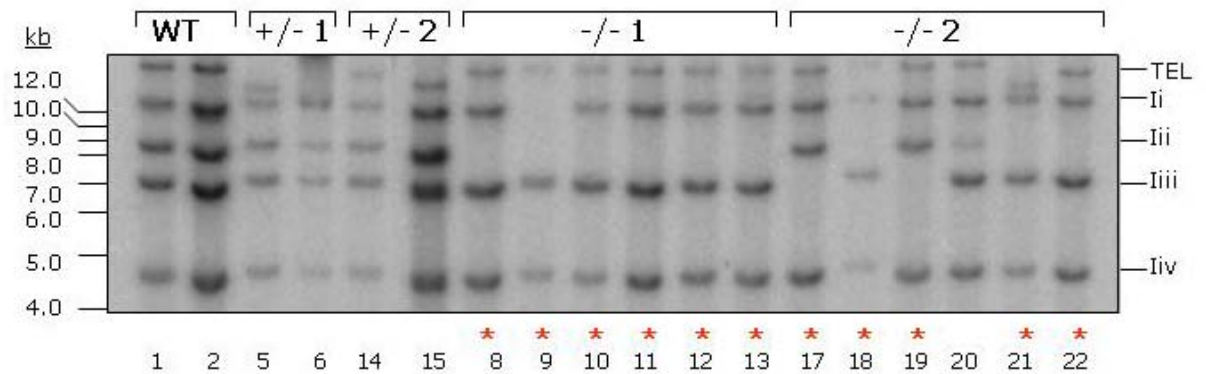
Figure 4.29 – *VSG121* **gene deletions resulting from BRCA2 inactivation.** Genomic DNA from wild type (WT), *BRCA2+/-* and *brca2-/-* clones was digested with *Xmn*I, separated in 0.8 % agarose and Southern blotted. The membrane was probed with DNA specific to the *VSG121* ORF. One telomeric (TEL) and four which are likely to be present in subtelomeric arrays (Ii, Iii, Iiii and Iiv) are shown. Clones 1-2: wild type; clones 5-6: *BRCA2+/-1*; clones 14-15: *BRCA2+/-2*; clones 8-13: *brca2-/-1*; clones 17-22: *brca2-/-2*. * indicates clones in which an internal gene copy has been lost. The clones are the same clones used in figure 4.30.

To investigate these rearrangements in *brca2-/-* mutants further, PFGE was performed to separate intact chromosomes of *T. brucei*. Genomic plugs were prepared from the same set of clonally derived trypanosomes (as described in section 2.2.4). The DNA was then separated on the CHEF-DR III system (Bio-Rad), using one genomic plug per lane in a 1.2 % agarose gel, in 0.089 M Tris borate, 0.1 mM EDTA (TB(0.1)E) at 85 V, 1400 seconds to 700 seconds pulse time, for 144 hours at 15 °C. Firstly, the DNA was visualised by staining with ethidium bromide to examine the karyotype. It was then Southern blotted and probed, sequentially for the genes *VSG121*, *VSG221* and glucose-6-phosphate isomerase (*GPI*) to gain a picture of rearrangements associated with specific genes and chromosomes. From the ethidium-bromide staining (figure 4.30), it is apparent that the mutation of one *BRCA2* allele has little, pronounced effect on genomic stability. The karyotypes from clones derived from one of the independent heterozygous mutants (*BRCA2+/-1*) appeared to be very similar to wild type cells. However, 3 of the clones derived from *BRCA2+/-2* (clones 14-16) appeared to display the apparent loss of a chromosome approximately 2.5 Mb in size. In fact, there appears to be progressive shortening of the megabase chromosomes in both heterozygous mutants. In this gel, the clearest resolution of the megabase chromosomes are the 4 – 5 bands ranging in size from ~ 1.7 to 2.4 Mb. Some shortening of at least the 3 larger bands can be seen in most of the *BRCA2+/-* clones relative to WT clones. This *BRCA2+/-* phenotype appears consistent with haploinsufficiency seen in measuring sensitivity to phleomycin (section 4.3.4) (comparing clones 5-7 and 14-16 with 1-4). The ethidium staining of the PFGs shows that mutation of both *BRCA2* alleles generates gross chromosomal rearrangements (GCRs). Clones derived from one of the independent homozygous mutants (*brca2-/-1*; clones 8-13) display a continuation of the progressive shortening of chromosome length seen in the *BRCA2+/-*. In addition, loss of some chromosome bands is apparent; for instance, in clone

9, a chromosome of approximately 2.2 Mb is no longer visible. Clones derived from the second independent homozygous mutant (*brca2-/-2*; clones 17-22) indicate a more dramatic display of GCRs, a highly dispersed karyotype relative to the WT and *BRCA2+/-* clones. Again, this is apparent as loss of some chromosome bands, though again the trend is suggestive of chromosome shortening, this being most apparent in clones 17 and 19.

Probing for *VSG121*, *VSG221* and *GPI* is displayed in figure 4.31, and revealed that the GCRs in *brca2-/-* mutant clones resulted primarily from a reduction in chromosome size and, occasionally, loss of hybridising signal. In all but 2 cases (clones 17 and 19, probed with *VSG221*) the chromosomes had diminished in size. *VSG121* hybridises to 2 chromosomes of approximately 2.1 and 2.3 Mb in WT cells, and both appear to have become smaller in all of the *brca2-/- clones* (as much as 100 kb in some cases), with clones 17 and 19 also displaying an extra hybridising fragment at approximately 800 kb. The same chromosomes also appear to have become smaller in 2-3 of the *BRCA2+/-* clones, although to a lesser extent. The blot probed with *VSG221,* displays a severe size change for 4 of the *brca2-/-* clones (clones 10-12 and 18), with a reduction of as much as 500 kb. The *GPI* probed blot continued to confirm these findings that GCRs arose from a reduction in chromosome size, with the smaller homologue of chromosome 1 reducing in size in all *brca2-/-* clones and 3-4 of the *BRCA2+/-* clones. Clone 9 displayed the greatest reduction in size, with an approximate 400 kb reduction, whilst clones 17 and 19, again appeared unusual by displaying an extra hybridising fragment at approximately 1.2 Mb.

The chromosome rearrangements visible when the *brca2-/-* clones were probed for *VSG221* could be due to the loss of DNA sequence associated with a VSG switching event, rather than general GCRs (in particular, clones 10-12, where the chromosomes had become ~ 500 kb smaller, and clones 17 and 19, where no obvious hybridisation was seen). In order to determine what the cause of the rearrangements was, all the clones were analysed for which VSG they were expressing. Whole cell extracts were prepared, and electrophoresed on a SDS-PAGE gel and probed for VSG221, which resides at the active VSG expression site. The resulting western blots of this analysis are displayed in figure 4.32 and indicate that all of the clones were still expressing VSG221. This confirms that none of the clones represent cells within the population that had inactivated VSG221 expression. This therefore means that the size alterations observed in figure 4.31 must be due to GCRs that are separate from VSG switching events that target the active ES.

The mini and intermediate sized chromosomes of *T. brucei* are not resolved well in the PFG separation, and it is not therefore possible to see if these also undergo GCRs.

Chromosomes of the same clones were therefore separated on the CHEF-DR III system (Bio-Rad), using distinct conditions (a 1 % agarose gel, in 0.0445 M Tris borate, 0.5 mM EDTA (0.5 x TBE) at 197.2 V, 20 seconds pulse time, for 16 hours at 14 °C) that allow separation of smaller DNA molecules. In this case the PFGs were only visualised by staining with ethidium bromide, as displayed in figure 4.33. Remarkably, no obvious chromosome rearrangements were observed in the intermediate or mini-chromosomes for any of the clones, even those for the *brca2-/-* cells.
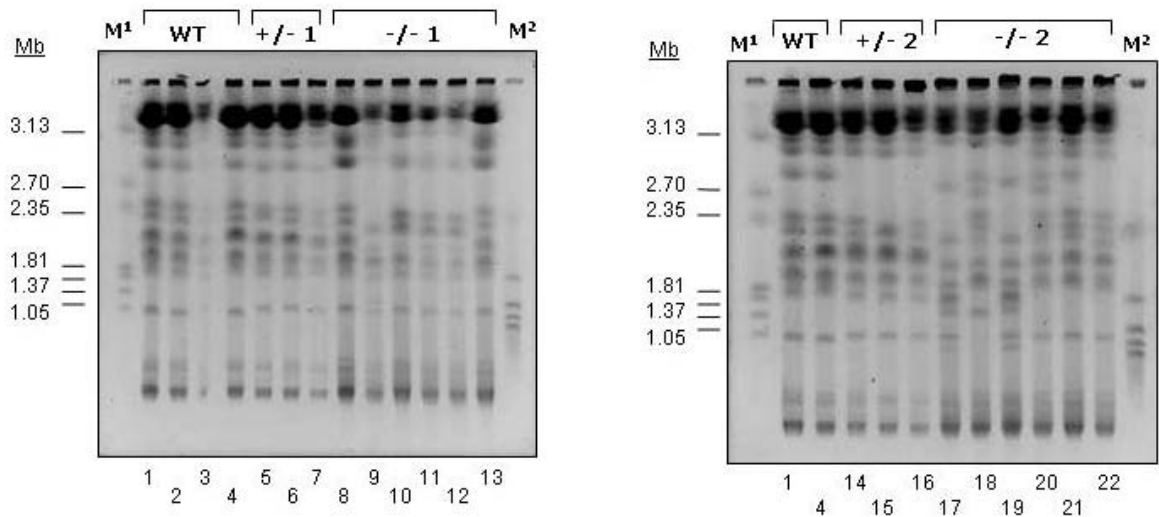
Figure 4.30 – Gross chromosomal rearrangements in *BRCA2* mutants. Ethidium bromide stained PFG electrophoresis separation of intact genomic DNA from wild type (WT), *BRCA2* heterozygous (+/-) and *brca2* homozygous (-/-) clones. Clones 1-4: wild type; clones 5-7 *BRCA2+/-1*; clones 14-16, *BRCA2+/-2*; clones 8-13, *brca2-/-1*; clones 17-22 *brca2-/-2*. Lanes containing marker DNA molecules (M) are shown; M[1] – *H. wingei*, M[2] – *S. cerevisiae* (Bio-Rad). The clones are the same clones used in figure 4.28.
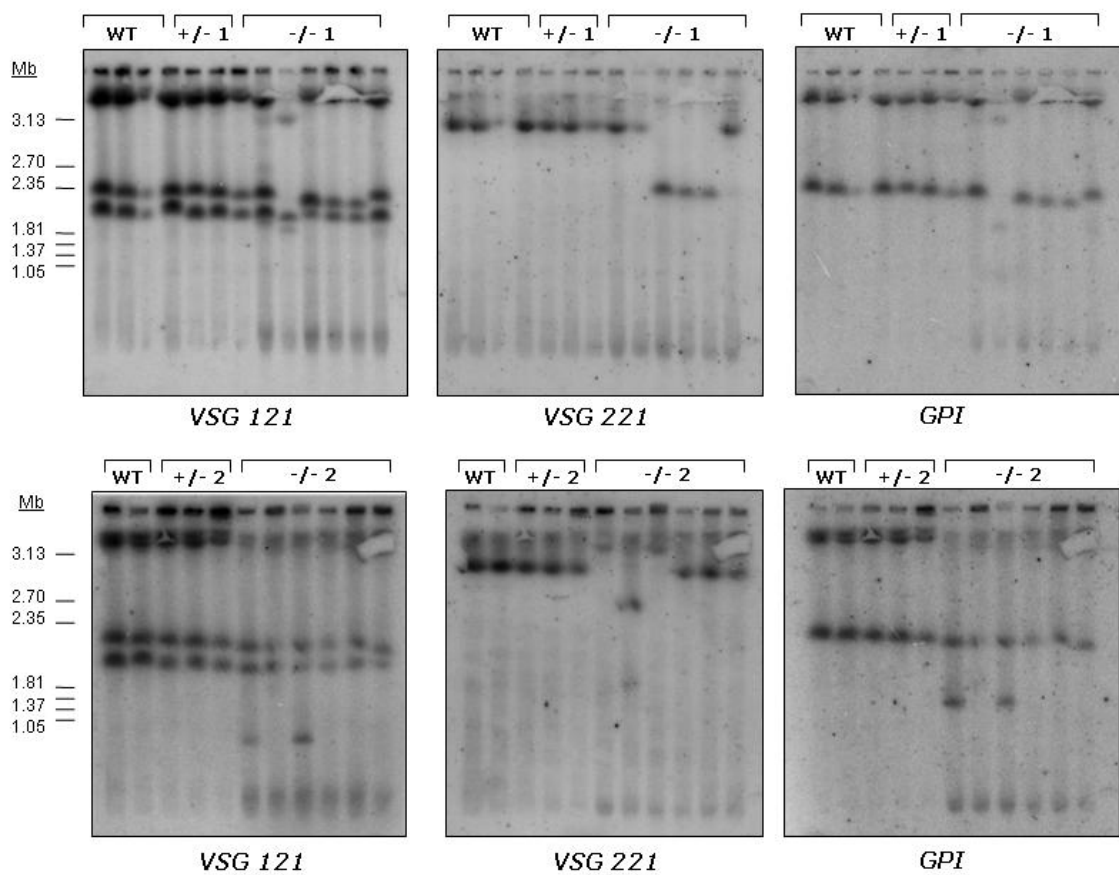
Figure 4.31 – Gross chromosomal rearrangements in *BRCA2* mutants. Southern blots of the PFG electrophoresis separation of intact genomic DNA from wild type (WT), *BRCA2* heterozygous (+/-) and *brca2* homozygous (-/-) clones shown in figure 4.29. The Southern blots have been sequentially probed with *VSG121, VSG221* and *GPI* (Glucose-6-isomerase).
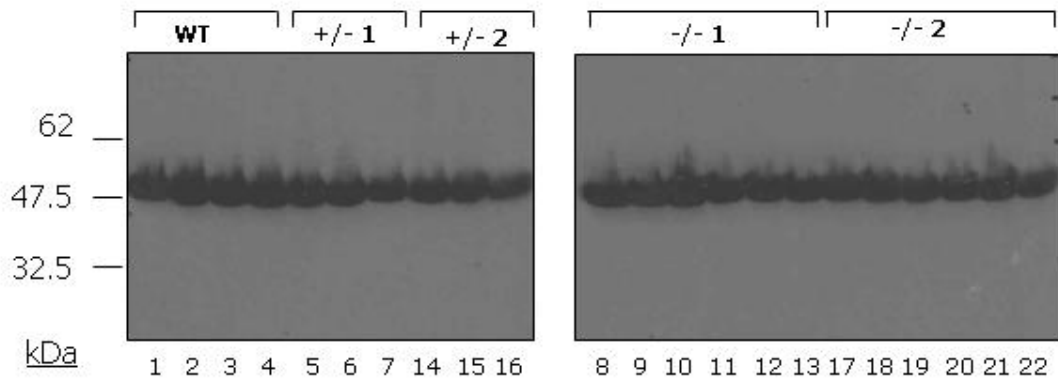
Figure 4.32 – Western blots of VSG221 expression in clonal cell lines. The western blots display total protein extracts from wild type (WT), *BRCA2* heterozygous (+/-) and *brca2* homozygous (-/-) clones probed with anti-VSG221 antiserum. VSG221 is visible at 47kDa, and the lanes represent the clones derived following extensive *in vitro* passaging, as depicted in figure 4.30. Clones 1-4: wild type; clones 5-7 *BRCA2+/-1*; clones 14-16, *BRCA2+/-2*; clones 8-13, *brca2-/-1*; clones 17-22 *brca2-/-2*.
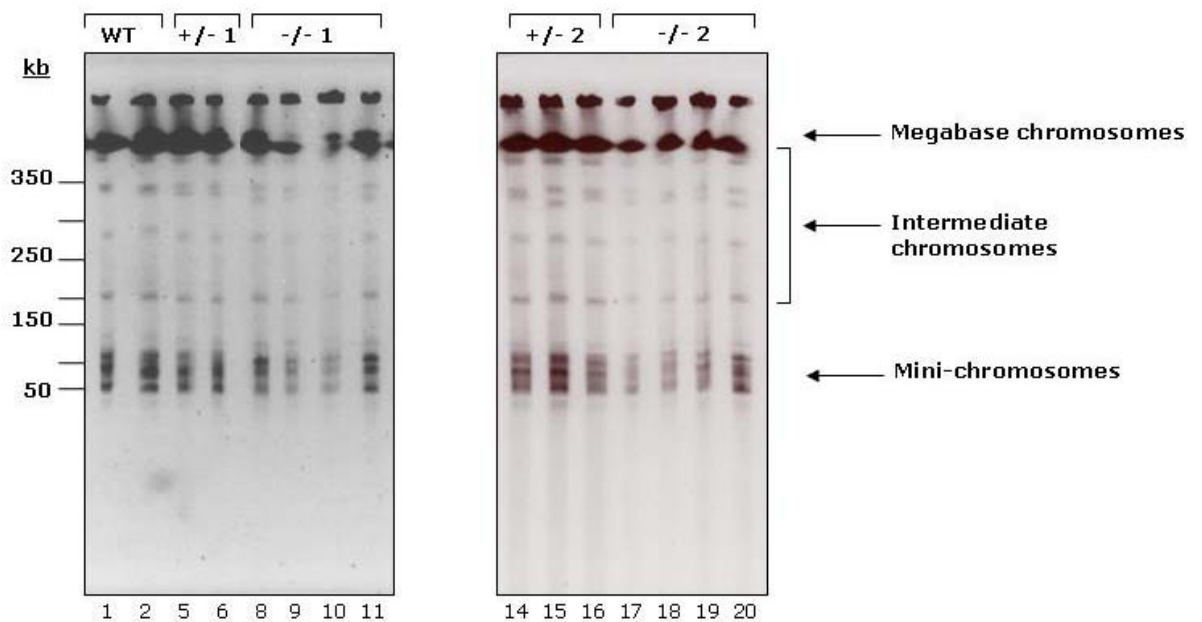
Figure 4.33 – PFGs showing the intermediate and mini-chromosomes. Ethidium bromide stained PFGs show the intermediate chromosomes and mini-chromosomes for wild type (WT), *BRCA2* heterozygous (+/-) and *brca2* homozygous (-/-) clones. Clones 1-4: wild type; clones 5-7 *BRCA2+/-1*; clones 14-16, *BRCA2+/-2*; clones 8-13, *brca2-/-1*; clones 17-22 *brca2-/-2*. The clones are the same clones used in figure 4.30.

4.4 Generation of re-expresser and over-expresser cell lines

In order to determine whether or not the phenotypic defects that had been observed in the *brca2-/-* mutants were due to the loss of *BRCA2,* and not the result of a secondary mutation, it was necessary to generate re-expression cell lines. This was performed in the second of the two independent *brca2-/-* homozygous mutants, (*brca2-/-2*), in both the Lister 427 and 3174.2 cell lines. Re-expression constructs were generated by PCR-amplifying the ORF of *BRCA2* from Lister 427 genomic DNA (primers *Tb BRCA2 for2* (which included an HA tag) and *Tb BRCA2 rev2*) using Herculase DNA polymerase (Stratagene). This PCR product was restriction digested with *Nru*I and ligated into the plasmids pRM481 and pRM482 (R.McCulloch, gift), which had been *Eco*RV-digested and CIP treated. This resulted in the generation of the constructs pRM481::*BRCA2* and pRM482::*BRCA2* (see figure 4.33), which contained the antibiotic resistance cassettes for phleomycin and G418, respectively. These constructs allowed *BRCA2* to be inserted into the tubulin array, where it was transcribed from the endogenous transcription. Splicing and polyadenylation was provided by 5' actin and 3' tubulin intergenic sequences respectively, as opposed to the natural processing of the gene, meaning that the level of mRNA, or its stability, could have been altered relative to endogenous *BRCA2*. Nevertheless, re-expression by this strategy was successful for RAD51-3 and RAD51-5 (Proudfoot and McCulloch, 2005).

The constructs were excised from the plasmid backbone by restriction digestion with *Xho*I and *Xba*I, before phenol: chloroform extraction and ethanol precipitation. Approximately 5 μg of digested DNA was introduced into *brca2-/-* cells. pRM481::*BRCA2* was transformed into 3174.2 *brca2-/-* cells and pRM482::*BRCA2* was transformed into Lister 427 *brca2-/-* cells. Antibiotic resistant transformants were selected by plating out 4×10^7 cells from each transformation at 1.25 μg.ml^{-1} phleomycin or 2.5 μg.ml^{-1} G418, respectively, over 48 wells with 1.5 mls per well. A number of antibiotic resistant transformants were recovered for each construct, and the re-introduction of *BRCA2* in the clones was confirmed by PCR-amplification of the entire ORF using Taq DNA polymerase and the primers '*Tb BRCA2 for2*' and '*Tb BRCA2 rev*' (data not shown) and Southern analysis (section 4.4.1). One transformant was chosen for each cell line and named 427 *BRCA2-/-/+* and 3174.2 *BRCA2-/-/+*.
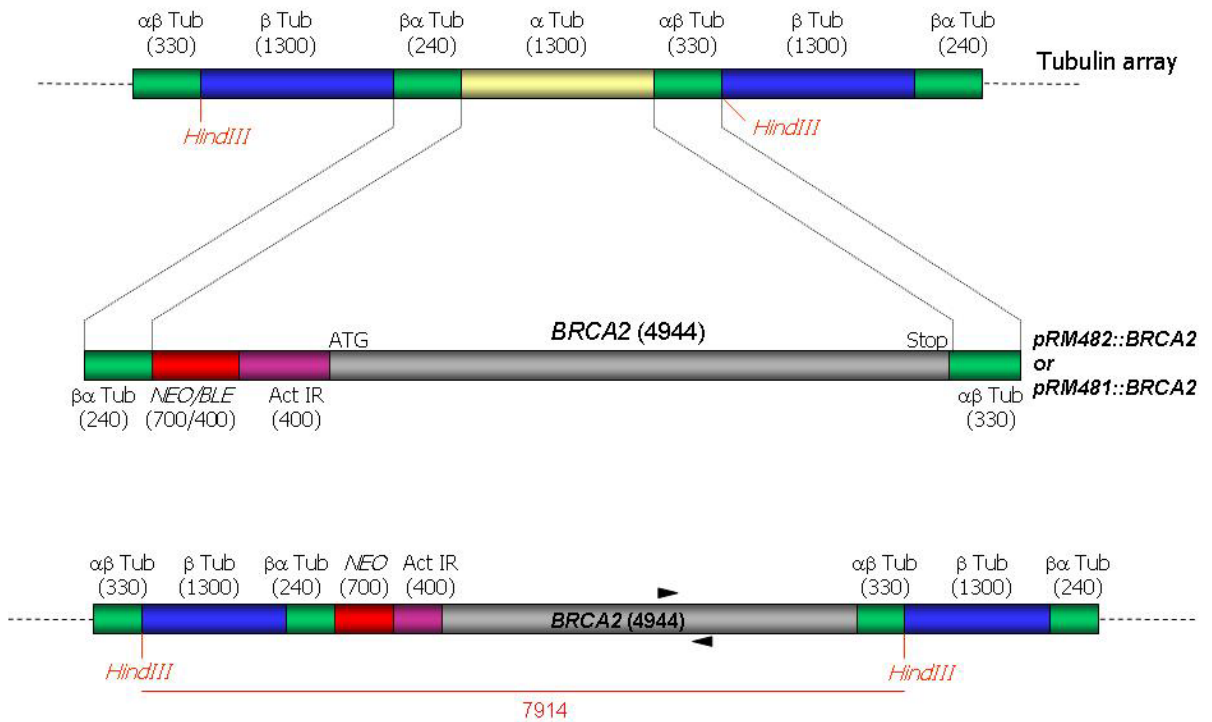
Figure 4.34 – pRM481::*BRCA2* and pRM482::*BRCA2* constructs generated to re-express BRCA2. BRCA2 was cloned into an *EcoRV* site between the actin intergenic (Act IR) and βα tubulin (βαTUB) intergenic sequences of the plasmids pRM481 and pRM482, which contain the antibiotic resistance cassettes for phleomycin (BLE) and G418 (NEO) respectively. The construct is flanked with tubulin intergenic regions (IR; αβ Tub and βα Tub), which allow homologous integration into the tubulin array, replacing an α tubulin ORF. The size of IR and ORFs are indicated (in bp).

The re-expression construct pRM482::*BRCA2* was also transformed into the wild type Lister 427 cell line. The purpose of this transformation was to generate a cell line which contained an extra copy of *BRCA2*, as a means of potentially over-expressing the protein, revealing information about transcription levels from the tubulin array as opposed to the endogenous locus of *BRCA2*. The construct was prepared as described above, before approximately 5 μg of digested DNA was transformed into wild type Lister 427 cells. Antibiotic resistant transformants were selected for by plating out 2×10^7 cells from the transformation at 2.5 μg.ml^{-1} G418, over 24 wells with 1.5 mls per well. A number of antibiotic resistant transformants were recovered and the introduction of an extra *BRCA2* in the clones was confirmed by Southern analysis (section 4.4.1). One transformant was chosen and named OE *BRCA2*.

4.4.1 Confirmation of BRCA2 re-expression by Southern analysis

In order to confirm that pRM481::*BRCA2* and pRM482::*BRCA2* had integrated into the tubulin array of the cell lines 3174.2 *BRCA2-/-/+*, 427 *BRCA2-/-/+* and OE *BRCA2* as expected, Southern analysis was carried out. Genomic DNA from wild type, *brca2-/-*, re-expression and over-expression cell lines were all subjected to Southern analysis, in order to allow direct comparison. Genomic DNA was digested with *Hin*dIII before being run out

on a 0.8 % agarose gel and transferred to a nylon membrane by Southern blotting. The blots were probed with a 378 bp region of the *BRCA2* ORF and the results are shown in figures 4.35 and 4.36. Two allelic variants for wild type cell lines, and the absence of *BRCA2* in the homozygous mutants, was seen as before (section 3.8.2). Integration of *BRCA2* in *brca2-/-* mutants in both Lister 427 and 3174.2 cell lines was confirmed by a hybridising fragment of approximately 7.5 kb, the size expected, when taking into account the number of BRC repeats (section 3.8.1). The blots were subsequently stripped and re-probed with a region of RNA polymerase I, as a positive loading control for DNA in the *brca2-/-* sample.
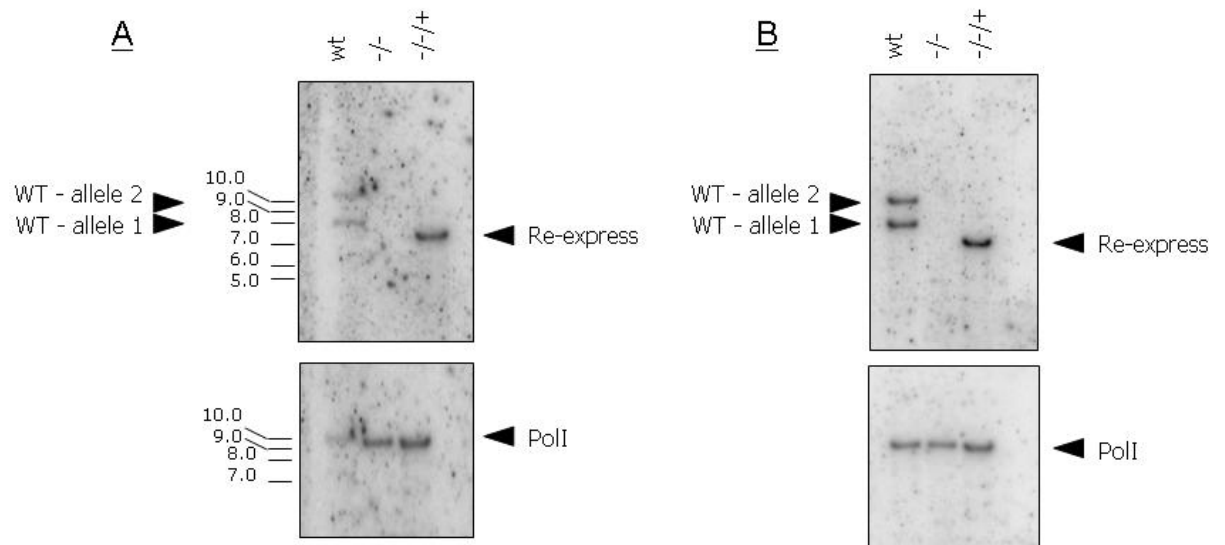
Figure 4.35 – Confirmation of the generation of re-expressers by Southern analysis.
Genomic DNA from (A) Lister 427 cell lines and (B) 3174.2 cell lines were digested with *Hin*dIII and 5μg run out on a 0.8 % agarose gel. The DNA was Southern blotted before being probed with a 378 bp region of the *BRCA2* ORF. WT refers to genomic DNA from untransformed cell lines. The homozygous mutants are indicated by -/- and re-expressers by -/-/+. Southern blots were stripped and re-probed with a bp fragment of the RNA polymerase I 452 bp (PolI) ORF.
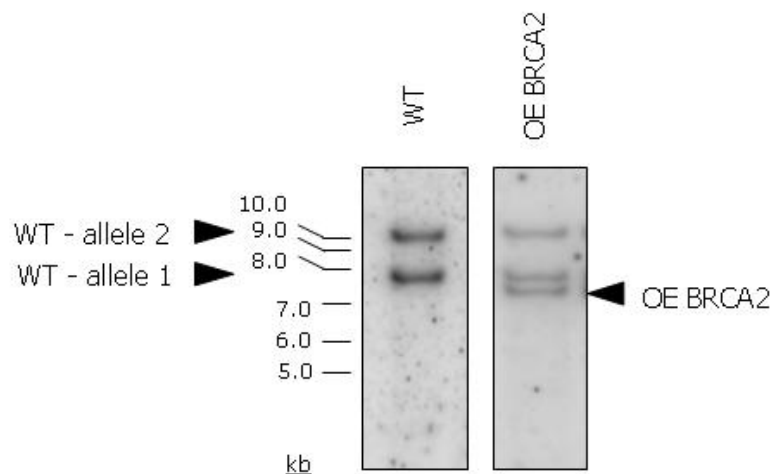
Figure 4.36 – Confirmation of the generation of a BRCA2 over-expresser by Southern analysis. Genomic DNA from Lister 427 cell lines were digested with *Hin*dIII and 5 μg run out on a 0.8 % agarose gel. The DNA was Southern blotted before being probed a 378 bp region of the *BRCA2* ORF. WT refers to genomic DNA from untransformed cell lines. Over-expressers are denoted by OE BRCA2.

The addition of a third copy of *BRCA2* in wild type Lister427 cells was revealed by the presence of a third BRCA2-hybridising band, corresponding to the re-expresser band in figure 4.35, in addition to the 2 endogenous BRCA2 alleles (figure 4.36).

4.4.2 Confirmation of BRCA2 re-expression by RT-PCR

To support the results of the Southern analyses, RT-PCR was carried out on the re-expresser cell lines (as described in section 4.2.4). A BRCA2-specific product, equivalent in size to that generated in wild type and *BRCA2+/-* cell lines, but absent in *brca2-/-* cells (figure 4.6), was seen (figure 4.37). This confirms that *BRCA2* mRNA is present in the re-expresser mutants for the cell lines 3174.2 *BRCA2-/-/+* and 427 *BRCA2-/-/+*. As this is non-quantitative PCR, it is not possible to determine if the amount of BRCA2 cDNA is equivalent in the -/-/+ cells to either the WT or *BRCA2+/-* cells. For the same reasons, the RT-PCR analysis was not performed on the *OE BRCA2* cell line.
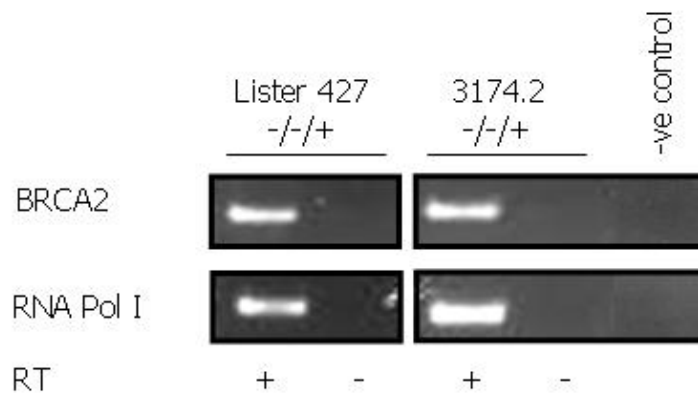
Figure 4.37 – Confirmation of the generation of re-expresser mutants by RT-PCR. RT-PCR was carried out on cDNA generated from total RNA from Lister 427 re-expressers (427 -/-/+) and 3174.2 re-expressers (3174.2 -/-/+). RNA polymerase I specific primers were used to control for the generation of intact cDNA. Primers specific for *BRCA2* were used to show the expression of that gene. The negative control contains no cDNA substrate. RT + denotes cDNA generated with reverse transcriptase, RT – denotes control reactions that were treated equivalently but no RT was added to the reactions.

It is worth noting that the re-expression constructs (pRM481::*BRCA2* and pRM482::*BRCA2*), both contained an N terminal HA tag. However, expression of the protein was undetectable for 3174.2 *BRCA2-/-/+*, 427 *BRCA2-/-/+* and *OEBRCA2* by western blot analysis with two anti-HA antibodies (Sigma and Roche). Expression of mRNA was shown to be present for 427 *BRCA2-/-/+* by northern blot (see section 5.2.3.3).

4.5 Phenotypic analysis of BRCA2 re-expresser and over-expresser cell lines

The re-expresser and over-expresser cell lines from Lister 427 (427 *BRCA2-/-/+* and OE *BRCA2*), were analysed for their *in vitro* population doubling times, cell cycle progression, DNA damage sensitivity, recombination efficiency and the ability to form RAD51 foci. The 3174.2 re-expresser cell line (3174.2 *brca2-/-/+*) was used only to analyse VSG switching rates.

4.5.1 Analysis of in vitro growth

Analysis of *in vitro* growth of the 427 *brca2-/-/+* and OE *BRCA2* cell lines was carried out to determine if transcription of the gene from the tubulin array alleviated the growth defect that was observed in the homozygous mutants, and if an extra copy of *BRCA2* in wild type cells altered the growth rate. The assay was carried out following the same protocol as described in section 4.3.1. Three repetitions of the growth assay were carried out for each cell line and are displayed in figure 4.38, compared with the values determined previously. Both the BRCA2 re-expresser and over-expresser cell lines were found to have growth rates that were essentially equivalent to that of wild type cells, with doubling times of 8.26 and 7.58 hours respectively, compared to 8.19 of wild type cells. This confirms that it is the absence of BRCA2 which causes *T. brucei brca2-/-* mutants to have an increased population doubling time. Providing an extra copy of *BRCA2* in wild type cells, appears to provide the cell with no growth advantage or impediment.
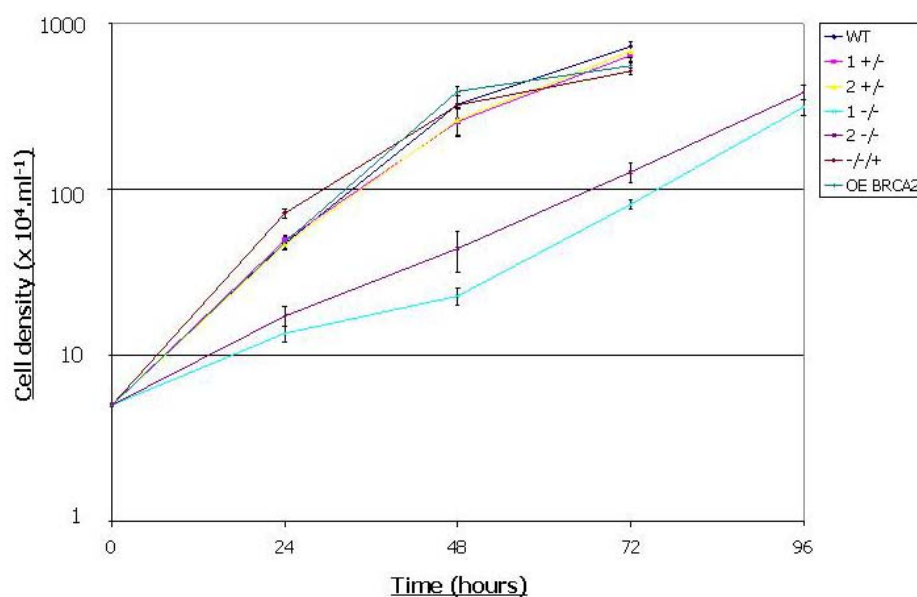
Figure 4.38 – Analysis of *in vitro* growth of *BRCA2* re-expressers and over-expressers. 5 ml cultures were set up at 5 x 10^4 cells.ml^{-1} and cell densities counted 24, 48, 72 and 96 hours subsequently. Standard errors are indicated for the counts using data from three repetitions. WT: wild type; +/-: heterozygote; -/-: homozygote; -/-/+: re-expresser; OE: over-expresser.

4.5.2 Analysis of the cell cycle

Analysis of cell cycle progression in the 427*BRCA2-/-/+* and OE *BRCA2* cell lines was next examined. The results of this are displayed in figure 4.39. When *BRCA2* was re-expressed, cell cycle progression occurred relatively normally, with the number of aberrant cells reducing from 11.8 % in the homozygous mutant in which the gene was integrated, to 4.8 % in the re-expresser cell line. This number of aberrant cells is relatively comparable to the 2 % found in wild type cells and 1-2 % found in the *BRCA2+/-* cells. The slight increase in these cell types is not a significant result (see section 5.3.2), but could result from GCRs that accumulated in the *brca2-/-* mutants.

The addition of an extra copy of *BRCA2* had no effect on cell cycle progression, with the distribution of cell cycle stages being indistinct from those in wild type cells and the number of aberrant cells being unaltered (2.8 %).

This result appears to confirm that it is the absence of BRCA2 which causes *T. brucei brca2-/-* mutants to accumulate aberrant cells.
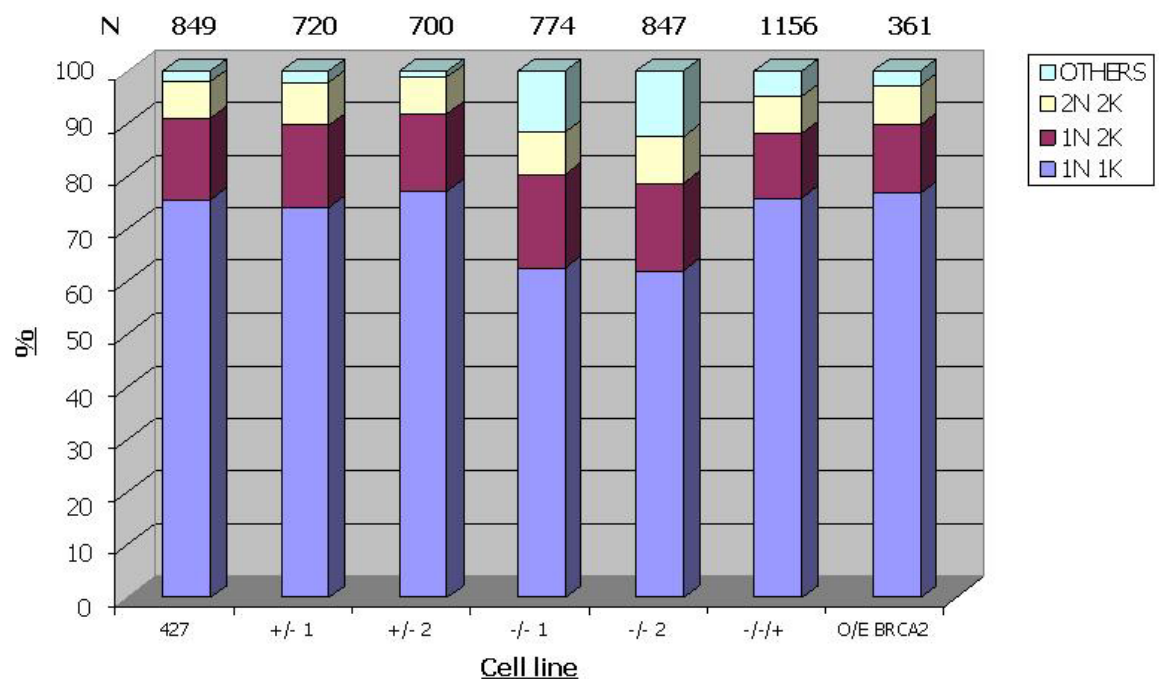
Figure 4.39 – DAPI analysis of the *BRCA2* mutants. The DNA content of wild type Lister 427, *BRCA2* heterozygous (+/-), homozygous (-/-), re-expresser (-/-/+) and over-expresser (OE) cell lines were visualised by DAPI. The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and unidentifiable cells (others) were counted and represented by their mean count as a percentage of the total cells counted. N = number of cells counted.

In order to examine if the abundance of cells in different cell cycle stages altered when BRCA2 was either re-expressed or over-expressed in the presence of damage, the DNA content of 427 *BRCA2-/-/+* and OE *BRCA2* cell lines was analysed after phleomycin

treatment, as described in section 4.3.3. The results are displayed in figure 4.40 and demonstrate that when *BRCA2* was re-expressed the cell cycle progression appears to occur normally, with the number of aberrant cells reducing from 19.4 % in *brca2-/-2* (at 1.0 µg.ml^{-1} of phleomycin) to 6.5 % in the re-expresser cell line, which is comparable to the 5.5 % found in wild type cells. Similarly, the number of 1N 2K cells decreased from 30.1 % in *brca2-/-2* to 15.1 % in the re-expresser cell line, indicating progression through G2 phase. Finally, the number of 1N 1K cells was seen to increase, from 45 % in *brca2-/-2* to 72.3 % in the re-expresser cell line, which is again comparable to the 76.3 % found in wild type cells.

The addition of an extra copy of *BRCA2* had no effect on abundance of cells in different cell cycle stages in the presence of damage, with the phenotype observed being indistinct from that of wild type or *BRCA2+/-* cells.

This result appears to confirm that it is the absence of BRCA2 which causes *T. brucei brca2-/-* mutants to accumulate 1N 2K and aberrant cells in the presence of DNA damage. Re-expression of BRCA2 therefore appears to alleviate the *brca2-/-* mutants' impairment in progression through nuclear DNA replication, most likely by allowing DNA damage to be repaired.
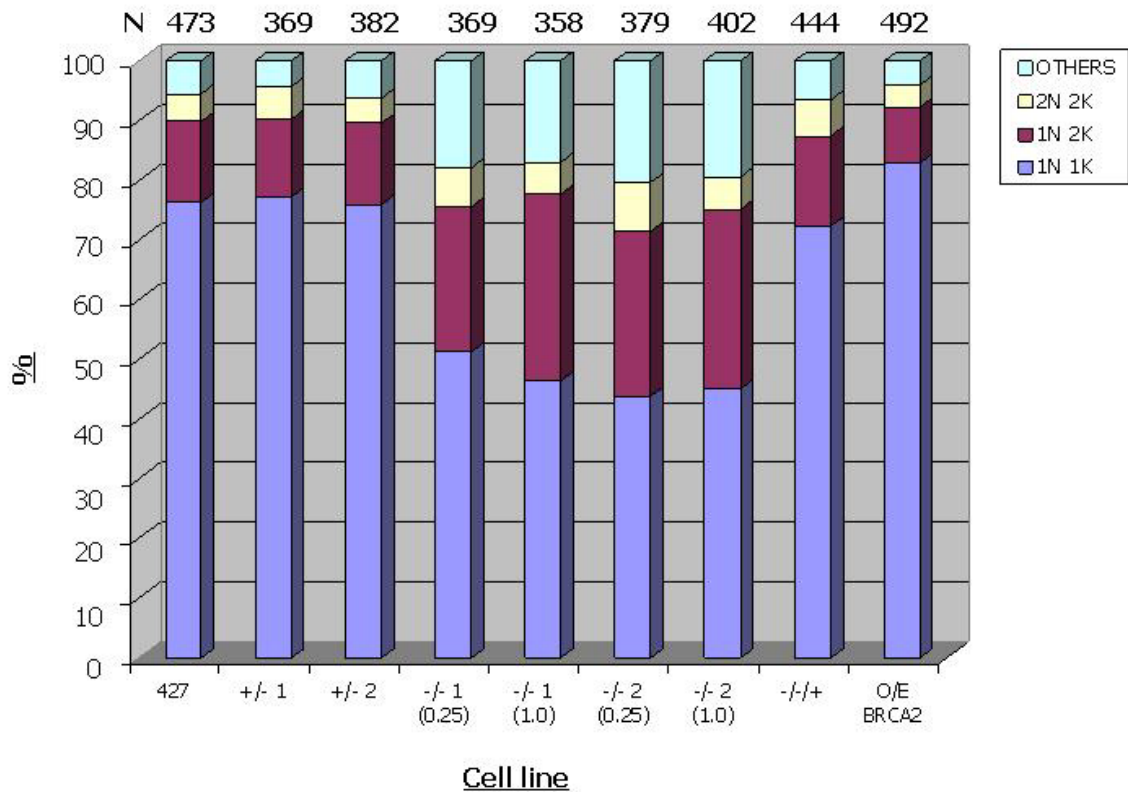
Figure 4.40 – DAPI analysis of *BRCA2* mutants after DNA damage. The DNA content of *BRCA2* heterozygous (+/-), homozygous (-/-), re-expresser (-/-/+) and over-expresser (OE) mutant cell lines were visualised by DAPI and compared with wild type Lister 427 cells, after cells had been damaged by phleomycin. The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications cells (others) were counted and represented by their mean count as a percentage of the total cells counted. Wild type, heterozygous cells, re-expresser and over-expresser cell lines were grown in media with 1.0 μg.ml^{-1} of phleomycin, whilst homozygous cells were grown in media with 0.25 μg.ml^{-1} and 1.0 μg.ml^{-1} of phleomycin. N = number of cells counted.

In order to examine if re-expression of *BRCA2* alleviated the *brca2-/-* mutants' difficulty in segregating the nuclei during mitosis, the number of 2N 2K cells with incompletely or completely separated nuclei in 427 *BRCA2-/-/+* mutants was analysed as described in section 4.3.3 (*OEBRCA2* was not assayed). The results are displayed in figure 4.41, relative to those generated previously, and demonstrate that when *BRCA2* is re-expressed, mitosis is able to progress at a similar rate to that of wild type cells, as the number of incompletely segregated nuclei reduced from the 37.6 % seen in the *brca2-/-2* mutant to 16.3 % in the re-expresser cell line, which is relatively comparable to the 12.1 %, 9.6 % and 11.5 % seen in wild type and *BRCA2+/-* cells. This phenotype is not entirely reverted to that of WT cells, but again, this could be due to GCRs that accumulated in the *brca2-/-* mutants.
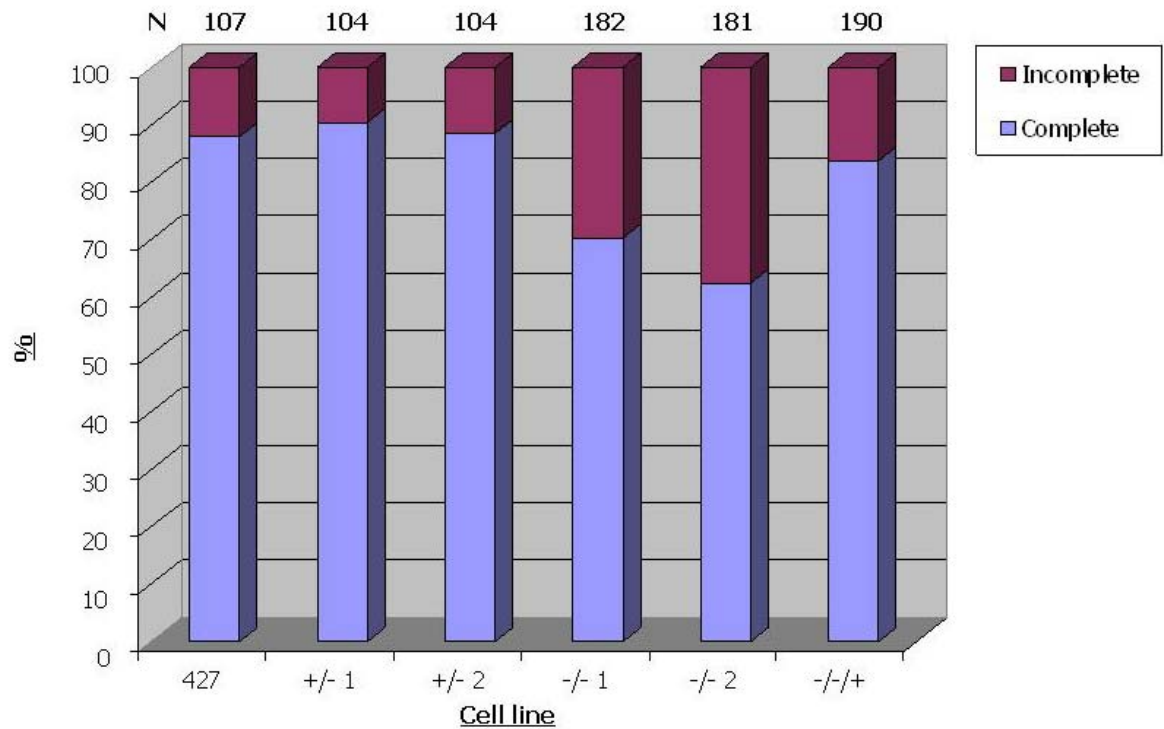
Figure 4.41 – Analysis of the number of 2N 2K cells that have completed nuclear division.
2N 2K cells in *BRCA2* heterozygous (+/-), homozygous (-/-) and re-expresser (-/-/+) mutant cell
lines were visualised by DAPI and compared with wild type Lister 427 cells. The cells were
analysed for the number that had completed nuclear division (complete), and those that were still
dividing the nucleus (incomplete). N = number of cells counted.

4.5.3 Analysis of DNA damage sensitivity

Analysis of DNA damage sensitivity in the 427 *BRCA2-/-/+* and OE *BRCA2* cell lines was
the next phenotype examined. To do this, the clonal survival and Alamar Blue assays were
carried out following the same protocols as described in section 4.3.4. The clonal survival
assay, using MMS as a DNA damaging agent, was performed only for 427 *BRCA2-/-/+*
cells, and the results are displayed in figure 4.42 alongside the wild type, *BRCA2*
heterozygous and *brca2* homozygous cell lines for comparison. The results demonstrate
that re-expression of *BRCA2* in the *brca2-/-2* mutant resulted in an elevated resistance to
MMS relative to WT or *BRCA2+/-* cells. There was little significant impairment in growth
at 0.0003 % MMS, where WT and *BRCA2+/-* cells had 39 – 54 % survival. More
tellingly, at 0.0004 % MMS, where the WT and *BRCA2+/-* cells showed only 12 – 24 %
survival, more than 60 % of *BRCA2-/-/+* cells survived MMS treatment. This was initially
taken as evidence for higher levels of *BRCA2* expression from the tubulin array compared
with the endogenous locus, and perhaps that elevated levels of BRCA2 would enable DNA
damage to be repaired faster than in wild type cells, through enhanced ability to transport
RAD51 to the sites of DNA damage, or enhancement of RAD51 strand exchange.
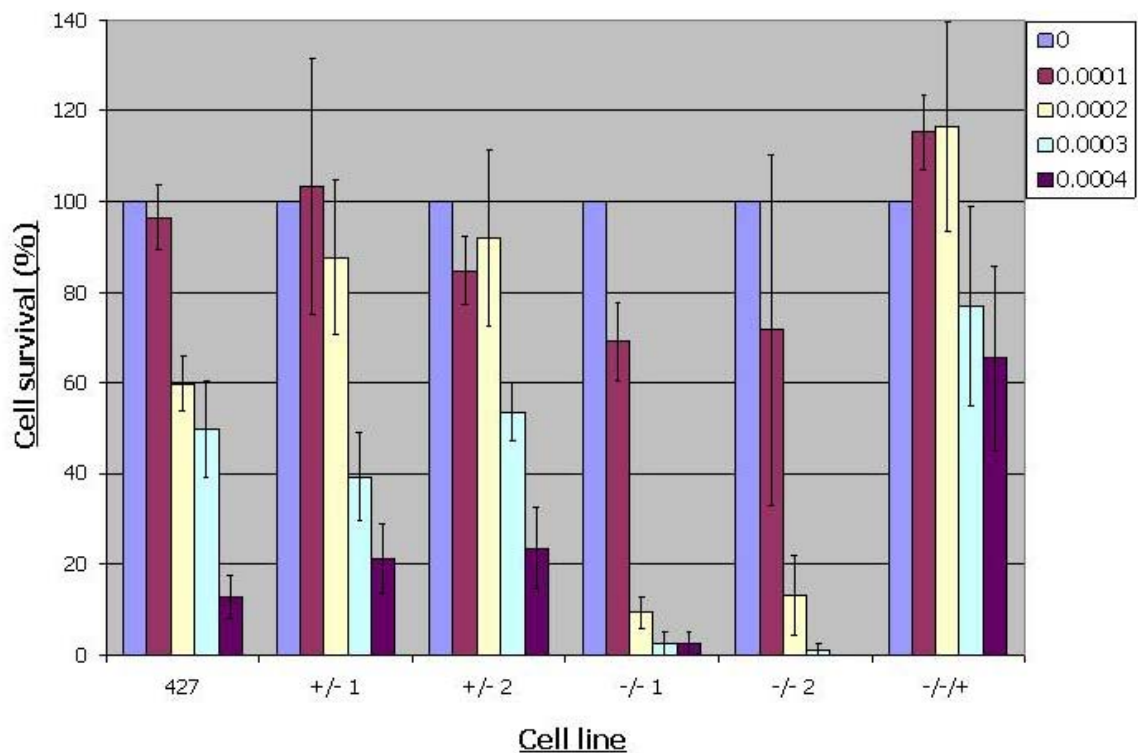
Figure 4.42 – Analysis of DNA damage sensitivity in the *BRCA2* mutants. Each strain was plated at one cell per well in five 96 well plates, each containing a different concentration of MMS: 0, 0.0001, 0.0002, 0.0003 and 0.0004 %. Four repetitions were carried out for each cell line. The mean number of cells to grow through for each cell line at each concentration was calculated and represented as a percentage of the number that had grown through on the 0 % plate for that cell line. Standard errors are indicated and the data is presented for the two independent heterozygous mutants (+/- 1, 2), the two independent homozygous mutants (-/- 1, 2), the re-expressers (-/-/+) and the wild type Lister 427 cell line.

In order to investigate this phenotype further, both the 427 *BRCA2-/-/+* and *OEBRCA2* cells were examined using the Alamar Blue assay (see section 2.1.6.2) using both MMS and phleomycin as DNA damaging agents. These results are displayed in figures 4.43 and 4.44, and confirm that the *BRCA2-/-/+* cell line is indeed more resistant to MMS than wild type cells, with a mean IC50 of 0.0031 % MMS compared to 0.0015 % in WT cells. Not only this, but the re-expresser cell line is also more resistant to phleomycin than wild type cells, with a mean IC50 of 0.144 μM compared to 0.095 μM in WT cells. Interestingly, when an extra copy of *BRCA2* is expressed from the tubulin array locus, the cells do not display an advantage over wild type cells in terms of DNA damage sensitivity: the IC50s determined for both MMS and phleomycin in the *OEBRCA2* cells are indistinguishable from WT or *BRCA2+/-* cells. This therefore leads us to reject the initial hypothesis that *BRCA2* is transcribed at a higher level in the tubulin array than in the endogenous locus, and the reason for the 427 *BRCA2-/-/+* mutant's advantage over wild type cells in terms of DNA damage sensitivity must therefore be explained by another phenomenon.
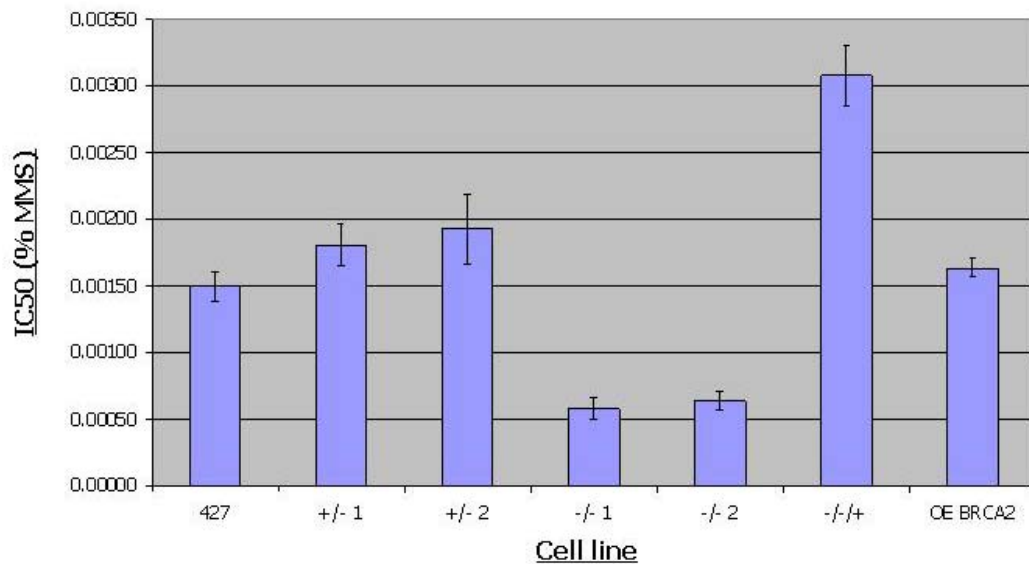
Figure 4.43 – IC50 of *T. brucei BRCA2* mutants exposed to MMS. Wild type, *BRCA2+/-*, *brca2-/-*, *BRCA2-/-/+* and OE *BRCA2* cell lines were placed in serially decreasing amounts of MMS and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments; bars indicate standard error.
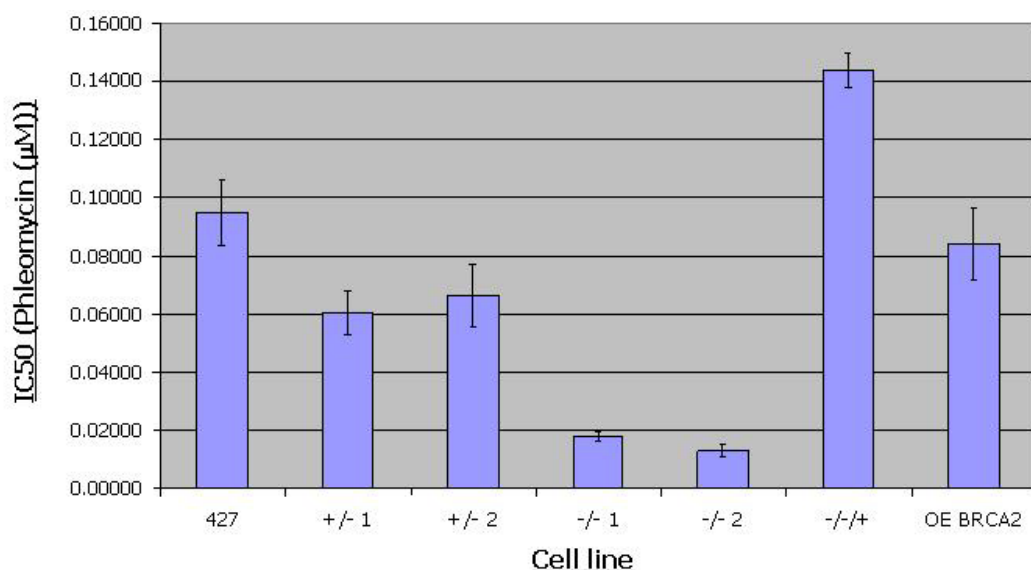
Figure 4.44 – IC50 of *T. brucei BRCA2* mutants exposed to phleomycin. Wild type, *BRCA2+/-*, *brca2-/-*, *BRCA2-/-/+* and *OE BRCA2* cell lines were placed in serially decreasing amounts of phleomycin and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments; bars indicate standard error.

One explanation could be that a mutation could have arisen during the passaging of the *brca2-/-* cells that resulted in increased resistance to MMS and phelomycin. Indeed, mutants with increased resistance to MMS have previously been reported in mismatch repair (Glaab *et al.*, 1998), amino acid biosynthesis (Kafer, 1987) and p53 function (Kuo *et al.*, 1997). Whether the same mutations lead to phleomycin resistance has not been

documented. Another possibility is that the uptake systems for the drugs could have been mutated, resulting in a decreased transfer across the parasites membrane. However, given that MMS and phleomycin damage DNA by distinct modes of action, it seems highly unlikely that the cells could have accumulated mutations which resulted in resistance to both of these agents. Similarly, the structure of the drugs is highly distinct, suggesting that uptake is likely to be *via* different pathways.

Given the above, another possibility is that during the prolonged growth of the *brca2-/-* mutants, another DNA repair pathway has become more active than usual. Following the re-integration of BRCA2 and re-instalment of active homologous recombination, cells could display an increased resistance to DNA damage if both of these pathways were active at the same time. We already know that in the absence of BRCA2, cells are still able to perform homologous recombination at a low level, presumably through another possibly unknown pathway. It could therefore be this pathway which allows this increased resistance to DNA damaging agents to exist. Having said this, this explanation is also not truly satisfactory, since it would be expected that the *brca2-/-* cells would themselves display enhanced tolerance to DNA damage, and this was not seen.

Since the re-expresser cell line was generated in only one *brca2-/-* mutant cell line, it would be interesting to generate a re-expresser cell line in the other independent *brca2-/-* mutant cell line in order to see if this phenomenon was again produced. Nevertheless, despite the complexities of these results, we can still be confident that it is the absence of BRCA2 which causes *T. brucei brca2-/-* mutants to be sensitive to DNA damaging agents, since re-expression reverts the phenotype.

4.5.4 Analysis of recombination efficiency

Analysis of recombination efficiency was next examined in the 427 *BRCA2-/-/+* and OE *BRCA2* cell following the same protocol described in section 4.3.5. Three repetitions of the transformation efficiency assay were carried out for each cell line and are shown in figure 4.45, alongside the transformation efficiency rates for wild type, *BRCA2* heterozygous and *brca2* homozygous cell lines for comparison. Both the re-expresser and over-expresser cell lines were found to generate transformants at essentially equivalent frequencies to that of wild type and *BRCA2+/-* cells, with mean transformation efficiency rates of 4.27 and 3.80 generated, respectively, compared to 4.53 in wild type cells. These results confirm that it is the absence of BRCA2 which causes *T. brucei brca2-/-* mutants to be impaired in their ability to transform constructs into their genome. Providing an extra

copy of *BRCA2* in wild type cells provided the cells with no detectable advantage over wild type cells.
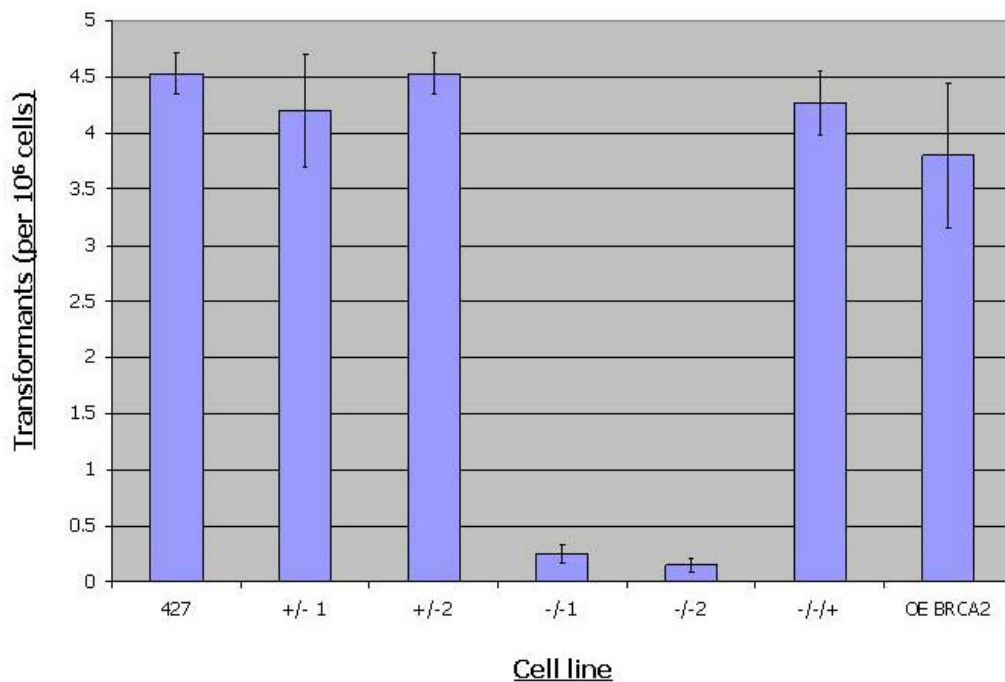
Figure 4.45 – Recombination efficiency in *BRCA2* mutants. Values are mean numbers of transformants obtained per 10^6 cells transformed; error bars are shown from 3 repetitions. The data are presented for two independent heterozygous mutants (+/-), two independent homozygous mutants (-/-), re-expressers (-/-/+), over-expressers (OE) and wild type Lister 427 cells (427).

4.5.5 Analysis of the ability to form RAD51 foci

To confirm that BRCA2 is central to the ability of *T. brucei* to form RAD51 sub-nuclear foci following DNA damage, the 427 *BRCA2-/-/+* cell line was treated with phleomycin and RAD51 localisation examined by indirect immunofluorescence, as described in section 4.3.6. 256 cells were counted and scored for the number of foci they contained. These results are displayed in table 4.15 and examples of these foci are shown in figure 4.46. The *BRCA2* re-expresser cell line was found to form RAD51 foci at least as efficiently as wild type cells. When the cells were treated with phleomycin to introduce DNA double strand breaks, the majority of cells (85 %) were found to contain RAD51 foci. This is slightly higher than was found for WT and *BRCA2+/-* cells (75-80 %). This finding could simply be due to counting differences because these experiments were performed on separate occasions. Another hypothesis could be that this effect is due to a hyper response to DNA induced damage, which could fit with the increased level of sensitivity (see section 4.3.4).

		Number of foci (%)						
	BLE	**0**	**1**	**2**	**3**	**4**	**5**	**6 or more**
-/-/+	**0.0**	87.5	3.5	2.3	2.3	1.9	1.6	0.8
	1.0	14.8	10.5	15.6	14.8	13.7	13.3	17.2

Table 4.15 – RAD51 foci formation in *BRCA2-/-/+* mutants. The percentages of cells showing foci at given concentrations of phleomycin (BLE) are shown. Phleomycin concentrations are shown in μg.ml^{-1}. Boxes shaded in light yellow contain foci, whilst boxes shaded in bright yellow contain the highest percentage of foci.
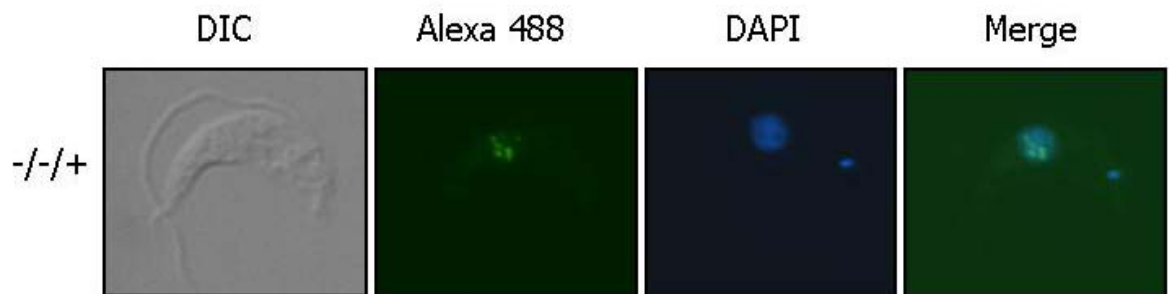
Figure 4.46 – RAD51 immunolocalisation in *BRCA2-/-/+* mutants. Representative images of T. brucei cells following growth in 1.0μg.ml^{-1} phleomycin for 18 hours are shown. Each cell is shown in differential interface contrast (DIC), after staining with DAPI and after hybridisation with anti-RAD51 antiserum and secondary hybridisation with Alexa Fluor 488 conjugate (Alexa 488). Merged images of DAPI and Alexa 488 cells are also shown. BRCA2 re-expresser cells (-/-/+) are shown.

4.5.6 Analysis of VSG switching

Finally, analysis of VSG switching was performed in the 3174.2 *BRCA2-/-/+* cell line. The assay was carried out following the same protocol as described in section 4.3.8, with three repetitions of the assay. The results of this are shown in figure 4.47, alongside the VSG switching frequencies for wild type, *BRCA2* heterozygous and *brca2* homozygous cell lines, to allow for comparison. The re-expresser cell line displayed a VSG switching frequency that was essentially equivalent to that of *BRCA2* heterozygous mutants, with a mean VSG switching frequency of 11.31×10^{-7}, compared to 11.37×10^{-7} and 11.67×10^{-7} in *BRCA2+/- 1* and *2*, respectively. This was a significant increase compared to 0.55×10^{-7} in the *brca2-/-2* from which the re-expresser was generated. The profile of the VSG switching mechanisms were not examined.
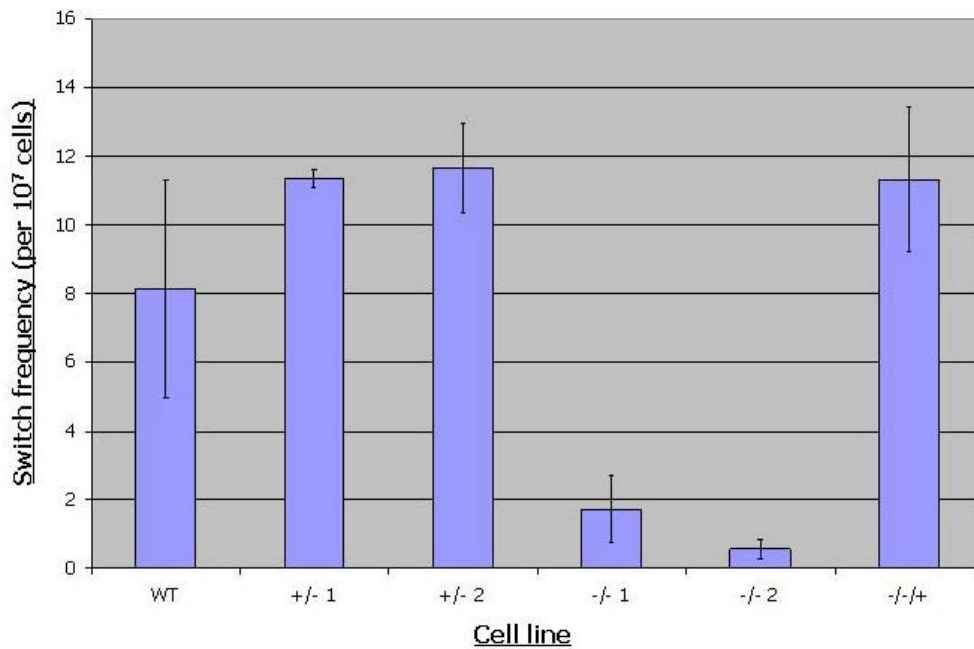
Figure 4.47 – VSG switching frequencies in *BRCA2* mutants. Values shown are the average switching frequencies for 3174.2 wild type cells (WT), the 2 independent (1 and 2) *BRCA2* heterozygous (+/-) and homozygous (-/-) mutants and the re-expresser (-/-/+) mutants. Data are from at least 2 experiments and standard error is indicated by bars.

4.6 Summary

The aim of this chapter was to determine if BRCA2 plays a role in DNA damage repair, recombination or VSG switching in *T. brucei*. To do this, a reverse genetics approach was taken and homozygous mutants generated. *brca2-/- T. brucei* bloodstream stage cells were shown to be viable, indicating that *BRCA2* is not essential in this life cycle stage, as has previously been found for *RAD51* and the *RAD51*-related genes *RAD51-3, RAD51-5* and *DMC1*.

Mutation of *BRCA2* caused a growth defect, observed by increased population doubling times *in vitro* and *in vivo*. DNA content analysis of *brca2-/-* mutants revealed an accumulation of cells with aberrant DNA content and an increase in the proportion of cells that were in the process of nuclear DNA segregation during mitosis. These phenotypes were not observed in other *T. brucei* recombination mutants, most notably *RAD51*, suggesting that BRCA2 has a role in the regulation of cell division. An interpretation of this is that BRCA2 mutation causes impaired replication or segregation of *T. brucei* nuclear DNA, but without a cell cycle stall, leading to aberrantly early cytokinesis and the accumulation of cells with incorrect DNA content.

brca2-/- mutants were also shown to have a number of DNA repair and recombination impairments. The mutants displayed genomic instability, detectable by gross chromosomal rearrangements in the megabase chromosomes, at least some of which arises due to deletions within the *VSG* arrays. Interestingly, the smaller chromosome classes of *T. brucei* appeared not to be susceptible to such instability. *brca2-/-* mutants were impaired in their ability to repair DNA damage, induced by both MMS and phleomycin, and in the ability to integrate transformed DNA constructs into their genome. Both phenotypes are consistent with a role for BRCA2 in *T. brucei* DNA repair and homologous recombination, due to an impairment in the interaction with RAD51, since re-localisation of this protein to sub-nuclear foci following phleomycin damage was compromised.

Finally, *brca2-/-* mutants displayed a reduced ability to switch their VSG coat. The extent of this phenotype was highly reminiscent of *rad51-/-* and *rad51-3-/-* mutants, both in the level of impairment and in the finding that all modes of switching appeared to be affected, with reduced levels of gene conversion and *in situ* switching mechanisms observed. This result highlights the importance of homologous recombination events in VSG switching. The re-expression of *BRCA2* confirmed the direct role of the protein in all the above results, whilst over-expression of *BRCA2* had no detectable effect in any of the phenotypes assayed.

CHAPTER 5

Complementation of *brca2-/-* mutants with variants of BRCA2

5.1 Introduction

In the previous chapter, *brca2-/-* mutants were generated in bloodstream stage *T. brucei* and shown to display impaired growth, both *in vitro* and *in vivo*; replication or cell cycle defects; genomic instability; sensitivity to MMS and phleomycin induced damage; an impaired ability to introduce constructs into their genome; and a reduction in VSG switching frequency. In previous research, *rad51-/-* and *rad51-3-/-* mutants were generated (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005) and displayed broadly comparable phenotypes, consistent with impaired DNA repair and recombination. The striking exception to this is the cell cycle defects that were observed in *brca2-/-* cells, as this phenotype has not previously been displayed in any of the recombination mutants generated in *T. brucei*, including *rad51-5-/-* and *mre11-/-* (Proudfoot and McCulloch, 2005;Robinson *et al.*, 2002).

The structure of *T. brucei* BRCA2 was examined (see Chapter 3) and shown to contain an unusually high number of BRC repeats. Furthermore, the DNA binding domain (DBD) in *T. brucei* BRCA2 was predicted to contain all 5 conserved motifs; an α helical domain, three OB domains and a tower domain. This is contrary to the prediction made by Lo *et al*. 2003, whereby the OB3 domain was thought to be absent, similar to BRCA2 homologues in other eukaryotes such as *U. maydis* Brh2 (Lo *et al.*, 2003;Kojic *et al.*, 2002). In addition, the DBD domain was also predicted to function in a similar manner to higher eukaryotes, due not only to conservation of the DBD motifs, but also to the presence of a DSS1 homologue. Indeed, the DBD domain in *U. maydis* Brh2 has been shown to function in a similar manner to that of BRCA2 in higher eukaryotes, where a functional homologue of DSS1 has also been shown to exist (Kojic *et al.*, 2003;Kojic and Holloman, 2004;Kojic *et al.*, 2005;Zhou *et al.*, 2007). In contrast to this broad conservation, little evidence was available for a C terminal RAD51 binding motif, apart from a putative CDK binding motive, leaving the question of whether *T. brucei* BRCA2 could bind RAD51 bimodally.

In this chapter, we wanted to investigate if the whole protein was needed for BRCA2 to function, or if certain motifs could function alone. In particular, we wished to localise the region of the protein which was responsible for the unusual phenotype of a cell cycle defect, and to ask if this was separable from the DNA repair function. To do this, attempts were made to generate mutants expressing various truncations of the protein.

Initially, it was hypothesised that the unusually high number of BRC repeats in *T. brucei* BRCA2 was a consequence of the high levels of recombination needed by *T. brucei* during antigenic variation. This hypothesis was tested by attempting to generate mutants with a reduced number of BRC repeats.

This chapter therefore describes the generation of cell lines with various truncations of BRCA2 expressed, in bloodstream stage *T. brucei*. The importance and contribution of specific motifs within BRCA2 for DNA damage repair, cell cycle progression, recombination and VSG switching were then tested using a number of assays.

5.2 Generation of mutants with reduced numbers of BRC repeats

Given the high number of BRC repeats in *T. brucei* BRCA2, and its function in VSG switching, we wanted to examine why the protein contains so many BRC repeats, since BRCA2 homologues in *U. maydis* and *C. elegans* are able to function similarly to higher eukaryotes with only 1 BRC repeat (Kojic *et al.*, 2002;Martin *et al.*, 2005). A number of different approaches were attempted, including utilising the naturally occurring low number of BRC repeats in BRCA2 orthologues from other *Trypanosoma* species, as well as more complicated cloning methods.

5.2.1 Generation of a re-expresser line with T. vivax BRCA2

It has been demonstrated that not all kinetoplastid BRCA2s have a high number of BRC repeats (see section 3.8.1). Indeed, this phenomenon appears to be potentially unique to *T. brucei* amongst the *Trypanosoma*. *T. congolense* TREU 1457 and *T. vivax* ILDAR 1 BRCA2 homologues have been shown to possess 2 and 1 BRC repeats respectively.

The first strategy to generate *BRCA2* with reduced numbers of BRC repeats was to express *T. vivax BRCA2* in the second of the two independent *brca2-/-* homozygous mutants (*brca2-/-2*) generated in the Lister 427 cell line. An expression construct was generated by PCR-amplifying the ORF of *T. vivax BRCA2* from *T. vivax* ILDAR 1 genomic DNA, using Herculase DNA polymerase (Stratagene) and the primers *TvivBRCA2for2* (which included a HA tag) and *TvivBRCA2rev2*. This PCR product was restriction digested with *Nru*I and ligated into the plasmid pRM482 (R. McCulloch, gift), which had been *Eco*RV-digested and CIP treated. This resulted in the generation of the construct pRM482::*T. vivax* BRCA2 (figure 5.1), which contained the antibiotic resistance cassette for G418, allowing selection for insertion into the tubulin array, where it was transcribed from the endogenous

transcription (as was performed for *BRCA2-/-/+,* section 4.4). Splicing and polyadenylation of the *BRCA2* mRNA was provided by 5' actin and 3' tubulin intergenic sequences, respectively, as opposed to the natural processing signals of the gene.

The pair-wise comparisons shown in section 3.6 demonstrated that the *T. brucei* BRCA2 polypeptide shared 26.2 % sequence identity and 35.2 % sequence similarity with the *T. vivax* BRCA2 polypeptide. The single *T. vivax* BRC repeat, however, contains 31.4 % sequence identity and 51.4 % sequence similarity with the 14 BRC repeats in *T. brucei BRCA2* and contains all the critical residues as highlighted by Lo *et al.*, 2003 (section 3.7.1), demonstrating that the BRC repeat of *T. vivax* should be able to interact with RAD51. However, the possibility remains that the *T. vivax* BRC repeats may not be able to interact with the *T. brucei* RAD51. Further investigations however, demonstrate that the *T. brucei* and *T. vivax* RAD51 polypeptides share 71.7 % sequence identity and 78.8 % sequence similarity, indicating a high probability of interaction (see section 5.3.5, figure 5.19 and table 5.9 for further discussion).
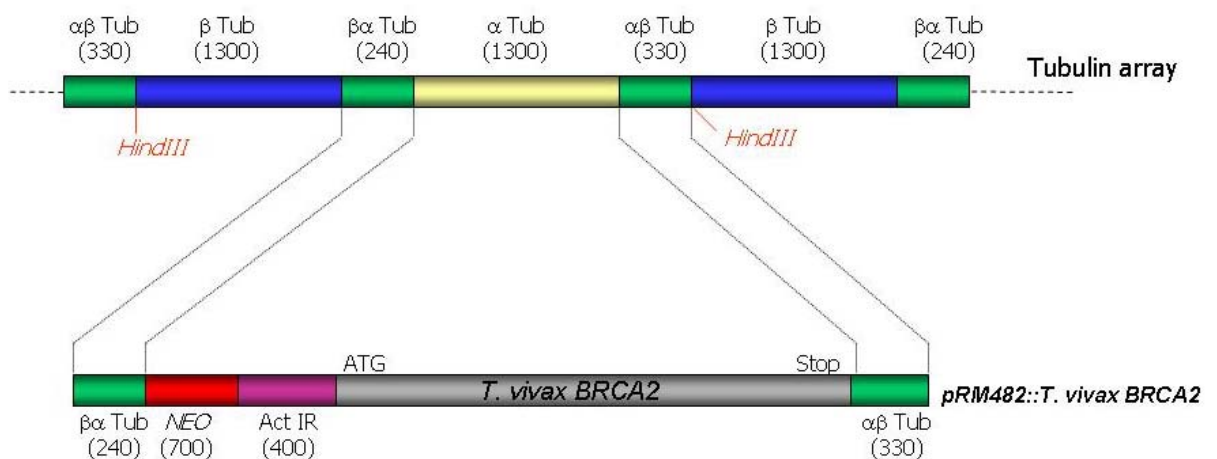
Figure 5.1 – pRM482::*T. vivax BRCA2* construct generated to express *T. vivax* BRCA2 in brca2-/- mutants. *T. vivax* BRCA2 was cloned into an *Eco*RV site between the actin intergenic (Act IR) and βα tubulin (βαTUB) intergenic sequences of the plasmid pRM482, which contains the ORF encoding neomycin phosphotransferase, providing resistance to the antibiotic G418 (NEO). The construct is flanked with tubulin intergenic regions (IR; αβ Tub and βα Tub), which allow homologous integration into the tubulin array, replacing an α tubulin ORF. The size of IR and ORFs are indicated (in bp).

The construct was excised from the plasmid backbone by restriction digestion with *Psp*OMI and *Xba*I, before phenol:chloroform extraction and ethanol precipitation. Approximately 5 μg of digested DNA was transformed into the Lister 427 *brca2-/-2* mutant cell line. Antibiotic resistant transformants were selected by plating out 4 x 10^7 cells from the transformation at 2.5 μg.ml^{-1} G418, over 48 wells with 1.5 mls per well. One G418 resistant transformant was recovered, and the introduction of *T. vivax BRCA2* was confirmed by PCR-amplification of the entire ORF using Taq DNA polymerase and

the primers '*TvivBRCA2for2*' and '*TvivBRCA2rev2*' (data not shown) and Southern
analysis (section 5.2.3.1). This transformant was named *T. vivax BRCA2-/-/+*.

5.2.2 Attempts at creating mutants with reduced numbers of BRC repeats

The second strategy to reduce the number of BRC repeats in *BRCA2,* was to use a
modified version of the *Δ BRCA2::PUR* construct, in which the 3' flank was replaced with
the 5' end of *BRCA2,* up to and including the first few BRC repeats. The prediction was
that when this construct was transformed into *BRCA2+/-* mutants it should be capable of
generating transformants with varying numbers of BRC repeats, due to the construct
integrating into the endogenous copy of *BRCA2* at various positions in the BRC repeat
array by homologous recombination.

To generate this construct, oligonucleotide primers were designed to amplify the 5' end of
the *BRCA2* ORF from Lister 427 genomic DNA, which resulted in a number of different
products being generated that contained varying numbers of BRC repeats (figure 5.2). A
forward oligonucleotide primer (*5'BRC VAR*) was designed, which contained an *Hpa*I
restriction site, a methionine, an HA tag and 21 bases of sequence that was homologous to
the start of *BRCA2*. A reverse oligonucleotide primer (*3'VARPLUSTAG*) was designed
that contained 14 bases of non-*BRCA2* sequence, an *Xba*I restriction site and 21 bases of
sequence that was complementary to amino acids 18-26 of each BRC repeat. A third
oligonucleotide primer (*3'VARTAG*) was also designed. This contained the same non-
BRCA2 sequence and restriction site present in *3'VARPLUSTAG,* so it could only PCR
amplify from products generated by the primer pairs *5'BRC VAR* and *3'VARPLUSTAG.*
The purpose of using 3 oligonucleotide primers in this reaction was to utilise tagged MVR
(mini-satellite variant repeat) mapping, as described in section 3.8.1.1, which allows larger
numbers of repeats to be amplified.

In the PCR reactions, the oligonucleotide primers *5'BRC VAR* and *3'VARTAG* were used at
5 μM, whilst the oligonucleotide primer *3'VARPLUSTAG* was used at a lower
concentration of 0.5 μM, in order to quickly become exhausted. Once amplification had
occurred with the primer pairs *5'BRC VAR* and *3'VARPLUSTAG*, amplification could then
occur using the primers *5'BRC VAR* and *3'VARTAG*. This tagged MVR mapping method
was preferable for generating this product, as using 2 oligonucleotide primers at a higher
concentration would be expected to preferentially amplify smaller numbers of repeats. In
addition to the primers, 25 μl MVR PCR reactions contained 2.5 U of Taq DNA

polymerase, 2.5 U of Herculase DNA polymerase, 2.5 µl of 10 x Herculase reaction buffer and 0.5 mM MgCl$_2$ (ABgene). PCR was performed for 28 cycles of 95 °C for 1 minute, 55 °C for 1 minute and 72 °C for 2.5 minutes, and the products separated by electrophoresis on a 1.5 % agarose gel.
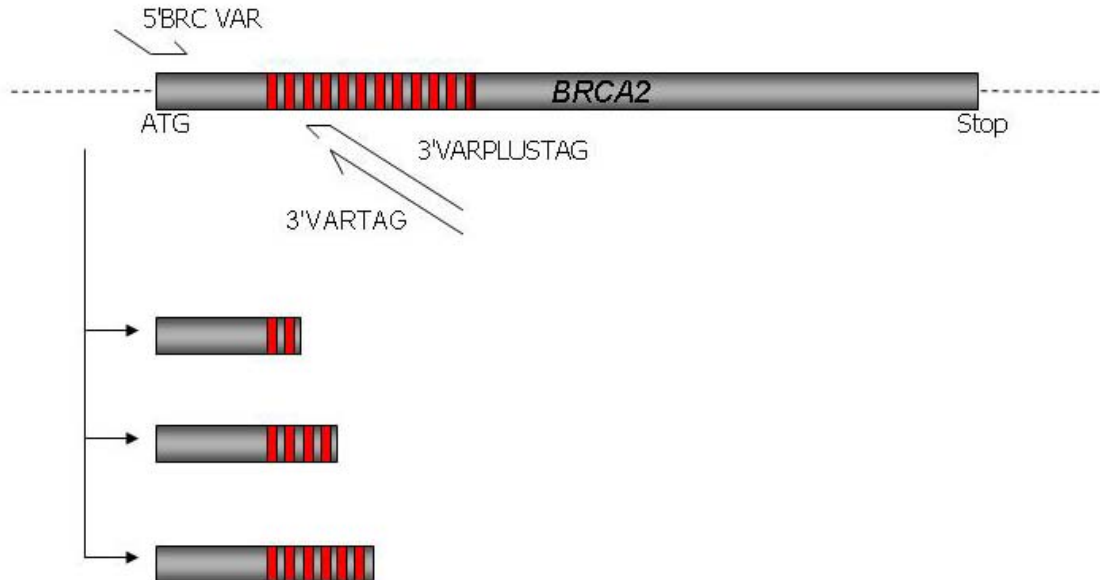
Figure 5.2 – Strategy for obtaining the 5' end of the *BRCA2* ORF containing varying numbers of BRC repeats. PCR amplification of the 5' end of BRCA2 was generated by tagged MVR (mini-satellite variant repeat) mapping, using the oligonucleotide primers 5'BRC VAR, 3'VARPLUSTAG and 3'VARTAG. The tagged MVR mapping should generate various sized products containing different numbers of BRC repeats.

The MVR-PCR generated a ladder of PCR products containing the 5' end of *BRCA2* including various numbers of BRC repeats (data not shown). Products with 1 and 4 BRC repeats were gel extracted and restriction digested with *Hpa*I and *Xba*I. These products were then cloned into the *Δ BRCA2::PUR* construct that had been restriction digested with the same enzymes. This generated the constructs *Δ BRCA2::PURb1* and *Δ BRCA2::PURb2* containing 1 and 4 BRC repeats respectively (figure 5.3).
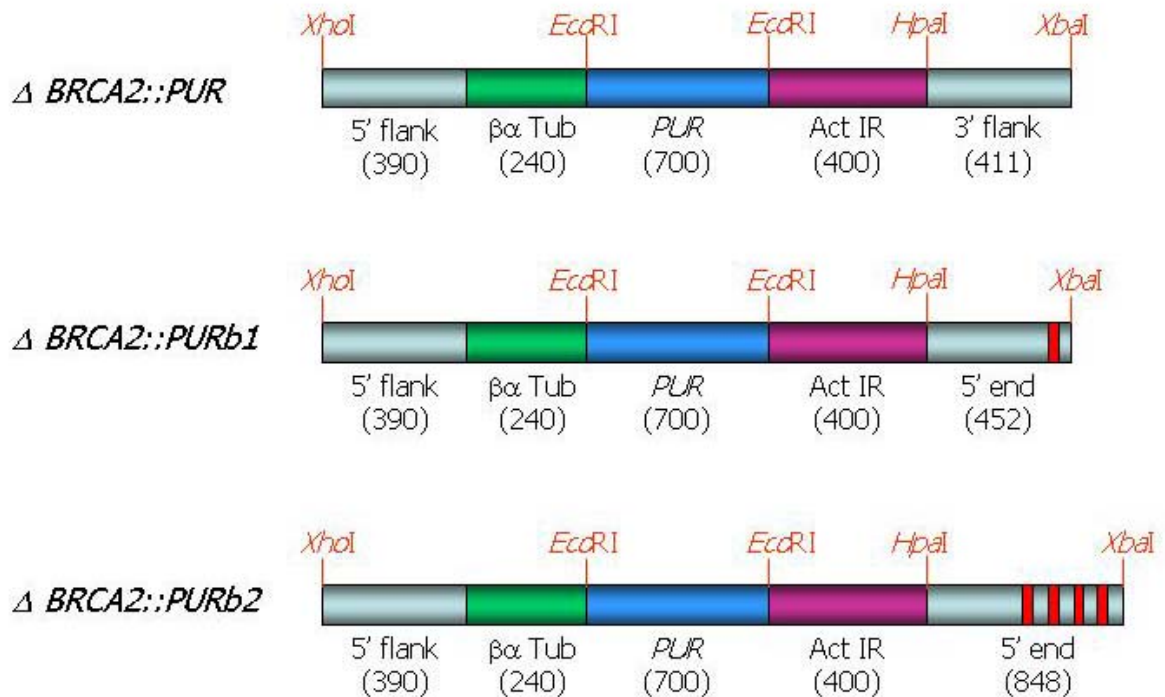
Figure 5.3 – Generation of constructs for transforming into *BRCA2+/-* mutants. Restriction maps of the constructs used for the reduction of BRC repeats by recombination. Sizes of the individual components are shown in base pairs. Constructs were generated by modifying the *Δ BRCA2::PUR* construct (top), replacing the 3' flank with the 5' end of the *BRCA2* ORF. 5' flank and 3' flank correspond to the regions upstream and downstream of the *BRCA2* ORF. In constructs b1 and b2, the 3' flank is replaced by 5' ends corresponding to the start of the *BRCA2* ORF up to and including 1 or 4 BRC repeats (represented by red bars). αβ Tub: αβ tubulin intergenic region (processing signal). ACT IR: Actin intergenic region (processing signal. PUR: puromycin resistance gene ORF.

The constructs were excised from the plasmid backbone by restriction digestion with *Xho*I and *Xba*I, before phenol:chloroform extraction and ethanol precipitation. Approximately 5 μg of digested DNA was transformed into Lister 427 *BRCA2+/-2* cells. Antibiotic resistant transformants were selected by plating out 4×10^7 cells from each transformation at 2.5 μg.ml^{-1} G418, over 48 wells with 1.5 mls per well.

Genomic DNA was prepared from 24 of the resulting transformants (12 from *Δ BRCA2::PURb1* and 12 from *Δ BRCA2::PURb2*) and MVR mapping was performed, as described in section 3.8.1.1 to determine the number of BRC repeats. These constructs should have integrated into the endogenous copy of *BRCA2* by homologous recombination, creating mutants with varying numbers of BRC repeats by recombining at different repeats (see figure 5.4). However, all of the transformants analysed had retained all twelve BRC repeats, suggesting that all recombination events had failed to reduce the number of BRC repeats in *BRCA2*, instead recombining preferentially with the most N-terminal BRC repeat.
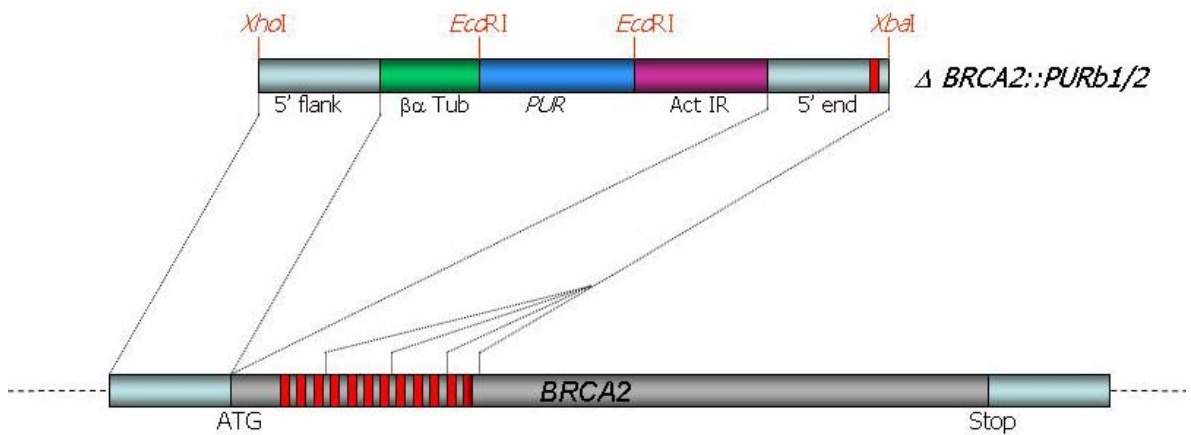
Figure 5.4 – Recombination strategy used to obtain cells with reduced numbers of BRC repeats. 5' flank corresponds to the region upstream of the *BRCA2* ORF, whilst the 5' end corresponds to the start of the *BRCA2* ORF up to and including 1 or 4 BRC repeats (represented by red bars). αβ Tub: αβ tubulin intergenic region (processing signal). ACT IR: Actin intergenic region (processing signal. PUR: puromycin resistance gene ORF. The construct should integrate into the endogenous copy of BRCA2 when introduced into *BRCA2+/-* mutants, recombining between the 5' flank and various BRC repeats. This should result in the generation of transformants with various numbers of BRC repeats.

The reasons that this approach failed are unclear. However, it could be speculated that the transformants which had retained all of the BRC repeats had a growth advantage over those with reduced BRC repeats, which were out competed during recovery. Alternatively, the plasmid flank sequence causes recombination preferentially at the N-terminal BRC repeats.

Given the failure of the above approach, a third strategy to reduce the number of BRC repeats in *BRCA2* was attempted in which the re-expresser construct (*pRM482::BRCA2*) was introduced into wild type *E. coli* cells (DS801, *recA+*) in the expectation that the recombination machinery of the bacteria would reduce the number of BRC repeats. Approximately 1 ng of uncut *pRM482::BRCA2* was transformed into DS801 *E. coli* cells by electroporation (section 2.10). Only 14 transformants were obtained, and all were analysed for the number of BRC repeats in *BRCA2* by MVR-PCR (section 3.8.1.1). In 2 of the transformants, all 12 BRC repeats were retained, whilst the remaining 12 transformants were found to have lost the entire plasmid (as evidenced by the lack of any plasmid DNA). These results suggested that this strategy was also not going to prove to be a successful method on reducing the number of BRC repeats in *BRCA2*.

5.2.3 Generation of a re-expresser line with 1 BRC repeat

The final strategy employed to reduce the number of BRC repeats in *BRCA2* was to generate altered versions of *BRCA2* by cloning methods. To do this the 5' end of the *BRCA2* ORF was PCR-amplified from Lister 427 genomic DNA, using Herculase DNA

polymerase (Stratagene) and the oligonucleotide primers *BRCVAR5'5* and *BRCVAR5'3* (figure 5.5). *BRCVAR5'5* consisted of an *Nru*I restriction site, a methionine, an HA tag and 21 bases of sequence homologous to the start of the ORF. *BRCVAR5'3* consisted of 21 bases of sequence that was complementary to the region immediately upstream of the first BRC repeat; this oligonucleotide was phosphorylated at the 5' end.

The 3' end of *BRCA2* was similarly PCR-amplified using the primers *BRCVAR3'5* and *BRCVAR3'3*. The oligonucleotide primer *BRCVAR3'5* consisted of 24 bases of sequence that was homologous to the 8 amino acids immediately upstream of each BRC repeat; like *BRCVAR5'3*, this oligonucleotide was also phosphorylated at the 5' end. The oligonucleotide primer *BRCVAR3'3* consisted of an *Nru*I restriction site and 24 bases of sequence that was complementary to the end of the *BRCA2* ORF.

PCR-amplification of the 5' end yielded a single DNA fragment of the expected 215 base pairs (data not shown). PCR-amplification of the 3' end yielded only a single DNA fragment of 2882 base pairs, corresponding to amplification of the 3' end, including only the most C terminal BRC repeat (data not shown). Amplification of products from the 3' end with more than 1 BRC repeat were unable to be generated, despite numerous efforts.

Both the above PCR products were gel-extracted and ligated together using the phosphorylated ends of the PCR products. This ligation reaction was subsequently TOPO cloned in order to recover successfully ligated products. The ligated product was excised from the TOPO vector by restriction digesting with *Nru*I and was subsequently ligated into the plasmid pRM482 (R. McCulloch, gift), which had been *Eco*RV-digested and CIP treated. This resulted in the generation of the construct *pRM482::1BRC BRCA2*, which contained the antibiotic resistance cassette for G418 and allowed integration into the tubulin array. The construct was excised from the plasmid backbone by restriction digestion with *Xho*I and *Xba*I, before phenol:chloroform extraction and ethanol precipitation. Approximately 5 µg of digested DNA was transformed into Lister 427 *brca2-/-2* cells. Antibiotic resistant transformants were selected by plating out 4×10^7 cells from each transformation at 2.5 µg.ml^{-1} G418, over 48 wells with 1.5 mls per well. Two G418 resistant transformants were recovered, and the introduction of *BRCA2* with 1 BRC repeat was confirmed by PCR-amplification of the entire ORF using Taq DNA polymerase and the primers '*TbBRCA2for2*' and '*TbBRCA2rev2*' (data not shown). This revealed the presence of a DNA fragment of the expected size, which was significantly smaller than the product generated from WT genomic DNA. This result was subsequently

confirmed by Southern analysis (section 5.2.3.1), and one of these transformants was selected and named *1 BRC BRCA2-/-/+*.
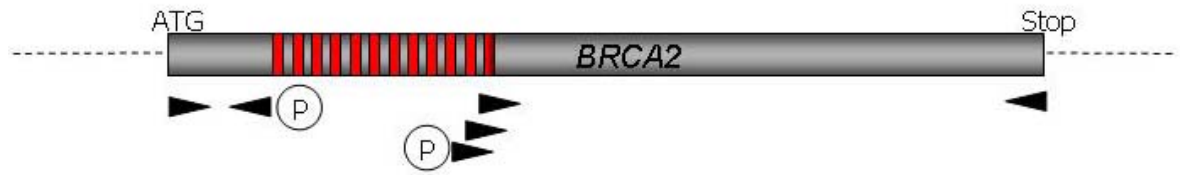
Figure 5.5 – Cloning strategy used to generate the construct *pRM482::1BRC BRCA2.* The 5' and 3' ends of *BRCA2* were PCR-amplified and ligated together using phosphorylated primers. Oligonucleotide primers are depicted by black triangles, (P) indicates that the primer was 5' phosphorylated, red bars represent BRC repeats. Once the 5' and 3' ends were ligated together, the product was cloned into the construct pRM482 to allow the product to be re-expressed in *brca2-/-* mutants.

The distinct variants of BRCA2 with reduced numbers of BRC repeats are depicted in figure 5.6. Ideally, more constructs would have been generated that express further versions with varying numbers of BRC repeats, but these could not be generated within the time scale, so only the effects of *BRCA2* containing 1 BRC repeat, either by using *T. vivax BRCA2* or using *T. brucei BRCA2* containing the most C terminal BRC repeat could be investigated.
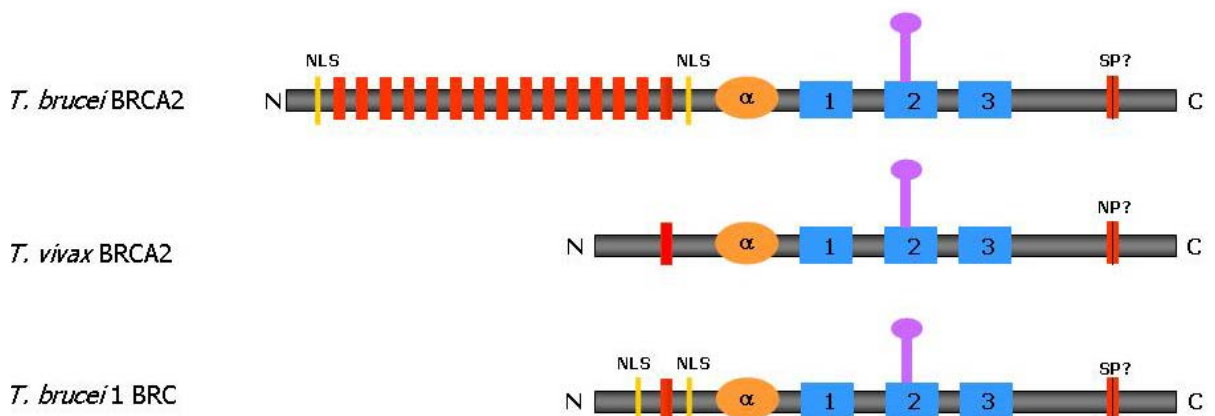
Figure 5.6 – Predicted functional domains of the BRCA2 variants examined in this study. The *T. vivax* BRCA2 homologue and the *T. brucei* BRCA2 containing the most C terminal BRC repeat were re-expressed by cloning into the construct pRM482. The full length *T. brucei* BRCA2 that was used to generate *BRCA2 -/-/+* mutants is shown for comparison. The figure represents the predicted domains of BRCA2 for the trypanosomatids investigated: red bars – BRC repeats; orange oval – alpha helical domain; blue squares – OB domains; purple bar – tower domain; yellow bars with NLS – nuclear localisation signals; red bar with SP? and NP? – possible CDK phosphorylation motif.

5.2.3.1 Confirmation of BRCA2 variant expressers by Southern analysis

In order to confirm that *pRM482::T. vivax BRCA2* and *pRM482::1 BRC BRCA2* had integrated into the tubulin array of the cell lines *T. vivax BRCA2-/-/+* and *1 BRC BRCA2-/-/+* as expected, Southern analysis was carried out. Genomic DNA from wild type, *brca2-/-2*, and expression cell lines was digested with *Hin*dIII before being run out on a 0.8 % agarose gel and transferred to a nylon membrane by Southern blotting. The blots were probed with either a 378 bp region of the *T. brucei* BRCA2 ORF or a 527 bp region of the *T. vivax* BRCA2 ORF (figure 5.7) and the results are shown in figure 5.8. Predicted fragment sizes of correctly integrated constructs are displayed in figure 5.7.
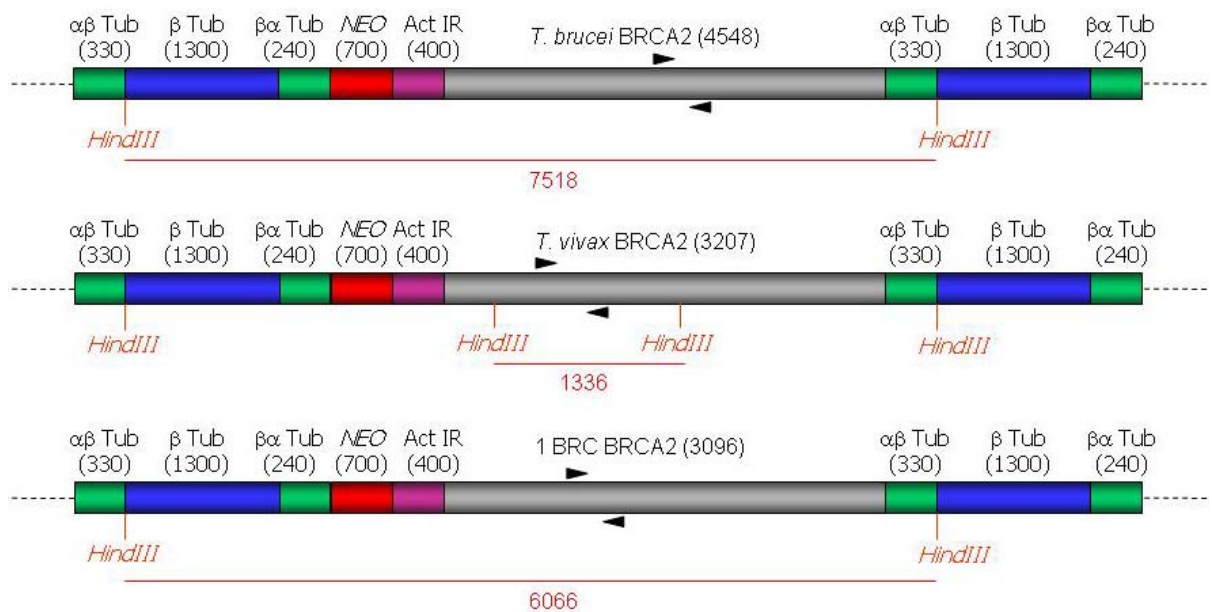
Figure 5.7 – Expressing BRCA2 with reduced BRC repeats from the tubulin array. The constructs generated for re-expressing BRCA2 with reduced numbers of BRC repeats were cloned into an *Eco*RV site between the actin intergenic (Act IR) and βα tubulin (βαTUB) intergenic sequences of the plasmid pRM482, which contains the antibiotic resistance cassettes for G418 (NEO). The constructs are flanked with tubulin intergenic regions (αβ Tub and βα Tub), which allow homologous integration into the tubulin array, replacing an α tubulin ORF. The sizes of IR and ORFs are indicated (in bp), with full length BRCA2 being represented as having 12 BRC repeats (as found in section 3.8.1.1). The predicted maps of each construct following integration is displayed, with the *Hind*III restriction sites and predicted size fragments used to confirm the mutants by Southern analyses shown. Primers used to generate DNA fragments for hybridisation are depicted by black triangles.
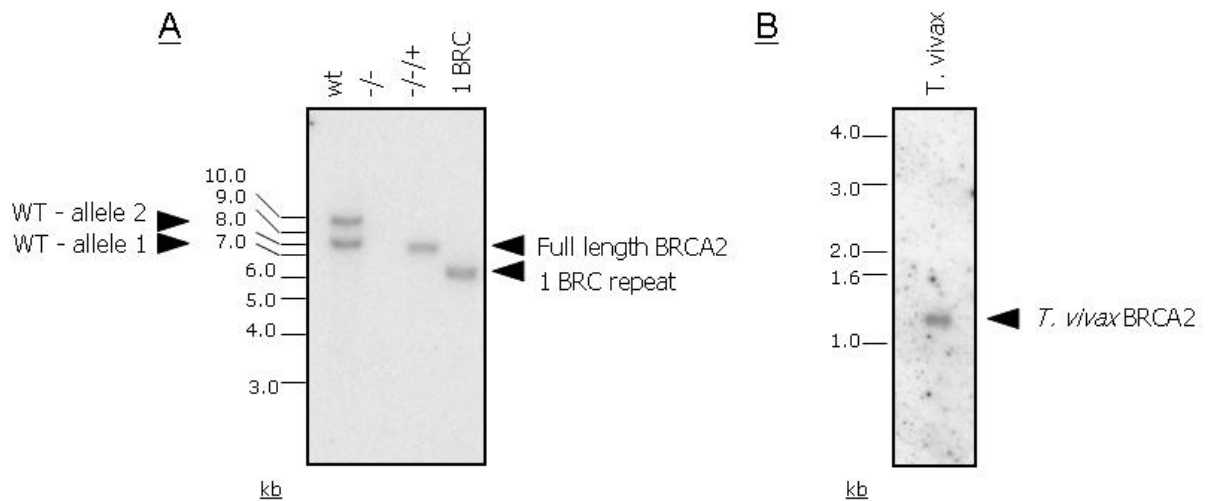
Figure 5.8 – Confirmation of the generation of BRCA2 variants with reduced numbers of BRC repeats by Southern analysis. Genomic DNA from all cell lines was digested with *Hin*dIII and 5 µg run out on a 0.8 % agarose gel. The DNA was Southern blotted before being probed either (A) a 378 bp region of the *BRCA2* ORF or (B) a 527 bp region of the *T. vivax BRCA2* ORF. WT refers to genomic DNA from untransformed cell lines, homozygous mutant is indicated by -/-, full length re-expresser by -/-/+, re-expresser with 1BRC repeat by 1 BRC and the re-expresser with *T. vivax* BRCA2 by T. vivax.

Two allelic *BRCA2* variants for wild type cell lines, and the absence of *BRCA2* in the homozygous mutants, was seen as before (section 4.4.1). Integration of *BRCA2* with 1 BRC repeat and *T. vivax BRCA2* in *brca2-/-* mutants was confirmed by hybridising fragments of the expected sizes. Direct comparison of *1BRC BRCA2 -/-/+* with *BRCA2 -/-/+* confirms the size difference of the *BRCA2* genes through variation in the BRC repeat number.

5.2.3.2 Confirmation of BRCA2 variant expressers by RT-PCR

To support the results of the Southern analyses, RT-PCR was carried out on the re-expresser cell lines with a reduced number of BRC repeats (as described in section 4.2.4). A *T. brucei* BRCA2-specific product, equivalent in size to that generated in wild type and *BRCA2+/-* cell lines, but absent in *brca2-/-* cells (section 4.2.5), was seen in the *1BRC BRCA2 -/-/+* mutant (figure 5.9). For the *T. vivax BRCA2 -/-/+* cells, an equivalent RT-PCR approach was adopted, but using the primers *vivax probe 5'* and *vivax probe 3'*, which revealed the presence of a *T. vivax* BRCA2-specific product (figure 5.9). These data confirm that the appropriate *BRCA2* mRNA was present in the re-expresser mutants for the cell lines *1 BRC BRCA2-/-/+* and *T. vivax BRCA2-/-/+*. As this analysis was non-quantitative, relative amounts of cDNA could not be determined.
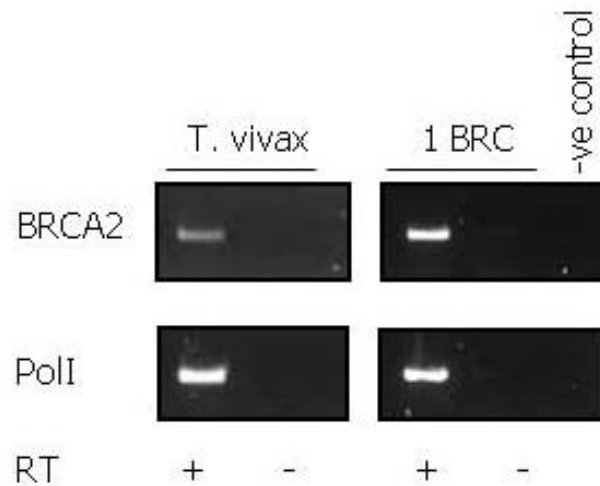
Figure 5.9 – Confirmation of the generation of BRCA2 variants with reduced numbers of BRC repeats by RT-PCR. RT-PCR was carried out on cDNA generated from total RNA from *T. vivax BRCA2 -/-/+* (T. vivax) and *1 BRC BRCA2 -/-/+* (1 BRC). RNA polymerase I specific primers were used to control for the generation of intact cDNA. Primers specific for *T. brucei* BRCA2 were used to show the expression of that gene in 1 BRC re-expressers, whilst primers specific for *T. vivax* BRCA2 were used to show the expression of that gene in *T. vivax* BRCA2 re-expressers. The negative control contains no cDNA substrate. RT + denotes cDNA generated with reverse transcriptase, RT – denotes control reactions that were treated equivalently but no RT was added to the reactions.

5.2.3.3 Confirmation of BRCA2 variant expressers by Northern analysis

Both the constructs *pRM482::T. vivax BRCA2* and *pRM482::1 BRC BRCA2*, included an in-frame N terminal HA epitope, in order to allow for the confirmation of protein expression through anti-HA antisera and to localise proteins using immunofluorescence. Western analysis was carried out on total protein extracted from *T. vivax BRCA2 -/-/+* and *1BRC BRCA2 -/-/+* cell lines. Protein extracts from 1 x 10^7 cells were separated on 10 % SDS-PAGE gels and probed with 2 different monoclonal anti-HA, peroxidise-conjugated antisera (Sigma, H6533 and Roche, 11667475001). Protein expression was undetectable with both antibodies, as was found previously for 3174.2 *BRCA2-/-/+*, 427 *BRCA2-/-/+* and *OEBRCA2* cell lines in section 4.4 (where HA epitopes were similarly used).

In order to determine if *BRCA2* was transcribed in the *T. brucei* re-expressers, northern blots were performed on total RNA isolated from wild type, *BRCA2-/-/+*, *T. vivax BRCA2 -/-/+* and *1BRC BRCA2 -/-/+* cell lines. To do this, total RNA was extracted (RNeasy Mini Kit, Qiagen) from 25 mls of bloodstream stage culture grown to a density of 2 x 10^6 cells.ml^{-1}. The RNA was quantified by spectrophotometry (Beckman DU650 spectrophotometer) before 20 μg samples were separated by electrophoresis on a denaturing formaldehyde gel. The RNA was transferred to a nylon membrane by capillary blotting and blots probed with a 378 bp fragment of the *T. brucei BRCA2* ORF (figure 5.7)

for wild type, *BRCA2-/-/+* and *1BRC BRCA2 -/-/+* cell lines. RNA from the *T. vivax BRCA2-/-/+* cell line was probed with a 527 bp fragment of the *T. vivax BRCA2* ORF (figure 5.7). The hybridising bands generated in each lane were assumed to be mature mRNA, based on their size, and are shown in figure 5.10.
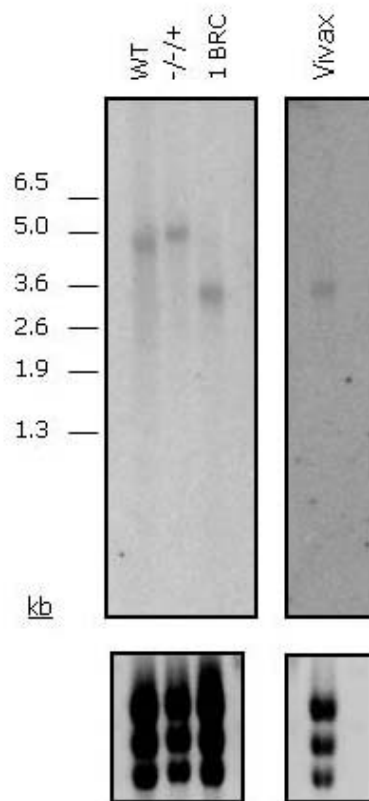
Figure 5.10 – Northern analysis of BRCA2 variants with reduced numbers of BRC repeats. Northern blots are shown of 20 µg of total RNA from wild type (WT), *BRCA2-/-/+* and *1BRC BRCA2-/-/+* probed with a 378 bp region of the *BRCA2* ORF (left blot), and *T. vivax -/-/+* probed with a 527 bp region of the *T. vivax BRCA2* ORF (right blot). Size markers are shown and ethidium stained gels are displayed below the Northern blots to demonstrate amount of RNA loaded.

The northern blot demonstrates that *T. brucei BRCA2* mRNA was detectable for WT, *BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines, whilst *T. vivax BRCA2* mRNA was detectable for *T. vivax BRCA2-/-/+*. Within the limits of the experiment, no substantial differences in mRNA abundance are detectable. Given this mRNA expression, it must be assumed that each protein was undetectable by anti-HA western blot analysis due to low levels of BRCA2 protein abundance, perhaps due to translation efficiency or protein turnover. It does not seem likely that it was due to lower levels of expression from the tubulin array as opposed to the endogenous locus for two reasons. Firstly, the abundance of BRCA2 mRNA in the WT cells is highly comparable with that of the *BRCA2-/-/+* and *1BRC BRCA2-/-/+* expressers. Secondly, the transformants generated in section 5.2.2, which also possessed an HA epitope tag, were transcribed from the endogenous BRCA2 locus and yet protein remained undetectable by western blot analysis (data not shown).

5.3 Phenotypic analysis

The BRCA2 variants with a reduced number of BRC repeats (*T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+*), were analysed for their *in vitro* population doubling times, cell cycle progression, DNA damage sensitivity, recombination efficiency, the ability to form RAD51 foci and VSG switching frequency.

5.3.1 Analysis of in vitro growth

Analysis of *in vitro* growth of the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines was carried out to determine if a reduction in the number of BRC repeats in BRCA2 affected population doubling times. The assay was carried out following the same protocol as described in section 4.3.1. Four repetitions of the growth assay were carried out for each cell line and the results are displayed in figure 5.11, in comparison with the values determined previously for wild type, *brca2-/-2* and *BRCA2-/-/+* cell lines.
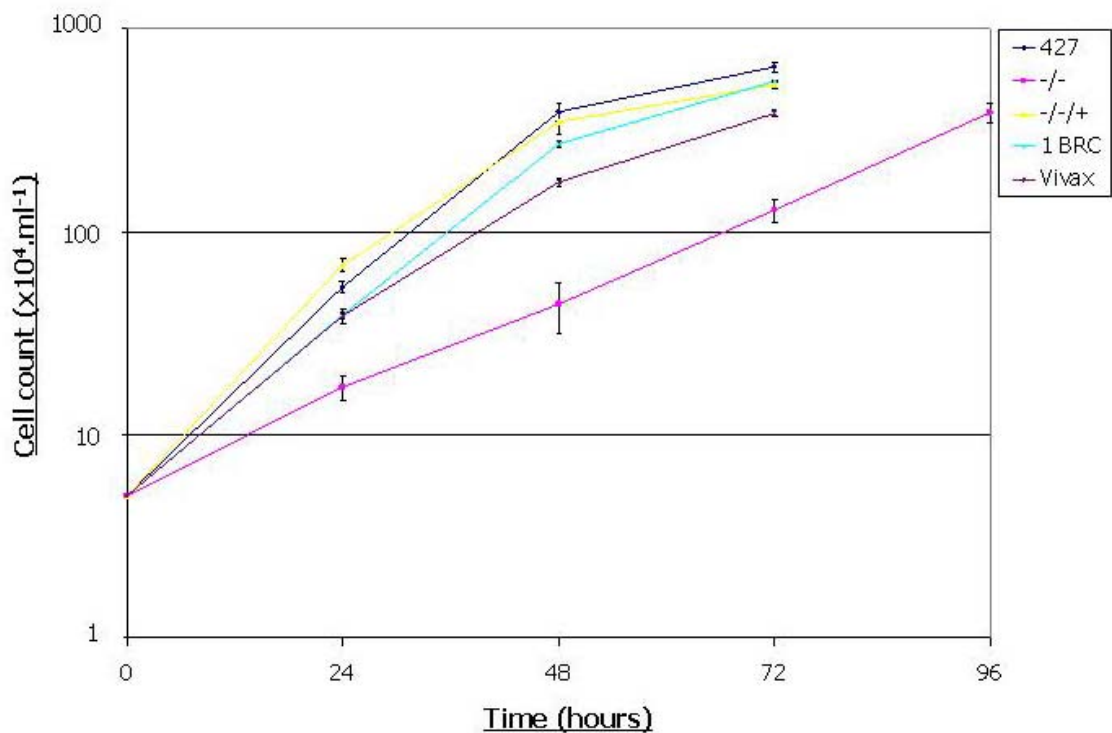
Figure 5.11 – Analysis of *in vitro* growth of BRCA2 variants with reduced number of BRC repeats. 5 ml cultures were set up at 5 x 10^4 cells.ml^{-1} and cell densities counted 24, 48, 72 and 96 hours subsequently. Standard errors are indicated for the counts using data from four repetitions. 427: wild type; -/-: homozygote (*brca2-/-2*); -/-/+: full length re-expresser; 1 BRC: *1BRC BRCA2-/-/+*; Vivax: *T. vivax BRCA2-/-/+*.

Cell line	427	-/- 2	-/-/+	Vivax	1 BRC
Doubling time	8.19+/-0.4	15.50+/-0.34	8.26+/-0.4	9.44+/-0.13	8.41+/-0.14

Table 5.1 – *in vitro* population doubling times of BRCA2 variants with reduced numbers of BRC repeats. The mean doubling time for each of the re-expresser mutants with reduced numbers of BRC repeats are displayed in hours and compared to the population doubling times for WT, *brca2-/-2* and *BRCA2-/-/+* cell lines. 427: wild type; -/-: homozygote (*brca2-/-2*); -/-/+: full length re-expresser; 1 BRC: *1BRC BRCA2-/-/+*; Vivax: *T. vivax BRCA2-/-/+*. Standard errors are indicated.

From the growth curves shown in figure 5.11 and the population doubling times shown in table 5.1, it was apparent that reducing the number of BRC repeats in *T. brucei* BRCA2 had no effect on growth. This is seen by the essentially equivalent growth rates of *BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines, with population doubling times of 8.26 and 8.41 respectively. This result was confirmed by the statistical tests displayed in table 5.2, which revealed that there was no statistical difference between the population doubling times of either WT or *BRCA2 -/-/+* and *1BRC BRCA2-/-/+* (p>0.05). However, expressing the *T. vivax BRCA2* did cause the cells to grow at a slightly slower rate, with the population doubling time increasing by a factor of 1.15 compared to *BRCA2-/-/+* cells. This was manifest as a significant difference in paired t-tests between *T. vivax -/-/+* and WT, *BRCA2 -/-/+* or *1BRC BRCA2 -/-/+* (p<0.05).

	-/-	-/-/+	Vivax	1 BRC
WT	0.0002	0.3820	0.0308	0.5804
-/-		0.0001	0.0001	0.0002
-/-/+			0.0288	0.8882
Vivax				0.0003

Table 5.2 – Statistical analysis of the population doubling times of BRCA2 variants with reduced numbers of BRC repeats. P values are shown for two sample T-tests comparing population doubling times of wild type cells, *brca2* homozygous mutant 2 (-/-), *BRCA2* re-expresser (-/-/+), *T. vivax BRCA2* re-expresser (Vivax) and *1BRC BRCA2* re-expresser (1 BRC). Areas shaded in yellow indicate a significant difference.

It should also be noted, however that the population doubling times from the BRCA2 variant cell lines with reduced numbers of BRC repeats were significantly faster than the *brca2-/-2* mutant. This demonstrates that the re-expression of either *T. vivax BRCA2* or *T. brucei BRCA2* with 1 BRC repeat recovered the impaired growth phenotype observed in *brca2-/-* mutants to a substantial degree.

5.3.2 Analysis of the cell cycle

Analysis of cell cycle progression in the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines was next examined. The assay was carried out following the same protocol as described in section 4.3.3. The results of this are displayed in figure 5.12, in comparison with the values determined previously for wild type, *brca2-/-2* and *BRCA2-/-/+* cell lines.
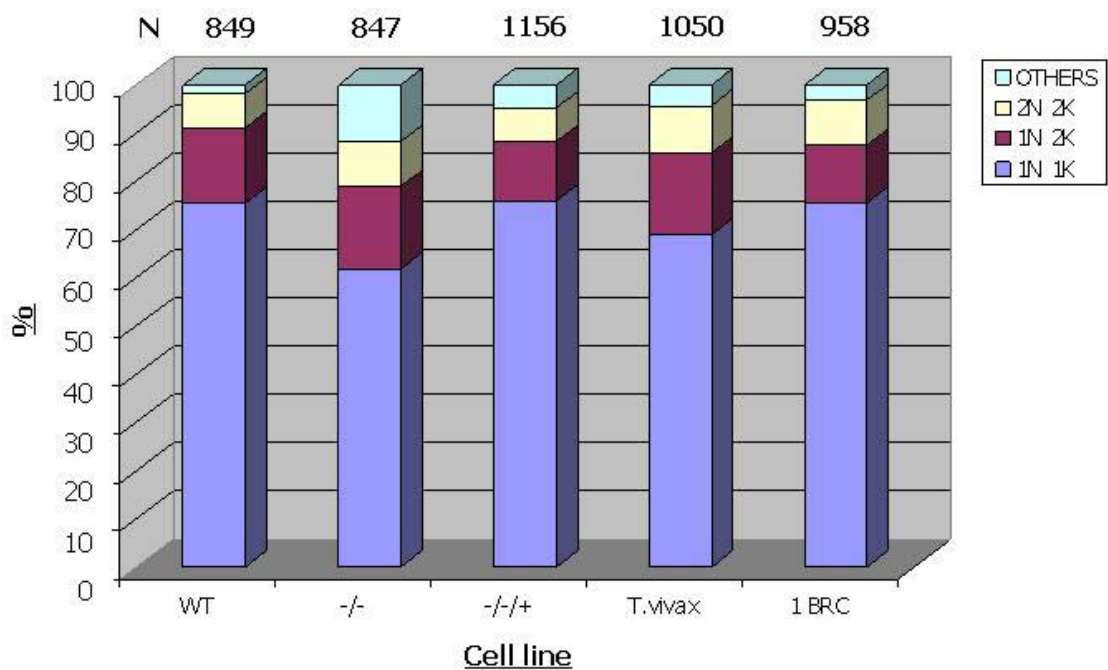
Figure 5.12 – DAPI analysis of BRCA2 variants with reduced number of BRC repeats. The DNA content of *1BRC BRCA2-/-/+* (1BRC) and *T. vivax BRCA2-/-/+* (T.vivax) cell lines were visualised by DAPI and compared with the DNA content of wild type Lister 427, homozygous *brca2* mutants (-/-) and full length *T. brucei BRCA2* re-expressers (-/-/+). The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications cells (others) were counted and represented by their mean count as a percentage of the total cells counted. N = number of cells counted.

This analysis demonstrated that expressing *T. brucei BRCA2* with a reduced number of BRC repeats allowed the cell cycle to progress normally, complementing the accumulation of aberrant cells seen in the *brca2-/-* mutants. The distribution of cells in the *1BRC BRCA2-/-/+* cell line appeared to be essentially equivalent to WT and *BRCA2-/-/+* cell lines, with 75.5 % of cells containing 1N1K compared to 75.4 % and 75.9 %, respectively. Indeed, the number of aberrant cell types was also comparable, with 3.2 % in the *1BRC BRCA2-/-/+* cell line, compared with 1.9 % and 4.7 % in WT and *BRCA2-/-/+* cell lines, respectively. This result was confirmed by the Chi squared analysis shown in table 5.3, which displays that there was no significant difference between the *1BRC BRCA2-/-/+* mutant and the WT and *BRCA2-/-/+* cell lines, with Chi squared values of 1.4 to 2.7 (at P = 0.7003 and 0.4484) respectively. However, a significant difference was observed between the *1BRC BRCA2-/-/+* mutant and the *brca2-/-2* mutant, with a Chi squared value of 10.74 at P = 0.0132.

From the distribution of cells in the *T. vivax BRCA2-/-/+* cell line it was less clear that cell cycle progression occurred normally. Though the number of aberrant cells (4.3 %) was significantly less than in the *brca2-/-2* mutant (11.8 %) and was comparable with WT or *BRCA2-/-/+* cells (1.9 % and 4.7 %, respectively), suggesting that the cell division defect

was complemented, chi-squared analysis suggested that the distribution of cell types overall in the *T. vivax BRCA2-/-/+* cell line was not significantly different from either *brca2-/-2* or WT cells. The basis for this is unclear, but may relate to the fact that the numbers of 1N1K cells (69 %) is intermediate between WT and *brca2-/-2* cells (75.4 % and 61.9 %, respectively). This, in turn could be due to increases in 1N2K and 2N2K cells, which may indicate an impediment in progression through G2 and M phases.

	-/-	-/-/+	Vivax	1 BRC
WT	**55.710** 0.0001	**5.438** 0.1424	**4.830** 0.1847	**2.652** 0.4484
-/-		**9.342** 0.0251	**5.662** 0.2192	**10.744** 0.0132
-/-/+			**3.693** 0.2966	**1.422** 0.7003
Vivax				**2.289** 0.5145

Table 5.3 – Statistical analysis of the cell cycle data for BRCA2 variants with reduced number of BRC repeats. Chi squared analysis of the cell cycle data for wild type cells, *brca2* homozygous mutant 2 (-/-), *1BRC BRCA2-/-/+* (1BRC) and *T. vivax BRCA2-/-/+* (Vivax). The numbers indicated in bold represent the Chi squared value, whilst the numbers below represent the P value at which it was calculated. Areas shaded in yellow indicate a significant difference.

In order to examine if reducing the number of BRC repeats in BRCA2 affected the response of the parasites to DNA damage, the abundance of cells in different cell cycle stages for *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines was measured following growth for 18 hours after phleomycin treatment, as described in section 4.3.3. These results are displayed in figure 5.13 and demonstrate that when *BRCA2* was re-expressed with a reduced number of BRC repeats, the cell cycle progressed somewhat more normally than was observed for *brca2-/-* mutants. The number of 1N 1K cells increased from 45 % in *brca2-/-2* (at 1.0 µg.ml^{-1} of phleomycin) to 66 % and 59.8 % in the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines (at 1.0 µg.ml^{-1} of phleomycin). Conversely, the number of 1N 2K cells decreased from 30.1 % in *brca2-/-2* to 15.9 % and 15.8 % in the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines, indicating enhanced progression through G2 phase. The number of aberrant cells reduced from 19.4 % in *brca2-/-2* to 11.7 % in the *T. vivax BRCA2-/-/+* cell line, but remained essentially equivalent in the *1BRC BRCA2-/-/+* cell line with 21.3 % aberrant cell types.

The statistical analysis shown in table 5.4 demonstrates a significant difference between *brca2-/-* mutants and *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines, with Chi squared values of 11.8 to 34.9 (at P = 0.0083 and 0.0001). This confirms that re-expressing *BRCA2* with a reduced number of BRC repeats, allowed the cells to progress through the cell cycle more normally than in *brca2-/-* mutants in the presence of

phleomycin-induced DNA damage. However, a significant difference was also displayed between WT cells and *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines (at 1.0 µg.ml[-1] of phleomycin), with Chi squared values of 9.4 to 50 (at P = 0.0242 and 0.0001), demonstrating that the distribution of cells in these mutants was not equivalent to wild type cells.

These results therefore indicate that the expression of *BRCA2* with a reduced number of BRC repeats appears, to a certain degree, to alleviate the impairment in nuclear DNA replication observed in the DNA damaged *brca2-/-* mutants. This phenotype most likely occurs by allowing DNA damage to be repaired more effectively than in *brca2-/-* mutants, but less effectively than either WT or *BRCA2-/-/+* cell lines, consistent with an increased sensitivity to DNA damaging agents, which is examined below (see section 5.3.3).
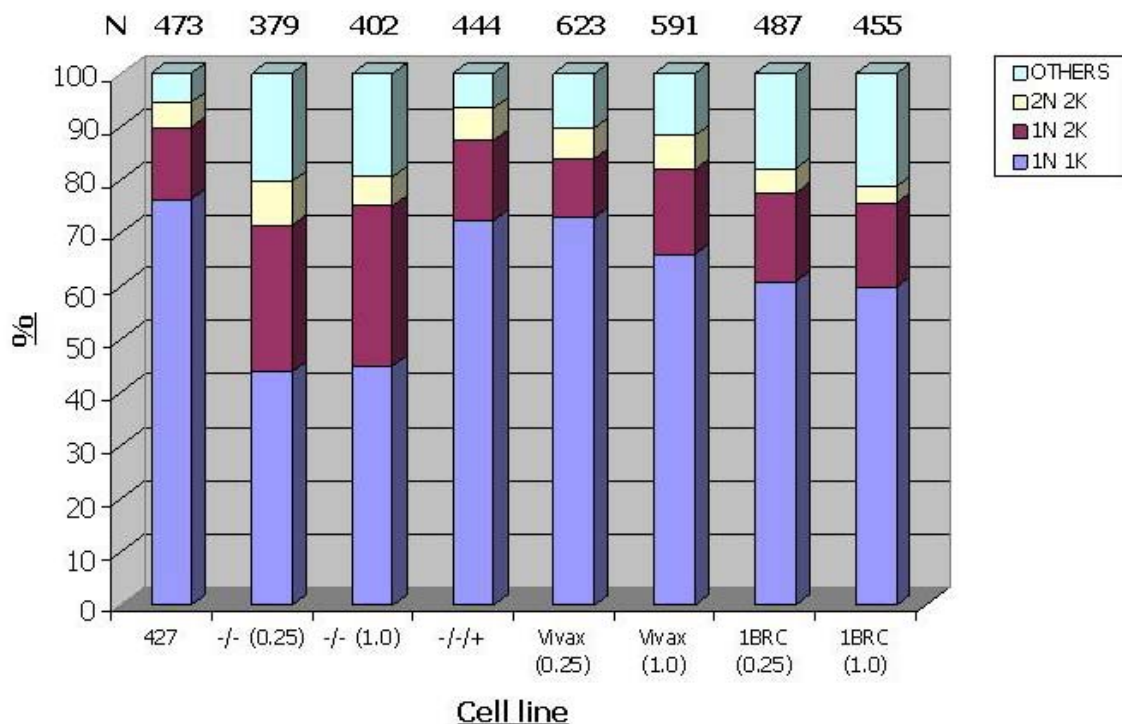
Figure 5.13 – DAPI analysis of BRCA2 variants with reduced number of BRC repeats after DNA damage. The DNA content of *1BRC BRCA2-/-/+* (1BRC) and *T. vivax BRCA2-/-/+* (T.vivax) cell lines were visualised by DAPI and compared with the DNA content of wild type Lister 427, homozygous (-/-) and full length *T. brucei BRCA2* re-expresser (-/-/+). The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications cells (others) were counted and represented by their mean count as a percentage of the total cells counted. Wild type and *BRCA2 -/-/+* cell lines were grown in media with 1.0 µg.ml[-1] of phleomycin, whilst *brca2-/-*, *1BRC BRCA2-/-/+* and *T. vivax BRCA2-/-/+* were grown in media with 0.25 µg.ml[-1] and 1.0 µg.ml[-1] of phleomycin. N = number of cells counted.

	-/- (0.25)	-/- (1.0)	-/-/+	Vivax (0.25)	Vivax (1.0)	1 BRC (0.25)	1 BRC (1.0)
WT	**71.310** 0.0001	**68.395** 0.0001	**1.0240** 0.7954	**5.0110** 0.1710	**9.4200** 0.0242	**31.6970** 0.0001	**50.0060** 0.0001
-/- (0.25)		**1.1820** 0.7573	**34.1830** 0.0001	**34.9400** 0.0001	**20.3170** 0.0001	**12.7660** 0.0052	**14.1540** 0.0027
-/- (1.0)			**32.6060** 0.0001	**33.5450** 0.0001	**19.6940** 0.0002	**11.7600** 0.0083	**12.8400** 0.0050
-/-/+				**3.2270** 0.3579	**4.6610** 0.1983	**21.9900** 0.0001	**37.1820** 0.0001
Vivax (0.25)					**3.0190** 0.3887	**10.6040** 0.0141	**17.5300** 0.0006
Vivax (1.0)						**4.1930** 0.2414	**10.3100** 0.0161
1 BRC (0.25)							**1.2920** 0.7311

Table 5.4 – Statistical analysis of the cell cycle data for BRCA2 variants with reduced number of BRC repeats after DNA damage. Chi squared analysis of the cell cycle data for wild type cells, *brca2* homozygous mutant 2 (-/-), *1BRC BRCA2-/-/+* (1BRC) and *T. vivax BRCA2-/-/+* (Vivax). The numbers indicated in bold represent the Chi squared value, whilst the numbers below represent the P value at which it was calculated. Areas shaded in yellow indicate a significant difference.

5.3.3 Analysis of DNA damage sensitivity

To follow up the above analysis, the DNA damage sensitivity of the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines was the next phenotype examined. To do this, Alamar Blue assays were carried out following the same protocols as described in section 4.3.4, using both MMS and phleomycin as DNA damaging agents. These results are displayed in figures 5.14 and 5.15, and demonstrate that both the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines are more sensitive to DNA damage than the full length *BRCA2* re-expresser.

When MMS was used as the DNA damaging agent, the *1BRC BRCA2-/-/+* mutant has a mean IC50 virtually identical to that observed in WT cells (0.0015 % MMS), but more sensitive than the 0.0031 % displayed in *BRCA2-/-/+* cells. This result was confirmed by the statistical analysis shown in table 5.5, which displays that the *1BRC BRCA2-/-/+* mutant was not significantly different from WT cells (p>0.05), but was significantly different than *BRCA2-/-/+* and *brca2-/-* mutants (p<0.05). The *T. vivax BRCA2-/-/+* cells, however, were more sensitive to MMS than WT and *BRCA2-/-/+* cells, with a mean IC50 of 0.0009 %. This result was confirmed by statistical analysis.
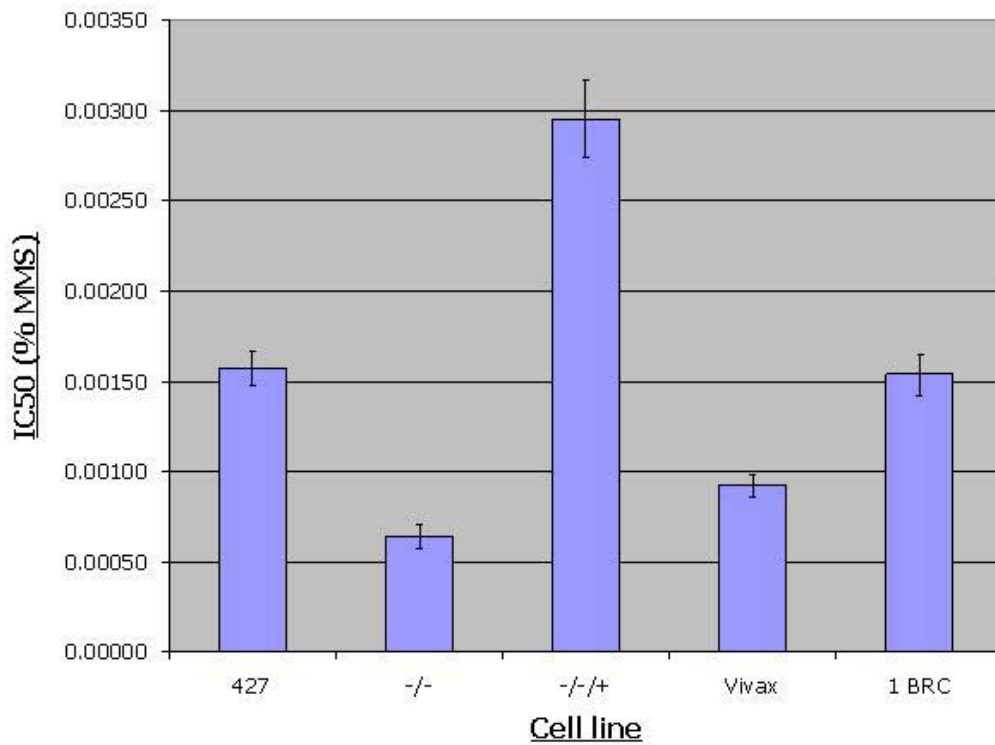
Figure 5.14 – IC50s of *T. brucei* BRCA2 variants with reduced number of BRC repeats exposed to MMS. *1BRC BRCA2-/-/+* (1BRC) and *T. vivax BRCA2 -/-/+* (Vivax) cell lines were placed in serially decreasing amounts of MMS and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments and are compared to the previous results from wild type (427), *brca2-/-2* (-/-) and *BRCA2-/-/+* (-/-/+) cell lines; bars indicate standard error.

	-/-	-/-/+	Vivax	1 BRC
WT	0.0025	0.0192	0.0101	0.8585
-/-		0.0024	0.0035	0.0146
-/-/+			0.0033	0.0047
Vivax				0.0383

Table 5.5 – Statistical analysis of the Alamar Blue results for MMS. P values are shown for two sample T-tests comparing the IC50s for MMS sensitivity of wild type cells (WT), *brca2* homozygous mutant 2 (-/-), *BRCA2-/-/+* (-/-/+), *1BRC BRCA2-/-/+* (1BRC) and *T. vivax BRCA2 -/-/+* (Vivax) mutants. Areas shaded in yellow indicate a significant difference.

When phleomycin was used as the DNA damaging agent, both the *1BRC BRCA2-/-/+* and *T. vivax BRCA2-/-/+* mutants displayed a greater level of sensitivity than was observed for *BRCA2-/-/+* or WT cells. Indeed, the mean IC50s appeared more reminiscent of the results obtained for the *brca2-/-* mutants, with 0.029 μM and 0.02 μM compared to 0.013 μM respectively. The statistical analysis shown in table 5.6 confirmed that there was no statistical difference between *brca2-/-* and *T. vivax BRCA2-/-/+* mutants (p>0.05), whilst both *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* were significantly different from WT and *BRCA2-/-/+* cells (p<0.05).
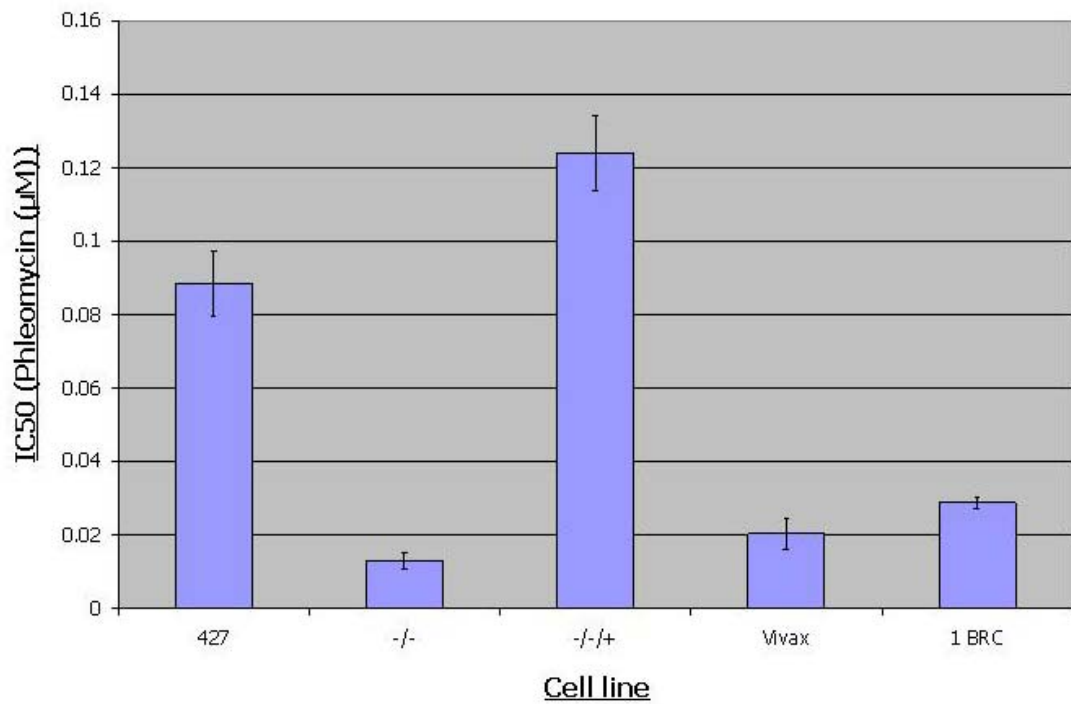
Figure 5.15 – IC50s of *T. brucei* BRCA2 variants with reduced number of BRC repeats exposed to phleomycin. *1BRC BRCA2-/-/+* (1BRC) and *T. vivax BRCA2 -/-/+* (Vivax) cell lines were placed in serially decreasing amounts of phleomycin and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments and are compared to the previous results from wild type (427), *brca2-/-2* (-/-) and *BRCA2-/-/+* (-/-/+) cell lines; bars indicate standard error.

	-/-	-/-/+	Vivax	1 BRC
WT	0.0065	0.0840	0.0206	0.0101
-/-		0.0002	0.2846	0.0029
-/-/+			0.0001	0.0005
Vivax				0.2378

Table 5.6 – Statistical analysis of the Alamar Blue results for phleomycin. P values are shown for two sample T-tests comparing the IC50s for phleomycin sensitivity of wild type cells (WT), *brca2* homozygous mutant 2 (-/-), *BRCA2-/-/+* (-/-/+), *1BRC BRCA2-/-/+* (1BRC) and *T. vivax BRCA2 -/-/+* (Vivax) mutants. Areas shaded in yellow indicate a significant difference.

Surprisingly from these data, the *T. vivax BRCA2-/-/+* cells were more sensitive to both MMS and phleomycin than the *1 BRC BRCA2-/-/+* mutant. The *1 BRC BRCA2-/-/+* cells may have been expected to be more sensitive to phleomycin than the *T. vivax BRCA2-/-/+* cells, given the results of the cell cycle analysis after phleomycin treatment (section 5.3.2), where the *1 BRC BRCA2-/-/+* cells displayed a more severe phenotype than *T. vivax BRCA2-/-/+*, which was similar to the *brca2-/-* mutants, indicative of a greater level of sensitivity to this DNA damaging agent. It is possible that this is due to the residual cell cycle abnormalities that were seen in the absence of damage in the *T. vivax BRCA2-/-/+*

cells (section 5.3.2), and the Alamar blue data give a clearer measure of repair efficiency. However, the basis for the cell cycle deficiencies in *T. vivax BRCA2-/-/+* are not clear, as noted.

The *1 BRC BRCA2-/-/+* cell lines DNA damage sensitivity phenotype is reminiscent of the *BRCA2+/-* mutants, and *mre11-/-* mutants (Robinson *et al.*, 2002), in that they were clearly sensitive to phleomycin but not to MMS. This phenomenon is most likely due to the different modes of action of the DNA damaging agents, whereby MMS-induced lesions can be repaired *via* base excision repair (BER) (Lindahl and Wood, 1999), but phleomycin creates DNA breaks directly that must be acted upon by recombination pathways. This does not, however, explain why the *T. vivax BRCA2-/-/+* mutant displays high levels of sensitivity to both DNA damaging agents.

Taken together, these results indicate that reducing the number of BRC repeats in *BRCA2* to a single repeat causes *T. brucei* cells to be more sensitive to DNA damaging agents, which leads us to suggest that the BRC repeat expansion in BRCA2 is necessary for efficient DNA repair.

5.3.4 Analysis of recombination efficiency

To examine if the relationship between BRC repeat organisation and DNA repair is mirrored in the role of *T. brucei* recombination, the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines were next examined for their transformation efficiency following the same protocol described in section 4.3.5. Three repetitions of the transformation efficiency assay were carried out for each cell line and are shown in figure 5.16, alongside the transformation efficiency rates determined previously for the wild type, *brca2-/-2* and *BRCA2-/-/+* cell lines, to allow for comparison.

Both the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines were found to have significantly worse transformation efficiency rates than wild type or *BRCA2-/-/+* cell lines, with mean transformation efficiency rates of 0.93 and 1.0 compared to 4.53 and 4.27, respectively. These differences were confirmed as being statistically significant by two sample T-tests, displayed in table 5.7. Notably, however, both *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cells also displayed significantly higher transformation efficiency rates than *brca2-/-* mutants, in which a transformation efficiency rate of only 0.15 was observed.

These results indicate that a reduction in BRC repeats in BRCA2 impairs the *T. brucei* cells' ability to transform DNA constructs into their genome. The presence of 1 BRC

repeat does, however, allow transformations to occur at a higher level than was observed for *brca2* null mutants. These results appear to support the hypothesis that the BRC repeat expansion in BRCA2 allows homologous recombination to occur at a high level, and appear to be consistent with the comparable efficiency of DNA repair described in 5.3.3
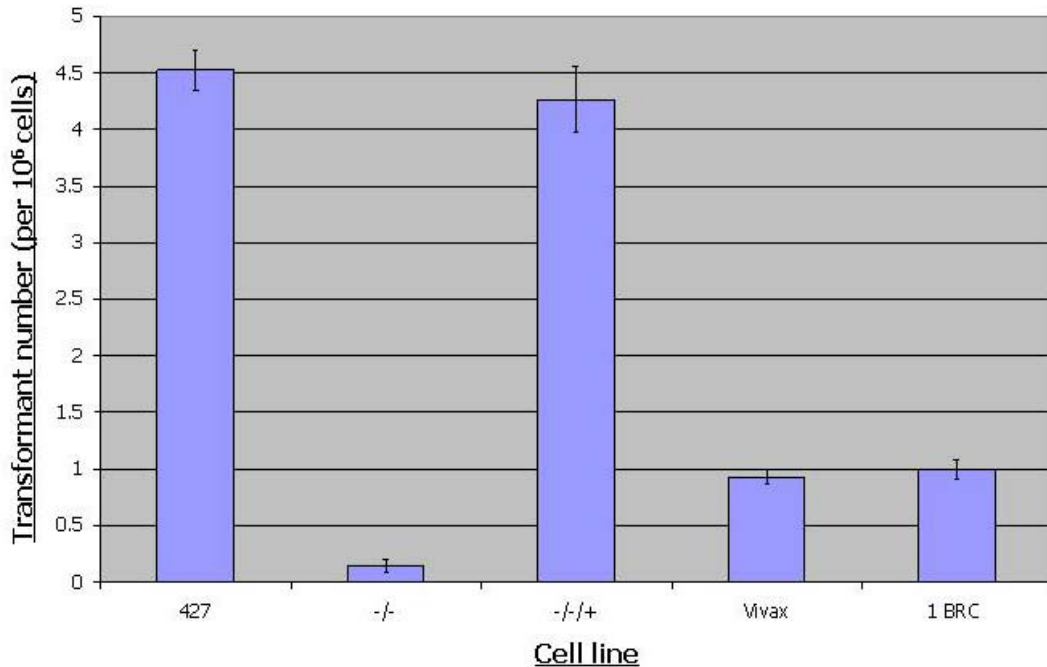
Figure 5.16 – Recombination efficiency in BRCA2 variants with reduced number of BRC repeats. Values are mean numbers of transformants obtained per 10^6 cells transformed; error bars are shown from 3 repetitions. The data are presented for wild type Lister 427 cells (427), *brca2* homozygous mutant 2 (-/-), *BRCA2* re-expresser (-/-/+), *T. vivax* BRCA2 re-expresser (Vivax) and *1 BRC BRCA2* re-expresser (1 BRC).

	-/-	-/-/+	Vivax	1 BRC
WT	0.0025	0.6254	0.0041	0.0067
-/-		0.0040	0.0124	0.0136
-/-/+			0.0052	0.0029
Vivax				0.4226

Table 5.7 – Statistical analysis of the recombination efficiency of BRCA2 variants with reduced number of BRC repeats. P values are shown for two sample T-tests comparing recombination efficiencies of wild type (WT) cells, *brca2* homozygous mutant 2 (-/-), *BRCA2* re-expresser (-/-/+), *T. vivax* BRCA2 re-expresser (Vivax) and *1 BRC BRCA2* re-expresser (1 BRC). Areas shaded in yellow indicate a significant difference.

5.3.5 Analysis of the ability to form RAD51 foci

In order to attempt to understand why the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cells display recombination and repair impairments, the ability of the cells to form RAD51 sub-nuclear foci following DNA damage was assessed. To do this, the cells were treated with phleomycin and allowed to grow for 18 hours, before RAD51 localisation examined by indirect immunofluorescence, as described in section 4.3.6.

Approximately 300 cells were counted and scored for the number of foci they contained after treatment with 2 concentrations of phleomycin (0.25 μg.ml^{-1} and 1.0 μg.ml^{-1}). As for the *brca2-/-* mutants, the cells were treated with an additional lower concentration of phleomycin due to their increased levels of sensitivity to this DNA damaging agent (section 5.3.3). The results of this analysis are displayed in table 5.8 and examples of cells displaying foci are shown in figure 5.17.

	BLE	Number of foci (%)						
		0	1	2	3	4	5	6 or more
WT	0.0	96.4	3.6	0.0	0.0	0.0	0.0	0.0
	1.0	24.8	22.6	18.8	16.5	13.5	2.3	1.5
-/-	0.0	98.2	1.3	0.4	0.0	0.0	0.0	0.0
	0.25	96.9	2.0	0.5	0.5	0.0	0.0	0.0
	1.0	98.6	0.7	0.7	0.0	0.0	0.0	0.0
-/-/+	0.0	87.5	3.5	2.3	2.3	1.9	1.6	0.8
	1.0	14.8	10.5	15.6	14.8	13.7	13.3	17.2
1BRC	0.0	99.7	0.3	0.0	0.0	0.0	0.0	0.0
	0.25	100.0	0.0	0.0	0.0	0.0	0.0	0.0
	1.0	99.7	0.3	0.0	0.0	0.0	0.0	0.0
Vivax	0.0	98.2	1.2	0.6	0.0	0.0	0.0	0.0
	0.25	98.4	1.0	0.6	0.0	0.0	0.0	0.0
	1.0	97.9	1.8	0.3	0.0	0.0	0.0	0.0

Table 5.8 – RAD51 foci formation in BRCA2 variants with reduced number of BRC repeats. The percentages of cells showing foci at given concentrations of phleomycin (BLE) are shown. Phleomycin concentrations are shown in μg.ml^{-1}. Boxes shaded in light yellow contain foci, whilst boxes shaded in bright yellow contain the highest percentage of foci.

Both the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines were found to have a greatly reduced ability to form RAD51 foci, with the majority of cells containing no foci, at either phleomycin concentration. This result is comparable with that observed for *brca2-/-* mutants, where there was also no clear induction of RAD51 foci formation at all, due to the very small percentage of cells with foci appearing to be no different in treated or untreated cells (around 1-3 %).

To ensure that these results do not simply result from decreased RAD51 levels in the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cells, western analysis was carried out on total

protein extracted from all *BRCA2* cell lines, before and after phleomycin-induced damage. Cell extracts from 1 x 10^7 cells were separated on 10 % SDS-PAGE gels, blotted and probed with polyclonal anti-RAD51 antiserum and detected with HRP-coupled anti-rabbit IgG. Equivalent quantities of cell extracts were also separated on 10 % SDS-PAGE gels and stained with coomasie to ensure the equivalent loading of samples. Figure 5.18 demonstrates that RAD51 was still clearly expressed in the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines, and there was no evidence for an increase or decrease in RAD51 levels after DNA damage.
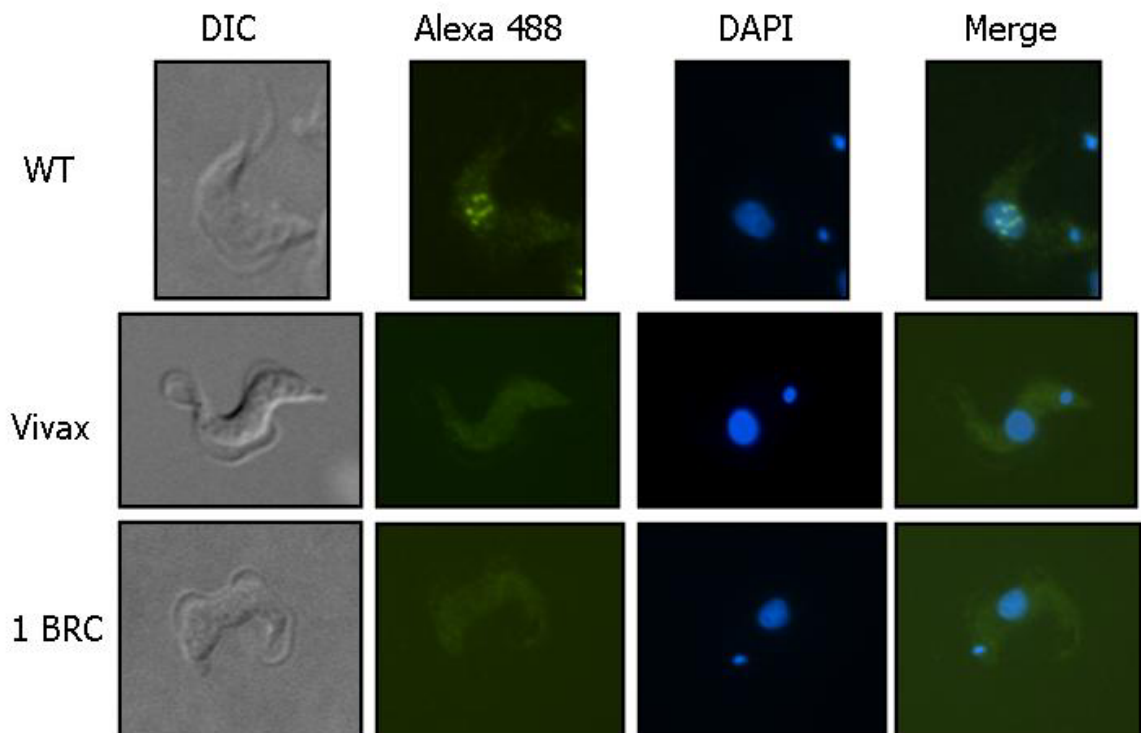
Figure 5.17 – RAD51 immunolocalisation in BRCA2 variants with reduced number of BRC repeats. Representative images of *T. brucei* cells following growth in 0.25 µg.ml^{-1} and 1.0µg.ml^{-1} phleomycin for 18 hours are shown. Each cell is shown in differential interface contrast (DIC), after staining with DAPI and after hybridisation with anti-RAD51 antiserum and secondary hybridisation with Alexa Fluor 488 conjugate (Alexa 488). Merged images of DAPI and Alexa 488 cells are also shown. Wild type (WT), *T. vivax BRCA2-/-/+* (Vivax) and *T. brucei 1BRC BRCA2-/-/+* (1 BRC) cells are shown.
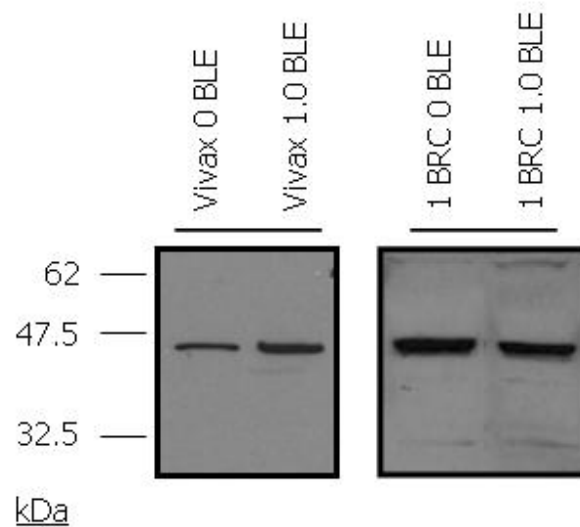
Figure 5.18 – Western blots of RAD51 in BRCA2 variants with reduced number of BRC repeats. The western blots display total protein extracts from *T. vivax BRCA2-/-/+* (Vivax) and *1 BRC BRCA2-/-/+* (1 BRC) cells probed with anti-RAD51 antiserum (RAD51). Protein extracts were prepared without damage (0μg.ml^{-1} BLE) and with damage (1.0μg.ml^{-1} BLE). Size markers are indicated. The endogenous copy of RAD51 is visible at 47kDa.

It has already been demonstrated that the *T. vivax* BRC repeat contains all of the critical residues highlighted by Lo *et al*., (2003), which should allow it to interact with *T. vivax* RAD51. However, we cannot be sure that a BRC repeat from *T. vivax* will interact with *T. brucei* RAD51. In order to address this, a global multiple alignment of the *T. brucei* RAD51 polypeptide (AAD51713) with RAD51 orthologues from *T. vivax* (tviv626b11.p1k_13)*, T. cruzi* (AAZ94621) and *L. major* (LmjF28.0550) was produced using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003). This was then visualised using the Boxshade server (http://www.ch.embnet.org/software/BOX_form.html), as shown in figure 5.19. The alignment shows that a high level of conservation is observed throughout the RAD51 polypeptides from the trypanosomatids, with the 8 critical residues for BRC binding, highlighted by Lo *et al.,* (2003) appearing to be very well conserved, indicating functional conservation.

Pair-wise comparisons were performed using AlignX (Vector NTI), and the percentage sequence identities calculated (see table 3.6). The pair-wise comparisons show that the *T. brucei* RAD51 polypeptide shares the highest level of sequence identity with the *T. vivax* orthologue, with 71.7 % sequence identity.

```
Tcruzi     1 --------------------------------------MNTRSKRGKRKG---VEEVEVHE
Lmajor     1 --------------------------------------MQTRSKAKGRRGRPSARPSEEVE
Tvivax     1 MPLVELYESGASASWQGCLLDSSSCIDKRGCLSAEKKMSTRTRGRKRTKP--AVEEEVHE
Tbrucei    1 --------------------------------------MNTRTKNKKRTKE--VIEDEVHD

Tcruzi    21 IANTSPDPVAAPEQQQQQQEDQQ---NVDGANNGGFRVIQVLESYGIASADIKKLMESG
Lmajor    24 VVESQPQEVLQNEEQEPQQQQQQQSTDMAEPNASGFRVIQILENYGVASSDIKKLMECG
Tvivax    59 IVESLNGDASMRGSQENIQDQLQQG--VDNASGNFRVIQTLENYGIASADIKKLMESG
Tbrucei   22 IDDTAFDDAAVDAVNDNTQEMQQQVG--DAAGGPSFRVLQIMENYGVASADIKKLMECG

Tcruzi    77 F YTVESVAYAPKKNILAVKGISETKADKIMAECAKLVPMGFTSAVVYHEARKEIIMVTT
Lmajor    83 F YTVESAAYAPKKAILAVKGISENKAEKIMAECAKLVPMGFTSAVAYHEARKEIIMVTT
Tvivax   116 F YTVEAVAYAPKKNLLAVKGISEAKVEKIMAECAKLVPMGFASAIVYHEARKEIIMVTT
Tbrucei   79 F LTVESVAYAPKKSILAVKGISEAKAEKIMAECCKLTPMGFTRATVFQEQRKETIMVTT

Tcruzi   136 GS REVDKLLGGGIETGGITELLGEFRTGKTQLCHTLCVTCQLPISQGGAEGMALYIDTE
Lmajor   142 GS REVDKLLGGGIETGSITELFGEFRTGKTQLCHTLCVTCQLPISQGGAEGMALYIDTE
Tvivax   175 GS REVDKLLGGGIETGSITELFGEFRTGKTQLCHTLCVTCQLPISNGGAEGMALYIDTE
Tbrucei  138 GS REVDKLLGGGIEVGSITELFGEFRTGKTQLCHTLCVTCQLPLSQGGGEGMALYIDTE

                                          A A        L Y  SAM
Tcruzi   195 GTF RPERLVAVAERYKLDPQDVLSNVACARAFNTDHQQQLLLQASAMMAENRFAIIIVD
Lmajor   201 GTF RPERLVAVAERYKLDPEDVLANVACARAFNTDHQQQLLLQASAMMAENRFALIIVD
Tvivax   234 GTF RPERLVAVAERYKLDAQDVLANVACARAYNSDHQQNLLVQASAMMAENRFAIIIVD
Tbrucei  197 GTF RPERLVAVAERYSLDPEAVLENVACARAYNTDHQQQLLLQASATMAEHRVAIIVVD

                                          M
Tcruzi   254 SATA LYRTDYSGRNELAARQMHLGKFLRSLHNLAEEYGVAVVVTNQVVANVDGSAQMFQ
Lmajor   260 SATA LYRTDYSGRNELAARQMHLGKFLRSLHNLAEEYGVAVVVTNQVVANVDGSAQMFQ
Tvivax   293 SATA LYRTDYNGRSELAARQMHLGKFLRSLQNLAEEYGVAVVVTNQVVANVDGSAQMFQ
Tbrucei  256 SATA LYRTDYNGRGELAARQMHLGKFLRSLRNLANEYNVAVVVTNQVVANVDGAAPTFQ

Tcruzi   313 ADAKK PIGGHIMAHASTTRLSLRKGRGEQRIMKVYDSPCLAEAEAIFGIYEDGVGDARD
Lmajor   319 ADSKK PIGGHIMAHASTTRLSLRKGRGEQRIIKVYDSPCLAEAEAIFGIYDDGVGDARD
Tvivax   352 ADPRK PIGGHIMAHASTTRLSLRKGRGEQRIIKVYDSPCLPEAEAIFGIYEDGVGDVRD
Tbrucei  315 ADSKK PIGGHIMAHASTTRLSLRKGRGEQRIIKVYDSPCLAESEAIFGIYENGVGDVRD
```

Figure 5.19 – Global multiple alignment of the *T. brucei* RAD51 polypeptide with a range of RAD51 orthologues. Multiple sequence alignment of the putative *T. brucei* RAD51 polypeptide with homologues of RAD51 from *T. vivax, T. cruzi* and *L. major.* Sequences were aligned using CLUSTAL W (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and shaded using the BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.html): residues that are identical in at least 50 % of the proteins are shaded in black and similarly conserved residues shaded in grey. Letters in red above the alignment represent the 8 key residues in *H. sapiens* RAD51 for interactions with the BRC repeats of BRCA2 (Yang *et al.*, 2002).

	T. brucei	*T. vivax*	*T. cruzi*	*L. major*
T. brucei	**100**	**71.7**	**79.6**	**77.5**
	100	78.8	85.0	83.6
T. vivax		**100**	**76.1**	**74.3**
		100	80.5	79.1
T. cruzi			**100**	**85.4**
			100	88.9
L. major				**100**
				100

Table 5.9 – Pairwise comparison of the putative *T. brucei* RAD51 polypeptide with a range of RAD51 homologues. The full length putative *T. brucei* RAD51 polypeptide was compared with homologues from *T.vivax, T. cruzi* and *L. major*. Pairwise alignments were performed using AlignX (Vector NTI) and the percentage identities and similarities calculated. The percentage identities are displayed in bold.

Despite the level of sequence identity being extremely high between the RAD51 polypeptides from *T. brucei* and *T. vivax*, we cannot conclude that the *T. vivax* BRC repeat is able to interact with the *T. brucei* RAD51. This could provide an explanation for why very few RAD51 foci were detected in the *T. vivax BRCA2-/-/+* cell line. However, this does not explain why very few RAD51 foci were detected in the *1BRC BRCA2-/-/+* cell line, where it might be argued that the BRC repeats have been selected for RAD51 interaction. Nevertheless, the single C terminal BRC repeat retained in this polypeptide is a degenerate copy, in that it is identical to all other BRC repeats, apart from the last 11 amino acids.

5.3.6 Analysis of VSG switching

Finally, analysis of VSG switching was performed in the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* cell lines to examine whether or not the DNA repair, recombination and RAD51 foci impairments are reflected in similar changes to antigenic variation. This assay was not performed using the same protocol described in section 4.3.8, as this would have required generating these mutants in the 3174.2 cell line. Instead, a modified version of the assay was used, in which only VSG switching frequency was analysed, using the Lister 427 cell lines.

Before analysing VSG switching rates, western analysis was performed in order to determine whether VSG221 continued to be expressed in each cell line, since the parasites may have undergone switching events whilst being cultured *in vitro*. Whole cell extracts were prepared, electrophoresed on a 10 % SDS-PAGE gel, blotted and probed for VSG221, which resides at the active VSG expression site. The resulting western blots of this analysis are displayed in figure 5.20 and indicate that all cell lines were expressing VSG221.
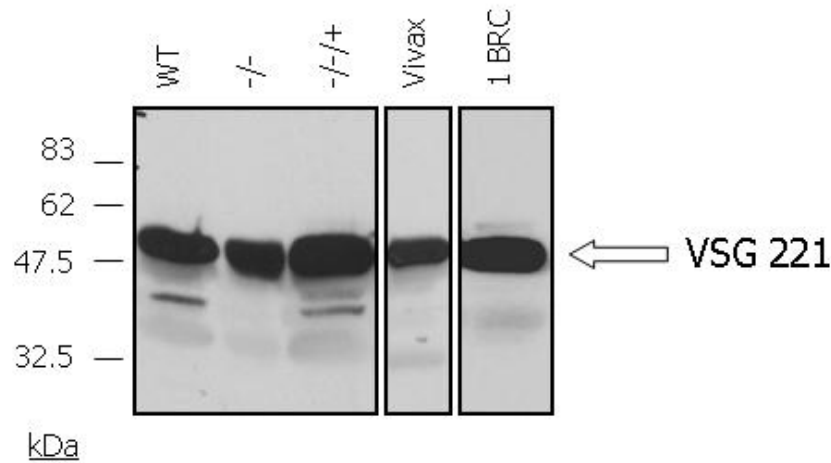
Figure 5.20 – Western blots of VSG221 in BRCA2 variants with reduced number of BRC repeats. The western blots display total protein extracts from wild type (WT), *brca2-/-2* (-/-), *BRCA2-/-/+* (-/-/+), *T. vivax BRCA2-/-/+* (Vivax) and *1 BRC BRCA2-/-/+* (1 BRC) cells probed with anti-VSG 221antiserum. Size markers are indicated.

Immunised mice were generated by infection with 2×10^5 cells of wild type, *brca2-/-2*, *BRCA2-/-/+*, *T. vivax BRCA2-/-/+* or *1BRC BRCA2-/-/+* cell lines. After parasitaemias reached $\sim 5 \times 10^7$ cells.ml^{-1}, the infections were cured by cymelarsan treatment, generating mice immune to all VSGs expressed in that cell line. Switched variants were then selected for by inoculating the immunised mice with $4\text{-}8 \times 10^7$ cells of each cell line (all mice were inoculated with the same cell line that they had previously been infected with). 24 hours after inoculation, the surviving *T. brucei* cells were recovered by exsanguination, and the number of switched variant clones in the blood was calculated as described in section 4.3.8.

The results of this assay are displayed in figure 5.21, and demonstrate that although there was greater variation in this assay, the switching frequency estimated for WT Lister 427 cells was comparable with that observed previously for WT 3174.2 cells: 9.5 switches per 10^{-7} cells compared to 8.13×10^{-7}, respectively. The Lister 427 *brca2-/-2* mutant also displayed an impaired ability to switch its VSG coat, though this was less pronounced in this assay, with a mean VSG switching frequency of 3.68×10^{-7} compared with 0.55×10^{-7} observed previously in 3174.2 *brca2-/-* cells. Surprisingly, both the *T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+* mutants were unaltered in their ability to switch their VSG coat relative to either the WT or *BRCA2-/-/+* cells, with mean VSG switching frequencies of 11.11×10^{-7} and 11.46×10^{-7} compared to 9.5×10^{-7} and 10×10^{-7}, respectively.

Statistical analysis suggested that only the *1BRC BRCA2-/-/+* cell line had a significantly higher VSG switching frequency than the *brca2-/-2* mutant (table 5.10). Most likely, the greater levels of variability in this assay are responsible for the lack of statistical significance being observed. Nevertheless, these results appear to indicate that the BRC

repeat expansion or organisation in *T. brucei* BRCA2 is not a critical determinant of VSG switching efficiency during an acute infection, despite the importance of this arrangement in DNA repair, recombination and in influencing RAD51 function.
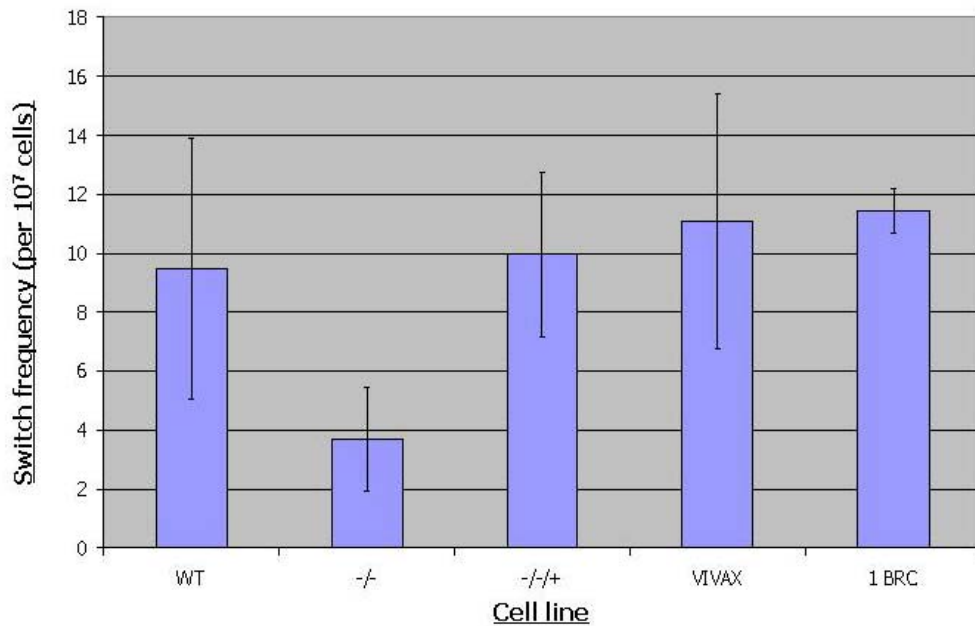
Figure 5.21 – VSG switching frequencies in BRCA2 variants with reduced number of BRC repeats. Values shown are the average switching frequencies for 427 wild type cells (WT), the *brca2* homozygous mutant 2 (-/-), the *BRCA2* re-expresser (-/-/+), the *T. vivax BRCA2* re-expresser (Vivax) and the 1BRC *BRCA2* re-expresser (1BRC) mutants. Data are from at least 3 experiments and standard error is indicated by bars.

	-/-	-/-/+	Vivax	1 BRC
WT	0.2182	0.2743	0.2909	0.0059
-/-		0.0675	0.3591	0.0178
-/-/+			0.8609	0.8825
Vivax				0.9254

Table 5.10 – Statistical analysis of the VSG switching frequencies in BRCA2 variants with reduced number of BRC repeats. P values are shown for two sample T-tests comparing VSG switching frequencies of 427 wild type cells (WT), the *brca2* homozygous mutant 2 (-/-), the *BRCA2* re-expresser (-/-/+), the *T. vivax BRCA2* re-expresser (Vivax) and the 1BRC *BRCA2* re-expresser (1BRC) mutants. Areas shaded in yellow indicate a significant difference ($P < 0.05$).

5.4 Generation of mutants with different truncations of BRCA2

Investigating the protein sequence of *T. brucei* BRCA2, and analysing shared regions of homology with BRCA2 orthologues has allowed a number of different putative functional motifs to be predicted (section 3.7). In addition to the BRC repeats, *T. brucei* BRCA2 has been predicted to possess all 5 conserved motifs in the DSS1-DNA binding domain (DBD): an alpha-helical domain, 3 oligosaccharide binding (OB) domains and a tower domain protruding from OB2. A C terminal RAD51 binding domain may also be present, but here homology is limited to a serine residue, which may be similar to a kinase regulatory sequence in the C terminal RAD51 binding domain in *H. sapiens* BRCA2 (Esashi *et al.*, 2005). Despite this, little is known about potential function, if any, of the other parts of the protein. Moreover, it is not known if the above domains provide further function.

In order to determine if the predicted domains of the protein are functional, perform the roles that have been predicted in *H. sapiens* and other eukaryotes, and to ask about other potential roles, it was decided to express further truncations and variants of the BRCA2 protein in *brca2-/-* mutants. Not only will this allow us to confirm the regions of the protein that allow for recombination and repair, but also to examine the replication-related phenotype observed previously (see above).

As for the 1BRC repeat BRCA2 variants, in all cases, constructs were generated by cloning DNA fragments into the plasmid pRM482 (R. McCulloch, gift), which had been *Eco*RV-digested and CIP treated, and were subsequently introduced into the second of the two independent *brca2-/-* homozygous mutants (*brca2-/-2*) in the Lister 427 cell line (expressers). In some cases, the constructs were also introduced into wild type Lister 427 cells, in order to generate cell lines with an additional copy of *BRCA2* (over-expressers). Again, these constructs allowed variations of BRCA2 to be inserted into the tubulin array, as was performed for *BRCA2-/-/+, T. vivax BRCA2* and *1 BRC BRCA2* (see sections 4.4 and 5.2).

5.4.1 Generation of expresser and over-expresser lines of the BRCA2 BRC repeat region

Previous research has demonstrated that six out of the eight BRC repeats from *H. sapiens* BRCA2 can directly interact with recombinant RAD51 when expressed *in vitro* (BRC1-4,

7-8) (Wong *et al.*, 1997;Chen *et al.*, 1998b;Marmorstein *et al.*, 1998). At high concentrations, peptides corresponding to BRC repeats 3 and 4 have been shown to bind RAD51 monomers and block RAD51-DNA filament formation. (Davies *et al.*, 2001). Indeed, when BRC3 was present in excess, it was found it could actually dissociate preformed RAD51 complexes. This inhibitory activity is proposed to occur by the BRC repeats binding to RAD51 and mimicking the recombinase's self association mode (Pellegrini *et al.*, 2002) and therefore preventing multimerisation. At lower concentrations, however, BRC3 and BRC4 can actually bind and form stable complexes with RAD51-DNA nucleoprotein filaments (Galkin *et al.*, 2005). A recent study purified the entire BRC repeat region of *H. sapiens* BRCA2 and, under conditions proposed to be typical of the nucleus, discovered that the BRC repeat region was capable of stimulating RAD51 strand exchange, suggesting that the BRC repeat region may be able to mediate homologous recombination in the absence of the C terminal region (Shivji *et al.*, 2006). It has not been made clear how this can be reconciled with *in vivo* studies, however, that have shown that cellular expression of single BRC repeats generated phenotypes that were highly reminiscent of *brca2-/-* mutants, indicating that the BRC repeats on their own interfere with RAD51 function (Chen *et al.*, 1998b;Stark *et al.*, 2004;Yuan *et al.*, 1999).

To address this in *T. brucei*, it was decided to attempt to express and over-express the isolated BRC repeat region of *T. brucei* BRCA2 (figure 5.22), in order to determine if the BRC repeats alone would be able to function in the absence of a DNA binding domain or if they would interfere with RAD51. A DNA fragment containing just the BRC repeat region of BRCA2 was generated by PCR-amplification from Lister 427 genomic DNA using Herculase DNA polymerase (Stratagene) and the primers *BRCA_TRUNC 5'* and *BRCA_TRUNC 3'*. The oligonucleotide *BRCA_TRUNC 5'* contained an *Nru*I restriction site, a methionine, an HA tag and 24 bases of sequence homologous to a region 17 amino acids upstream of the first BRC repeat. *BRCA_TRUNC 3'* consisted of an *Nru*I restriction site, a stop codon and 23 bases of sequence complementary to the bipartite NLS sequence, downstream of the most C terminal BRC repeat. The resulting PCR product was 1707 base pairs in length, the expected size containing all 12 BRC repeats and the bipartite NLS sequence. This PCR product was restriction digested with *Nru*I before ligating into the plasmid pRM482, resulting in the generation of the construct *pRM482::Trunc BRCA2*, which contained the antibiotic resistance cassette for G418.

Figure 5.22 – Cloning strategy used to generate the construct *pRM482::Trunc BRCA2*. The BRC repeat region of *BRCA2* was PCR-amplified using a 5' primer containing 24 bases of sequence homologous to a region 17 amino acids upstream of the most N terminal BRC repeat and a 3' primer containing 23 bases that was complementary to the bipartite NLS sequence downstream of the BRC repeats. Oligonucleotide primers are depicted by black triangles, red bars represent BRC repeats and the yellow bar represents the bipartite NLS sequence. The product was cloned into the construct pRM482 to allow the product to be re-expressed in *brca2-/-* mutants and over-expressed in wild type cells.

The construct was excised from the plasmid backbone by restriction digestion with *Xho*I and *Xba*I, before phenol:chloroform extraction and ethanol precipitation. Approximately 5 µg of digested DNA was transformed into the Lister 427 *brca2-/-2* mutant and the wild type Lister 427 cell line. Antibiotic resistant transformants were selected by plating out 4 x 10^7 cells from the transformation at 2.5 µg.ml^{-1} G418, over 48 wells with 1.5 mls per well. A number of antibiotic resistant transformants were recovered for each transformation and the introduction of the BRC repeat region of *BRCA2* confirmed by Southern analysis (section 5.4.3.1). One transformant was chosen from each transformation and named *Trunc BRCA2-/-/+* and *OE Trunc* for the expresser and over-expresser cell lines, respectively.

5.4.2 Generation of expresser and over-expresser lines of the BRCA2 C terminal domain

The COOH-terminal region of *H. sapiens* BRCA2 has been shown not only to contain a DSS1-DNA binding domain (DBD) (Yang *et al.*, 2002), but also to possess a RAD51 binding domain, (Sharan *et al.*, 1997;Mizuta *et al.*, 1997;Esashi *et al.*, 2005), which functions differently from the BRC repeats in that it binds RAD51 filaments, but not monomers like the BRC repeat region (Esashi *et al.*, 2007;Davies and Pellegrini, 2007).

It is unknown whether or not the C terminus of *T. brucei* BRCA2 contains a RAD51 binding domain, despite the existence of a potential CDK target site (section 3.7.3). It was decided to attempt to express and over-express the isolated C terminal domain of *T. brucei* BRCA2, in order to determine if this region can function in the absence of the BRC repeat region, and, indeed, to ask whether a RAD51 binding domain exists at the C terminus of *T. brucei* BRCA2.

The C terminal region of *T. brucei BRCA2* was PCR-amplified from Lister 427 genomic DNA using Herculase DNA polymerase (Stratagene) and the primers *BRCA_noBRC 5'* and *TbBRCA2 rev2* (figire 5.23). The oligonucleotide *BRCA_noBRC 5'* contained an *Nru*I restriction site, a methionine, an HA tag and 24 bases of sequence homologous to the region immediately downstream of the most C terminal BRC repeat. *BRCA_TRUNC 3'* consisted an *Nru*I restriction site and 24 bases complementary to the end of the *BRCA2* ORF. The resulting PCR product was a DNA fragment of 1746 base pairs, which included the bipartite NLS sequence. This PCR product was restriction digested with *Nru*I before ligating into the plasmid pRM482, resulting in the generation of the construct *pRM482::CtermBRCA2*, which contained the antibiotic resistance cassette for G418.

Figure 5.23 – Cloning strategy used to generate the construct *pRM482::Cterm BRCA2*. The C terminal region of *BRCA2* was PCR-amplified using a 5' primer containing 24 bases of sequence homologous to bipartite NLS sequence downstream of the BRC repeats and a 3' primer containing 24 bases that was complementary to the end of the ORF. Oligonucleotide primers are depicted by black triangles, red bars represent BRC repeats and the yellow bar represents the bipartite NLS sequence. The product was cloned into the construct pRM482 to allow the product to be re-expressed in *brca2-/-* mutants and over-expressed in wild type cells.

The construct was excised from the plasmid backbone by restriction digestion with *Xho*I and *Xba*I, before phenol:chloroform extraction and ethanol precipitation. Approximately 5 μg of digested DNA was transformed into the Lister 427 *brca2-/-2* mutant and the wild type Lister 427 cell line. Antibiotic resistant transformants were selected by plating out 4 x 10^7 cells from the transformation at 2.5 μg.ml^{-1} G418, over 48 wells with 1.5 mls per well. A number of antibiotic resistant transformants were recovered for the transformation in wild type cells, whilst only a single transformant was recovered from the transformation into the *brca2-/-2* mutant. The introduction of the C terminal region of *BRCA2* was confirmed by Southern analysis (section 5.4.3.1). One transformant was chosen from each transformation and named *C term BRCA2-/-/+* and *OE C term* for the expresser and over-expresser cell lines respectively.

5.4.3 Generation of expresser and over-expresser lines with the BRC repeat region of BRCA2 fused to the RPA70 subunit

Replication protein A (RPA) contains four OB folds (DBD-A, DBD-B, DBD-C and DBD-D), two of which bind to ssDNA with high affinity (DBD-A and DBD-B) (Bochkareva *et al.*, 2002). The DBD region of *H. sapiens* BRCA2 contains 3 oligonucleotide/oligosaccharide binding (OB) folds, of which OB2 and OB3 possess close

structural homology to DBD-A and DBD-B in replication protein A (RPA) 70, the largest subunit of RPA (Bochkarev *et al.*, 1999).

A recent study showed that the cellular expression of single or multiple *H. sapiens* BRC repeats, fused to the large RPA subunit could function in homology-directed repair, RAD51 binding and suppression of chromosomal abnormalities (Saeki *et al.*, 2006), implying that the DNA binding function of the C terminal region of BRCA2 could be provided by the ssDNA-binding protein RPA. A similar study carried out in *U. maydis* found comparable results, whereby the cells recovered from DNA damage sensitivity, but formed RAD51 foci and performed mitotic recombination more efficiently than WT cells (Kojic *et al.*, 2005).

It was decided to attempt a similar strategy in *T. brucei*, by translationally fusing the *T. brucei* BRC repeat region plus the downstream bipartite NLS sequence to the *T. brucei* Replication Protein A (RPA) 50 kDa subunit (Tb11.01.0870), which is homologous to the 70 kDa RPA protein in other eukaryotes. The *Leishmania* RPA 70 homologue has been shown to bind single-stranded DNA *via* its conserved OB fold domain (Neto *et al.*, 2007), providing confidence that the *T. brucei* homologue would do the same. This strategy should determine whether the C terminus of *T. brucei* BRCA2 simply functions as a DNA binding domain, or whether it has another, possibly undiscovered role.

An expression construct containing the BRC repeat region of *T. brucei* BRCA2 translationally fused to the large RPA subunit was generated by a few cloning steps (figure 5.24). The BRC repeat region of BRCA2 was PCR-amplified from Lister 427 genomic DNA using Herculase DNA polymerase (Stratagene) and the primers *BRCA_RPA 5'* and *BRCA_RPA 3'. BRCA_RPA 5'* consisted of an *Eco*RV restriction site, a methionine, an HA tag and 24 bases of sequence homologous to a region 17 amino acids upstream of the first BRC repeat. *BRC_RPA3'* consisted of an *Eco*RV restriction site and 23 bases of sequence that was complementary to a region immediately downstream of the most C terminal BRC repeat; this oligonucleotide was phosphorylated at the 5' end.

The gene encoding the large (50kDa) subunit of RPA was PCR-amplified from Lister 427 genomic DNA using the primers *RPA5'* and *RPA3'*. The oligonucleotide primer *RPA5'* consisted of 21 bases of sequence that was homologous to the start of the ORF, but excluding the start codon. This oligonucleotide was also phosphorylated at the 5' end. The oligonucleotide primer *BRCVAR3'3* consisted of an *Eco*RV restriction site and 20 bases of sequence that was complementary to the end of the ORF.

PCR-amplification of the BRC repeat region yielded a single DNA fragment of the expected 1707 base pairs that included the bipartite NLS sequence immediately down stream of the most C terminal repeat. PCR-amplification of the *RPA50* subunit yielded a single DNA fragment of the expected 1392 base pairs. PCR products were gel-extracted and ligated together using the phosphorylated ends of the PCR products, and the ligation reaction was subsequently TOPO cloned.

The DNA fragment of the successfully ligated PCR products was next excised from the TOPO vector by restriction digesting with *Eco*RV and was subsequently ligated into the plasmid pRM482 (R. McCulloch, gift), which had been *Eco*RV-digested and CIP treated. This resulted in the generation of the construct *pRM482::BRC+RPA*, which contained the antibiotic resistance cassette for G418.

Figure 5.24 – Cloning strategy used to generate the construct *pRM482:: BRC+RPA*. The BRC repeat region of *BRCA2* was PCR-amplified using a 5' primer containing 24 bases of sequence homologous to a region upstream of the first BRC repeat and a 3' primer containing 23 bases that was complementary to the bipartite NLS sequence downstream of the BRC repeats. The *RPA50* ORF was PCR-amplified using a 5' primer containing 24 bases of sequence homologous to the start of the ORF and a 3' primer containing bases of sequence that was complementary to the end of the ORF. The 2 PCR products were ligated together using phosphorylated primers. Oligonucleotide primers are depicted by black triangles, (P) indicates that the primer is 5' phosphorylated, red bars represent BRC repeats and the yellow bar represents the bipartite NLS sequence. Once the PCR products were ligated together, the product was cloned into the construct pRM482 to allow the product to be re-expressed in *brca2-/-* mutants and over-expressed in wild type cells.

The construct was excised from the plasmid backbone by restriction digestion with *Xho*I and *Xba*I, before phenol:chloroform extraction and ethanol precipitation. Approximately 5 μg of digested DNA was transformed into the Lister 427 *brca2-/-2* mutant and the wild type Lister 427 cell line. Antibiotic resistant transformants were selected by plating out 4 x 10^7 cells from the transformation at 2.5 μg.ml^{-1} G418, over 48 wells with 1.5 mls per well. A number of antibiotic resistant transformants were recovered for each transformation and the introduction of the BRC repeat region of BRCA2 translationally fused to the RPA50 subunit was confirmed by Southern analysis (section 5.4.3.1). One transformant was chosen from each transformation and named *BRC+RPA -/-/+* and *OE BRC+RPA* for the re-expresser and over-expresser cell lines respectively.

For reference, the expected polypeptides generated by each of the expressing and over-expressing transformants of various truncations of BRCA2 are depicted in figure 5.25.
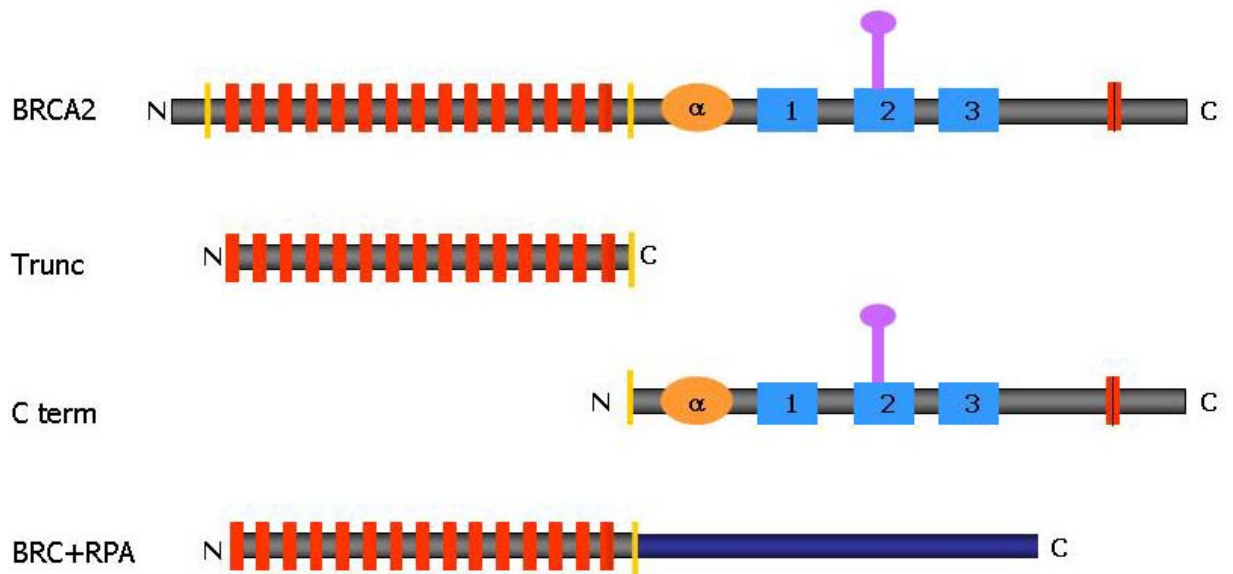
Figure 5.25 – Representation of the various truncated BRCA2 variants analysed. The BRC repeat region, the C terminal region and the BRC repeat regions translationally fused to *RPA50* were expressed by cloning into the construct pRM482. The full length *T. brucei* BRCA2 that was used to generate *BRCA2 -/-/+* mutants is shown for comparison. The figure represents the predicted, conserved domains of BRCA2: red bars – BRC repeats; orange oval – alpha helical domain; blue squares – OB domains; purple bar – tower domain; yellow bars – nuclear localisation signals; red bar with vertical line – putative CDK signal. The dark blue box in BRC+RPA represents the *T. brucei* RPA 50 kDa subunit.

5.4.3.1 Confirmation of BRCA2 variant expressers by Southern analysis

In order to confirm that *pRM482::Trunc, pRM482::C term* and *pRM482::BRC+RPA* had integrated into the tubulin array of the cell lines *Trunc BRCA2-/-/+*, *C term BRCA2-/-/+* and *BRC+RPA-/-/+* as expected, Southern analysis was carried out. Genomic DNA from wild type, *brca2-/- 2*, and the expresser cell lines were all digested with *Hin*dIII before being run out on a 0.8 % agarose gel and transferred to a nylon membrane by Southern blotting. The blots were probed with either a 378 bp region of the *T. brucei BRCA2* ORF, a 130 bp region of the BRC repeats, a 196 bp region of the C terminal region of *BRCA2* or a 456 bp region of the *RPA50* ORF (figure 5.26) and the results are shown in figure 5.27. Predicted fragment sizes of correctly integrated constructs are displayed in figure 5.26.
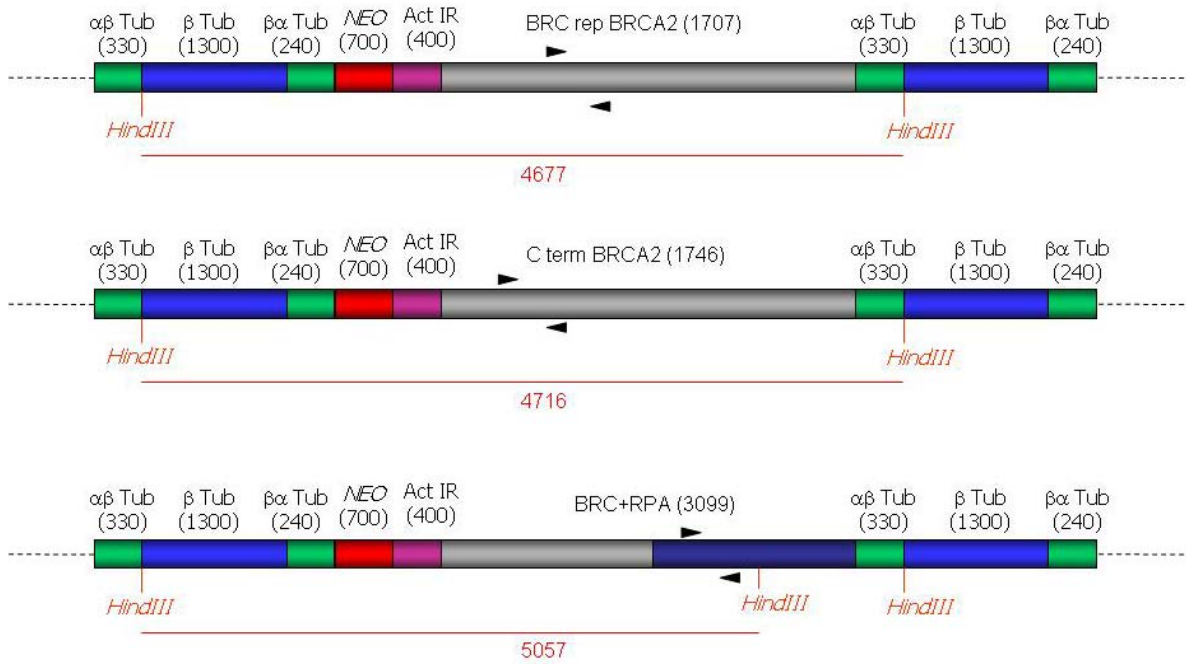
Figure 5.26 – Expressing BRCA2 with different truncations in the tubulin array. The constructs generated for expressing BRCA2 with different truncations were cloned into an *Eco*RV site between the actin intergenic (Act IR) and βα tubulin (βαTUB) intergenic sequences of the plasmid pRM482, which contains the antibiotic resistance cassettes for G418 (NEO). The constructs are flanked with tubulin intergenic regions (αβ Tub and βα Tub), which allow homologous integration into the tubulin array, replacing an α tubulin ORF. The sizes of IR and ORFs are indicated (in bp). This integration is displayed above with the restriction sites and predicted size fragments used to confirm the mutants by Southern analyses being shown. Primers used to generate DNA fragments for hybridisation are depicted by black triangles.
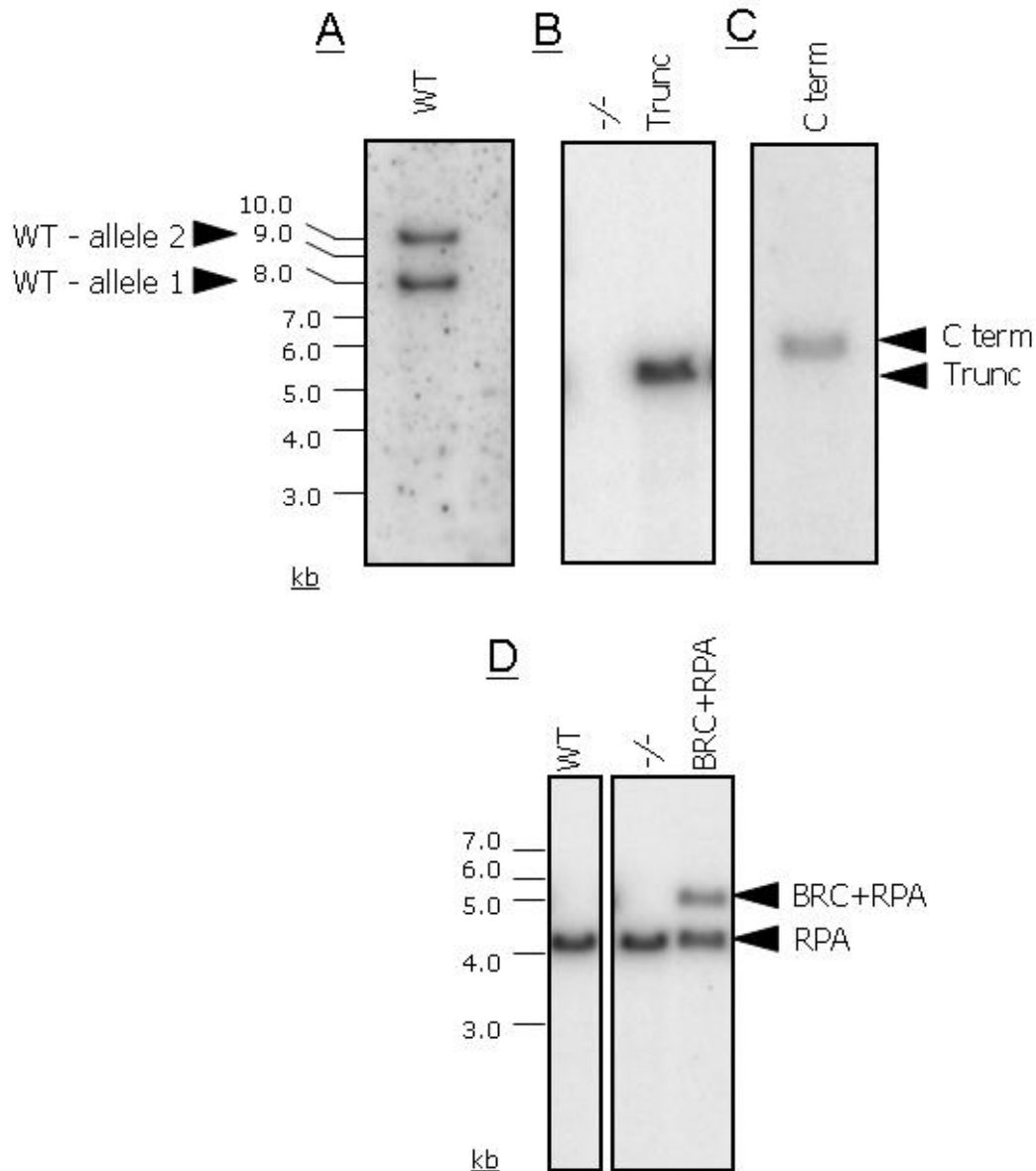
Figure 5.27 – Confirmation of the generation of BRCA2 variant expressers by Southern analysis. Cell lines were digested with *Hin*dIII by taking 5 µg of genomic DNA from each cell line and restriction digesting for 12 hours before running out on a 0.8 % agarose gel. The DNA was Southern blotted before being probed either (A) a 378 bp region of the *BRCA2* ORF, or (B) a 130 bp region of the BRC repeats, or (C) a 196 bp region of the C terminus of *BRCA2*, or (D) a 456 bp region of the large *RPA* subunit ORF . WT refers to genomic DNA from untransformed cell lines, homozygous mutant is indicated by -/-, expresser with just the BRC repeat region of *BRCA2* by Trunc, expresser with just the C terminal region of *BRCA2* by C term and the expresser with the BRC repeats translationally fused to RPA50 by BRC+RPA. The wild type copy of RPA50 is indicated by RPA.

Two allelic variants of *BRCA2* were observed for the wild type cell line, as observed previously (section 4.2.4). The absence of BRCA2 was again confirmed in the *brca2-/-2* mutant and the integration of the BRC repeat region and the C terminal region of *BRCA2* in the *brca2-/-2* mutant was confirmed by hybridising fragments of the expected sizes. The integration of the BRC repeats translationally fused to *RPA50* in the *brca2-/-2* mutant was confirmed by a hybridising fragment of the expected size, which was present in addition to the endogenous copy of *RPA50*.

5.4.3.2 Confirmation of BRCA2 variant expressers by RT-PCR

To support the results of the Southern analyses, RT-PCR was carried out on the expresser cell lines as described in section 4.2.4, using the primers *C term probe 5'* and *BRCA2 probe 3'* for the *C term-/-/+* cell line and the primers *BRC probe 5'* and *BRC probe 3'* for the *BRC+RPA -/-/+* and *Trunc BRCA2 -/-/+* cell lines. A ladder of products was seen in the *BRC+RPA -/-/+* and *Trunc BRCA2 -/-/+* cell lines, which correspond to the *T. brucei BRCA2* BRC repeats, whilst a *T. brucei BRCA2* C terminal-specific product was seen in the *C term-/-/+* cell line (figure 5.28).

This confirms that regions of *T. brucei BRCA2* mRNA are present in the expressers for the cell lines *BRC+RPA -/-/+, C term BRCA2-/-/+* and *Trunc BRCA2 -/-/+*. Relative amounts of cDNA could not be determined, as this analysis was non-quantitative.
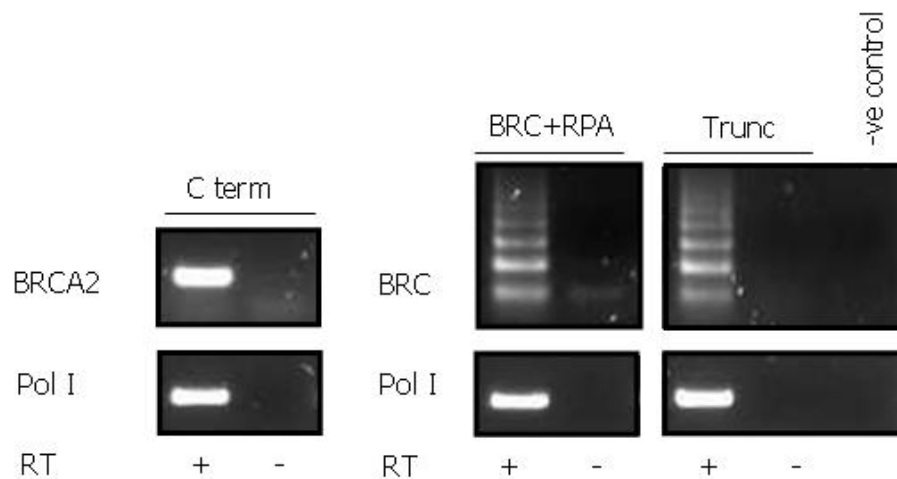
Figure 5.28 – Confirmation of the generation of expressers by RT-PCR. RT-PCR was carried out on cDNA generated from total RNA. RNA polymerase I-specific primers were used to show the generation of intact cDNA. Primers specific for the BRC repeat of BRCA2 were used to show the expression of genes in BRC+RPA and Trunc expressers, whilst primers specific for the C terminal region of BRCA2 were used to show the expression of genes in C term expressers. The negative control contains no cDNA. RT + and – indicates the presence or absence of reverse transcriptase.

5.4.3.3 Confirmation of BRCA2 variant expressers by Northern analysis

All of the constructs generated to re-express truncated versions of BRCA2 contained an N terminal HA tag. However, as with the *T. vivax BRCA2 -/-/+* and *1BRC BRCA2 -/-/+* cell lines, protein expression was undetectable with 2 different monoclonal anti-HA peroxidise conjugated antisera (Sigma, H6533 and Roche, 11667475001) for the cell lines *BRC+RPA -/-/+, C term BRCA2-/-/+* and *Trunc BRCA2 -/-/+*. Therefore, in order to determine if these truncations of *BRCA2* were being transcribed from the tubulin array as expected, northern blots were performed. Total RNA was extracted from *BRCA2-/-/+, Trunc*

BRCA2 -/-/+, C term BRCA2-/-/+ and *BRC+RPA -/-/+* cell lines (as in section 5.2.3.3) before 20 μg samples were separated by electrophoresis on a denaturing formaldehyde gel. The RNA was transferred to a nylon membrane by capillary blotting and blots probed with a 196 bp fragment of the C terminal region of the *BRCA2* ORF for *BRCA2-/-/+* and *C term BRCA2 -/-/+* cell lines. The *BRC+RPA-/-/+* and *Trunc BRCA2-/-/+* cell lines were probed with a 130 bp fragment from the BRC repeat of the *BRCA2* ORF. The hybridising bands generated in each lane were assumed to be mature mRNA, based on their size, and are shown in figure 5.29.
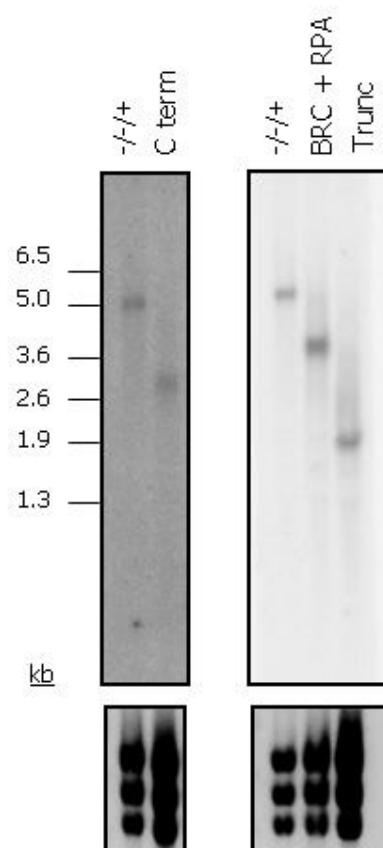
Figure 5.29 – Northern analysis of expresser mutants. Northern blots of 20 μg of total RNA from *BRCA2-/-/+* and *C term BRCA2-/-/+* probed with a 196 bp fragment of the C terminal region of the *BRCA2* ORF (left blot) and *BRC+RPA -/-/+* and *Trunc BRCA2-/-/+* probed with a 130 bp fragment of the BRC repeat region of the *BRCA2* ORF (right blot). Size markers are shown and ethidium stained gels are displayed below the Northern blots to demonstrate amount of RNA loaded.

The northern blot demonstrates that regions of *T. brucei BRCA2* mRNA were detectable for *BRCA2-/-/+, C term BRCA2-/-/+, BRC+RPA-/-/+* and *Trunc BRCA2-/-/+* cell lines.

This analysis, taken together with the results from the Southern analyses and RT-PCR, demonstrate that the *C term BRCA2-/-/+, BRC+RPA-/-/+* and *Trunc BRCA2-/-/+* cell lines had integrated the constructs as predicted, and the expected mRNA was generated.

5.4.3.4 Confirmation of over-expresser BRCA2 variants by Southern analysis

In order to confirm that *pRM482::Trunc, pRM482::C term* and *pRM482::BRC+RPA* had integrated into the tubulin array of the cell lines *OE Trunc, OE C term* and *OE BRC+RPA* as expected, Southern analysis was carried out in the same manner as was described in section 5.4.3.1.
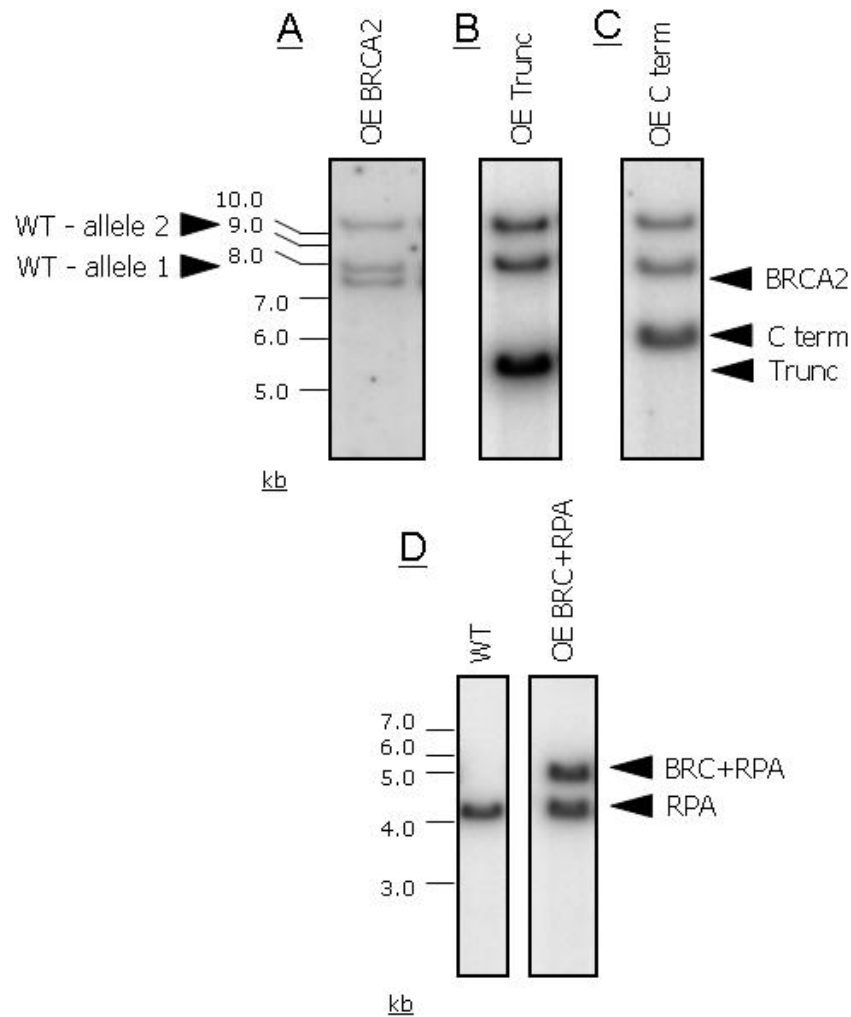
Figure 5.30 – Confirmation of the generation of over-expressers by Southern analysis. Cell lines were digested with *Hin*dIII by taking 5 µg of genomic DNA from each cell line and restriction digesting for 12 hours before running out on a 0.8 % agarose gel. The DNA was Southern blotted before being probed either (A) a 378 bp region of the *BRCA2* ORF, (B) a 130 bp region of the BRC repeats, (C) a 196 bp region of the C terminus of *BRCA2*, or (D) a 456 bp region of the large *RPA* subunit ORF . OE BRCA2 refers to the full length BRCA2 over-expressed in Lister 427 cell lines, OE Trunc to the over-expression of the BRC repeat region, OE C term to the over-expression of the C terminal region and OE BRC+RPA to the over-expression of the BRC repeats translationally fused to RPA.

Three hybridising fragments were seen for the *OE BRCA2* cell line. Two of these fragments represent the two allelic variants of *BRCA2*, as was observed previously (section 4.4.1), whilst the third, 7.5 kb band represents the integration of an extra copy of *BRCA2* in the tubulin array. The *OE Trunc* and *OE C term* cell lines were also seen to possess 3 hybridising fragments. In each case, the smallest bands represent the sizes expected

following integration of the BRC repeat region and the C terminal region of BRCA2, respectively, into the tubulin array. The *OE BRC+RPA* cell line was seen to possess two hybridising fragments, one of these represents the wild type copy of *RPA50*, whilst the second represents the integration of the BRC repeats translationally fused to *RPA50*.

5.5 Phenotypic analysis

The *brca2-/-* mutants expressing truncated versions of BRCA2 were analysed for their *in vitro* population doubling times, cell cycle progression, DNA damage sensitivity, recombination efficiency, the ability to form RAD51 foci and VSG switching frequency. The cell lines in which these truncations of BRCA2 had been over-expressed were only analysed for their *in vitro* population doubling times, cell cycle progression, DNA damage sensitivity and recombination efficiency.

5.5.1 Analysis of in vitro growth

In vitro growth rates of the cell lines were compared in order to determine if the absence of specific regions of the protein or presence of an additional copy of protein variants affected population doubling times. The assay was carried out following the same protocol as described in section 4.3.1. Four repetitions of the growth assay were carried out for each cell line and the results are displayed in figure 5.31, in comparison with the values determined previously for wild type, *brca2-/-2*, *BRCA2-/-/+* and *OE BRCA2* cell lines.
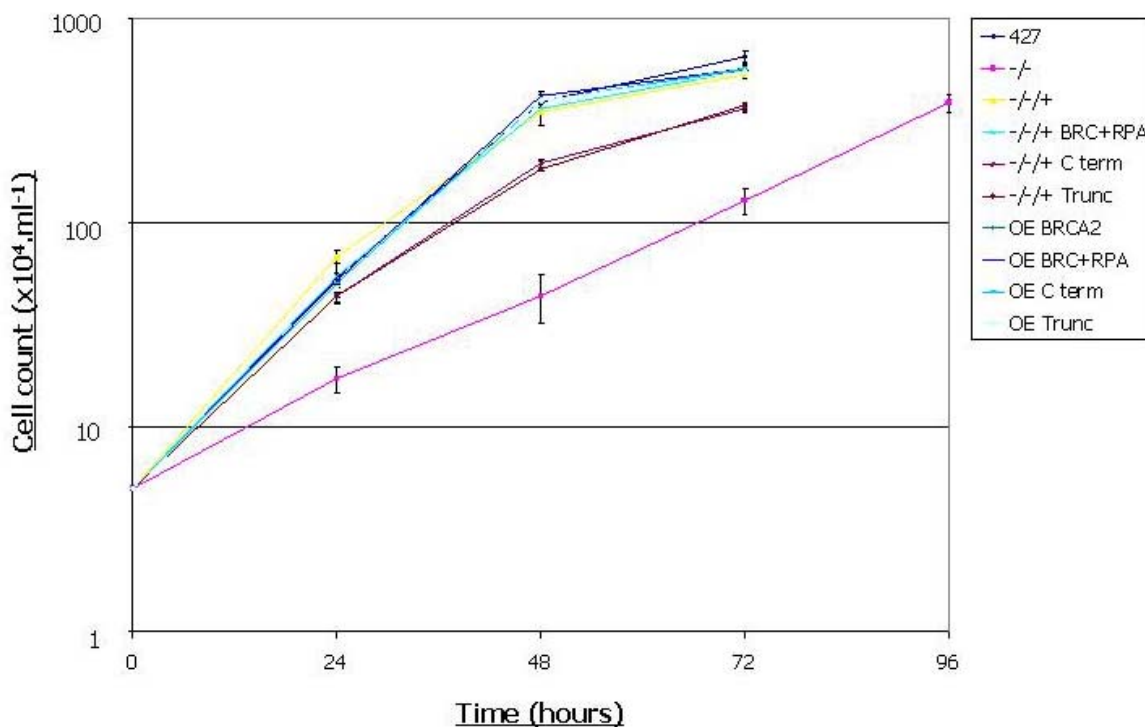
Figure 5.31 – Analysis of *in vitro* growth of BRCA2 variant expressers and over-expressers.
5 ml cultures were set up at 5×10^4 cells.ml^{-1} and cell densities counted 24, 48, 72 and 96 hours
subsequently. Standard errors are indicated for the counts using data from four repetitions. 427:
wild type; -/-: homozygote (*brca2-/-2*); -/-/+: full length BRCA2 expresser; -/-/+ BRC+RPA:
BRC+RPA expresser; -/-/+ C term: C term expresser; -/-/+ Trunc: Trunc expresser; OE BRCA2:
BRCA2 over-expresser; OE BRC+RPA: BRC+RPA over-expresser; OE C term: C term over-
expresser; OE Trunc: Trunc over-expresser.

Cell line	427	-/-	-/-/+	BRC+RPA	C term	Trunc
Doubling time	8.19 +/- 0.4	15.50 +/- 0.34	8.26 +/- 0.4	7.85 +/- 0.1	9.16 +/- 0.13	9.31 +/- 0.07

Table 5.11 – *in vitro* population doubling times of BRCA2 variant expressers. The mean
doubling time for each of the re-expresser mutants with reduced numbers of BRC repeats are
displayed in hours and compared to the population doubling times for WT, *brca2-/-2* and *BRCA2-/-
/+* cell lines. 427: wild type; -/-: homozygote (*brca2-/-2*); -/-/+: full length BRCA2 expresser;
BRC+RPA: BRC+RPA expresser; C term: C term expresser; Trunc: Trunc expresser. Standard
errors are indicated.

From the growth curves shown in figure 5.31 and the population doubling times shown in

table 5.11, it was apparent that expressing the BRC+RPA fusion protein rescued the

impaired growth phenotype observed in the *brca2-/-2* mutant, as the population doubling

time increased from 15.5 to 7.85 hours, which was comparative to the population doubling

times of the full length *BRCA2* re-expresser and also wild type cells. This result was

confirmed by the statistical analysis shown in table 5.12, which displayed that there was no

statistical difference between the *BRC+RPA-/-/+* mutant and the *BRCA2-/-/+* or wild type

cells, with P values of 0.1717 and 0.4980 respectively (p>0.05).

Expressing either the BRC repeat region (*Trunc BRCA2-/-/+*) or the C terminal region (*C term BRCA2-/-/+*) of *BRCA2*, appeared to rescue the impaired growth phenotype observed in the *brca2-/-2* mutant to a certain degree, with population doubling times increasing from 15.5 to 9.31 and 9.16 hours, respectively. However, these population doubling times were not comparable with either the *BRC+RPA-/-/+* or the *BRCA2-/-/+* cell lines, with statistically significant differences being displayed (p<0.05). This therefore indicated that expressing just these isolated regions of BRCA2 was not sufficient to provide cell functions of the full length protein.

	-/-	-/-/+	BRC+RPA	C term	Trunc
WT	0.0002	0.3820	0.4980	0.1075	0.0792
-/-		0.0001	0.0003	0.0001	0.0003
-/-/+			0.1717	0.0589	0.0491
BRC+RPA				0.0083	0.0017
C term					0.1254

Table 5.12 – Statistical analysis of the population doubling times of BRCA2 variant expressers. The P values are shown for two sample T-tests comparing population doubling times of wild type cells, *brca2* homozygous mutant 2 (-/-), *BRCA2* expresser (-/-/+), *BRC+RPA* expresser (BRC+RPA), *C term BRCA2* expresser and *Trunc BRCA2* expresser (Trunc). Areas shaded in yellow indicate a significant difference.

The over-expression of any of these truncated versions of *BRCA2* appeared to have no effect on the *T. brucei* growth rates. Indeed, the population doubling times were essentially equivalent to those seen when full length *BRCA2* was over-expressed, with population doubling times of 7.45 for *OE BRC+RPA*, 7.54 for *OE C term* and 7.71 for *OE Trunc* compared to 7.58 for *OE BRCA2*. This appears to indicate that the putative presence of excess, truncated BRCA2 protein does not impede growth.

5.5.2 Analysis of the cell cycle

Cell cycle progression was next examined in order to determine if the absence of specific regions of the protein or presence of an additional copy of protein variants affected the distribution of cell types. The assay was carried out following the same protocol as described in section 4.3.3 and the results are displayed in figure 5.32, in comparison with the values determined previously for wild type, *brca2-/-2*, *BRCA2-/-/+* and *OE BRCA2* cell lines.
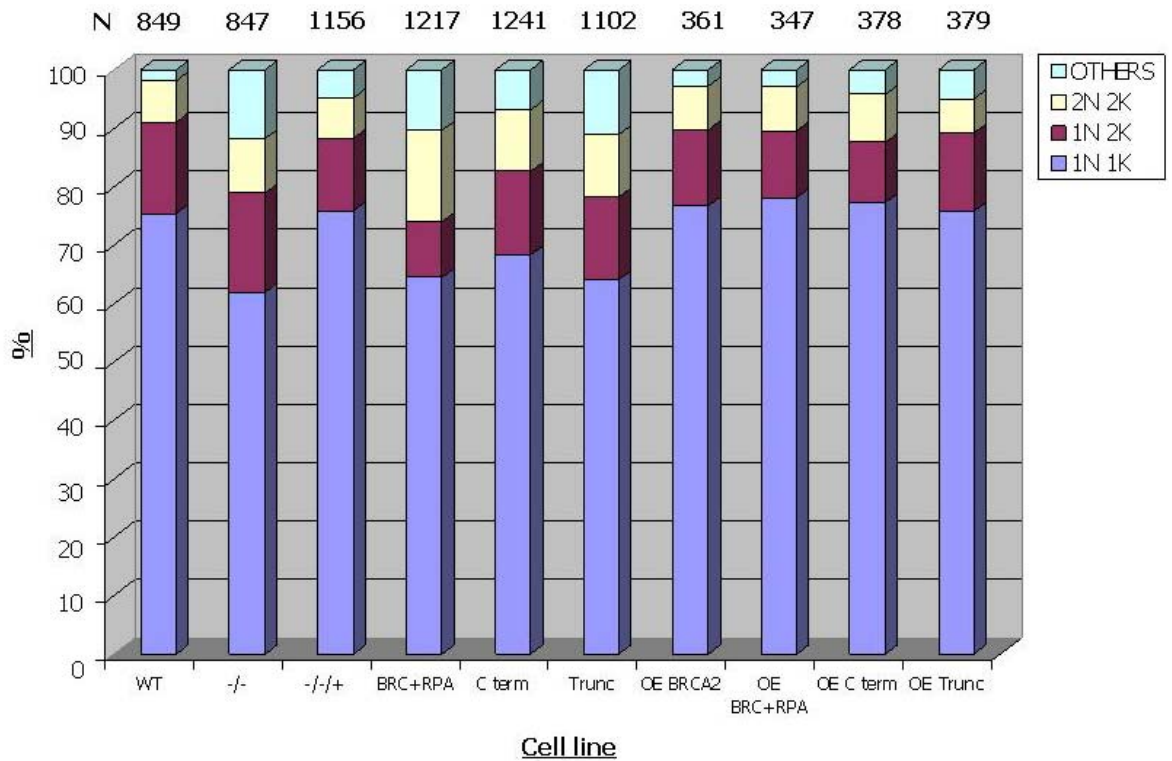
Figure 5.32 – DAPI analysis of BRCA2 variant expressers and over-expressers. The DNA content of *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term), *Trunc BRCA2-/-/+* (Trunc), *OE BRC+RPA*, *OE C term* and *OE Trunc* were visualised by DAPI and compared with the DNA content of wild type Lister 427, *brca2* homozygous mutant 2 (-/-),*BRCA2-/-/+* (-/-/+) cell lines and *BRCA2* over-expresser (OE BRCA2) cell lines. The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications cells (others) were counted and represented by their mean count as a percentage of the total cells counted. N = number of cells counted.

Comparison of the overall distribution of cell types by Chi-squared analysis demonstrated that the expression of each of the truncations of BRCA2 did not allow cell cycle progression to occur normally (table 5.13). Only when the full length *BRCA2* was re-expressed, was the distribution of cells comparative to that of wild type cells. Nevertheless, the relative ratios of different cell cycle stages appear to vary for the different expressers.

The accumulation of aberrant cells observed in the *brca2-/-2* mutant was rescued to a certain degree when the C terminus of *BRCA2* was expressed, since the number of 1N1K cells rose from 61.9 % to 68.3 % and the number of aberrant cell types reduced from 11.8 % to 6.8 %. However, the number of aberrant cells was still less than in the WT or *BRCA2-/-/+* cell lines, suggesting that the C-terminus cannot fully perform this function.

When the isolated BRC repeat region of *BRCA2* was expressed, the distribution of cells appeared comparatively similar to those previously observed for the *brca2-/-2* mutant, with 64.1 % of 1N1K cells, 14.2 % of 1N2K cells, 10.7 % of 2N2K cells and 11.1 % of aberrant cell types, compared to 61.9 %, 17 %, 9.3 % and 11.8 %, respectively. This is reflected in

the Chi squared analysis which shows that the distribution of cell types observed in the *Trunc BRCA2-/-/+* cell line were not significantly different from the *brca2-/-2* mutant, but were significantly different from the *BRCA2-/-/+* and wild type cell lines. This result suggests that the BRC repeat region of BRCA2 does not function to complement the cell cycle progression of *brca2-/-* mutants.

The distribution of cells observed for the *BRC+RPA-/-/+* mutant was found to be significantly different from WT, *BRCA2-/-/+* and also *brca2-/-* cell types, with Chi squared values ranging from 7.9 to 52.7 (at P = 0.0477 and 0.0001). Indeed, the distribution of cells appeared somewhat different to any of the other cell lines, with an apparent increase in the number of cells with 2N 2K content (15.7 %, compared with 5.8-10.7 % in the other cell lines). This may indicate a cell cycle stall during mitosis, the reasons for which are as yet unknown. It is also notable that the number of aberrant cells in *BRC+RPA-/-/+* are indistinguishable from the *brca2-/-* mutants, indicating that these cells, despite having a WT population doubling time, retain the cell division abnormality seen in the absence of BRCA2.

The over-expression of any of these truncated versions of *BRCA2* appeared to have no apparent effect on the distribution of *T. brucei* cells. Indeed, the cell cycle phenotypes appeared essentially equivalent to when the full length *BRCA2* was over-expressed.

	-/-	-/-/+	BRC+RPA	C term	Trunc
WT	**55.710** 0.0001	**5.438** 0.1424	**52.694** 0.0001	**15.649** 0.0013	**48.643** 0.0001
-/-		**9.342** 0.0251	**7.918** 0.0477	**3.322** 0.3447	**0.803** 0.8487
-/-/+			**19.773** 0.0002	**3.923** 0.2699	**12.297** 0.0064
BRC+RPA				**5.522** 0.1373	**3.888** 0.2739
C term					**2.861** 0.4135

Table 5.13 – Statistical analysis of the cell cycle data for BRCA2 variant expressers. Chi squared analysis of the cell cycle data for wild type cells, *brca2* homozygous mutant 2 (-/-), *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term) and *Trunc BRCA2-/-/+* (Trunc). The numbers indicated in bold represent the Chi squared value, whilst the numbers below represent the P value at which it was calculated. Areas shaded in yellow indicate a significant difference.

In order to examine if the distribution of cells was affected by the presence of DNA damage, the DNA content of the above expresser and over-expresser cell lines were analysed after phleomycin treatment, as described in section 4.3.3.

The results of this analysis are displayed in figure 5.33. These data demonstrate that when the BRC-RPA fusion was expressed in the *brca2-/-* cells, the cell cycle distribution was broadly comparable with that seen in full length *BRCA2-/-/+* cells and significantly different from that observed for *brca2-/-* mutants in the presence of DNA damage. The number of 1N 1K cells increased from 43.8 % in *brca2-/-2* (at 0.25 μg.ml^{-1} of phleomycin) to 72.8 % in the *BRC+RPA-/-/+* cell line (at 0.25 μg.ml^{-1} of phleomycin), whilst the number of 1N 2K cells were seen to reduce from 27.7 % to 12.2 %, indicating enhanced progression through G2 phase. The number of aberrant cell types were also seen to decrease, from 20.3 % in *brca2-/-2* to 10.2 % in the *BRC+RPA-/-/+* cell line, though the *BRC+RPA-/-/+* cell line appeared to yield greater numbers of such cells than the *BRCA2-/-/+* cells after damage. The statistical analysis shown in table 5.14 confirms these results, showing a significant difference between the *brca2-/-2* mutant and the *BRC+RPA-/-/+* cell line, with Chi squared values of 32.3 to 37.2 (at P = 0.0001). Moreover, no such significant difference was observed between the *BRC+RPA-/-/+* cell line and WT and *BRCA2-/-/+* cell lines, at 0.25 μg.ml^{-1} of phleomycin, though at 1.0 μg.ml^{-1} a difference was observed. Most likely, the BRC-RPA fusion is capable of allowing DNA damage to be repaired more effectively than in *brca2-/-* mutants, most probably due to an increased resistance to DNA damaging agents (see section 5.5.3).

Expression of just the BRC repeat or C terminal regions of BRCA2 allowed the cell cycle to progress somewhat more normally than the *brca2-/-* mutants in the presence of DNA damage, since significant differences were observed between the *C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines and the *brca2-/-2* mutant, with Chi squared values ranging from 17.36 to 23.6 (at P = 0.0006 and 0.0001). However, this appeared not to be as pronounced as for the *BRC+RPA-/-/+* cell line. At either phleomycin concentration, the numbers of 1N1K cells was lower than WT, *BRCA2-/-/+* and *BRC+RPA-/-/+*, indicating deficiencies in other cell cycle stages. This appeared primarily to be due to greater numbers of aberrant cells accumulating, indicating damage, but was not manifest as increased 1N2K cells that would represent a block in G2 progression (as seen in *brca2-/-* cells). Together, these data most likely indicate that DNA damage is repaired more effectively than in *brca2-/-* mutants, but less effectively than either WT, *BRCA2-/-/+* or *BRC+RPA-/-/+* cell lines, probably due to an increased sensitivity to DNA damaging agents (see section 5.5.3).

Unusually, when the phleomycin treatment of the *C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines was increased from 0.25 to 1.0 μg.ml^{-1} of phleomycin, the cell cycle

phenotypes appeared more similar to wild type and *BRCA2-/-/+* cell lines. The reason for this phenomenon is unknown, but could possibly be due to the increase in DNA damage causing a higher proportion of cell death, causing fewer aberrant cell types to be counted. In support of this, the same phenomenon appeared to be observed in the *brca2-/-* cells. In contrast, in the *BRC+RPA-/-/+* cells the numbers of aberrant cells increased with greater damage. If this latter cell line is more resistant to DNA damage than both the *C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* mutants (see section 5.5.3), the proportion of dead cells would be expected to be lower at equivalent concentrations.
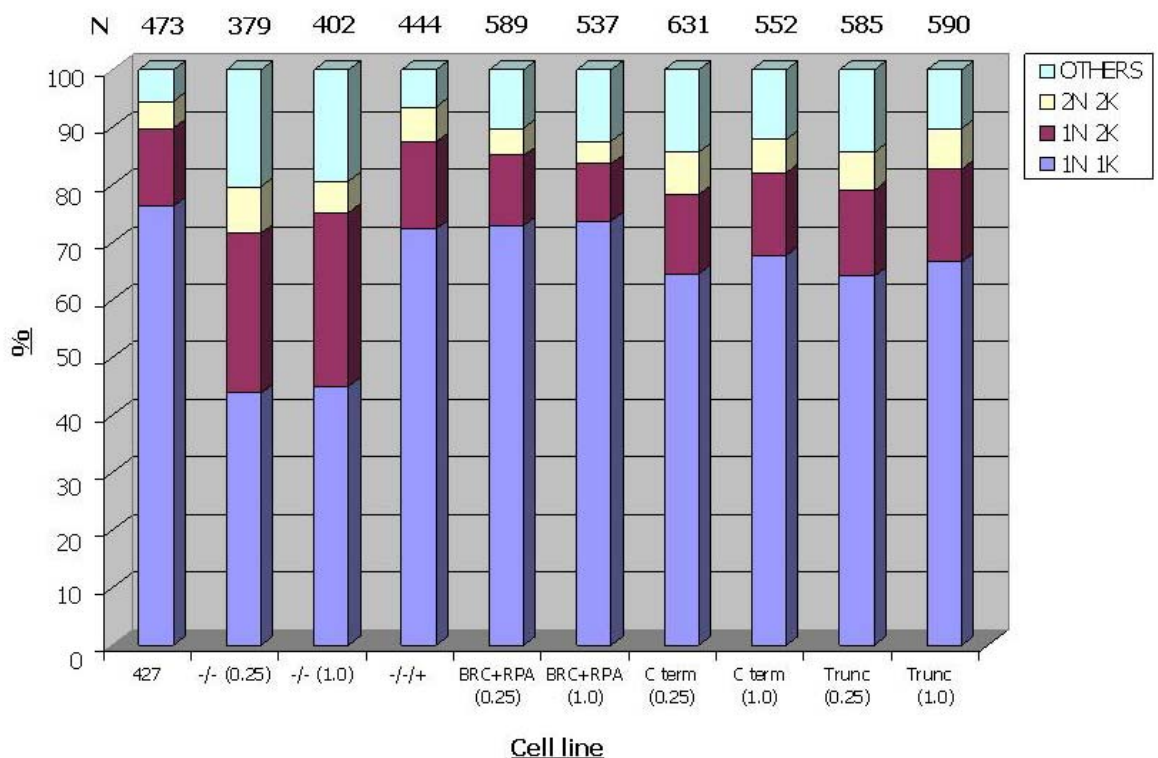
Figure 5.33 – DAPI analysis BRCA2 variant expressers after DNA damage. The DNA content of *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term), *Trunc BRCA2-/-/+* (Trunc), *OE BRC+RPA*, *OE C term* and *OE Trunc* were visualised by DAPI and compared with the DNA content of wild type Lister 427, *brca2* homozygous mutant 2 (-/-), *BRCA2-/-/+* (-/-/+) cell lines and *BRCA2* over-expresser (OE BRCA2) cell lines. The numbers of cells with 1 nucleus and 1 kinetoplast (1N 1K); 1 nucleus and 2 kinetoplasts (1N 2K); 2 nuclei and 2 kinetoplasts (2N 2K); and cells that do not fit into the expected classifications cells (others) were counted and represented by their mean count as a percentage of the total cells counted. Wild type and *BRCA2 -/-/+* cell lines were grown in media with 1.0 μg.ml^{-1} of phleomycin, whilst *brca2-/-*, *1BRC BRCA2-/-/+* and *T. vivax BRCA2-/-/+* were grown in media with 0.25 μg.ml^{-1} and 1.0 μg.ml^{-1} of phleomycin. N = number of cells counted.

The generation of DNA damage through addition of phleomycin had no apparent effect on the ability of the over-expresser cell lines to progress through the cell cycle (data not shown). This result would fit with the previous results that when truncated versions of BRCA2 were over-expressed, the *T. brucei* cells were unaffected in their ability to

progress through the cell cycle and in their population doubling times. Indeed, this is most likely due to a lack of increased or decreased sensitivity to DNA damaging agents compared to wild type cells (see section 5.5.3).

	-/- (0.25)	-/- (1.0)	-/-/+	BRC+ RPA (0.25)	BRC+ RPA (1.0)	C term (0.25)	C term (1.0)	Trunc (0.25)	Trunc (1.0)
WT	**71.310** 0.0001	**68.395** 0.0001	**1.024** 0.7954	**4.287** 0.2321	**10.053** 0.0181	**17.801** 0.0005	**9.308** 0.0255	**17.239** 0.0006	**7.001** 0.0719
-/- (0.25)		**1.182** 0.7573	**34.183** 0.0001	**34.393** 0.0001	**37.197** 0.0001	**18.578** 0.0003	**23.554** 0.0001	**17.378** 0.0006	**21.944** 0.0001
-/- (1.0)			**32.606** 0.0001	**32.263** 0.0001	**34.814** 0.0001	**19.140** 0.0003	**22.483** 0.0001	**17.355** 0.0006	**21.743** 0.0001
-/-/+				**2.892** 0.4085	**8.030** 0.0454	**10.728** 0.0133	**5.159** 0.1605	**10.379** 0.0156	**2.722** 0.4364
BRC+RPA (0.25)					**1.126** 0.7707	**4.271** 0.2337	**1.318** 0.2748	**4.106** 0.2503	**2.947** 0.3998
BRC+RPA (1.0)						**5.938** 0.1147	**3.407** 0.3331	**6.025** 0.1104	**7.685** 0.053
C term (0.25)							**0.848** 0.8379	**0.140** 0.9866	**1.713** 0.6341
C term (1.0)								**0.755** 0.8602	**0.863** 0.8343
Trunc (0.25)									**1.543** 0.6932

Table 5.14 – Statistical analysis of the cell cycle data for BRCA2 variant expressers after DNA damage. Chi squared analysis of the cell cycle data for wild type cells, *brca2* homozygous mutant 2 (-/-2), *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term) and *Trunc BRCA2-/-/+* (Trunc). The numbers indicated in bold represent the Chi squared value, whilst the numbers below represent the P value at which it was calculated. Areas shaded in yellow indicate a significant difference.

5.5.3 Analysis of DNA damage sensitivity

Sensitivity of the BRCA2 variant cell lines to DNA damaging agents was next examined by Alamar Blue assays that were carried out following the same protocols as described in section 4.3.4, using both MMS and phleomycin as DNA damaging agents. These results are displayed in figures 5.34 and 5.35, in comparison with the values determined previously for wild type, *brca2-/-2*, *BRCA2-/-/+* and the *OE BRCA2* cell lines.

When MMS was used as the DNA damaging agent, the *BRC+RPA-/-/+* cell line was essentially equivalent to the *BRCA2-/-/+* cells in terms of sensitivity, with a mean IC50 of 0.0030 % compared to 0.0031 %. This result was confirmed by the statistical analysis shown in table 5.14, which displayed that the *BRC+RPA -/-/+* mutant was not significantly different from the *BRCA2-/-/+* mutant (p>0.05), but was significantly different from the WT and *brca2-/-2* cell lines (p<0.05). This therefore demonstrates that the BRC repeat region of BRCA2 is capable of repairing damage induced by MMS when it is fused with a

distinct DNA binding domain, consistent with findings in mammalian cells and *U. maydis* (Saeki *et al.*, 2006;Kojic *et al.*, 2005;Kojic *et al.*, 2006).
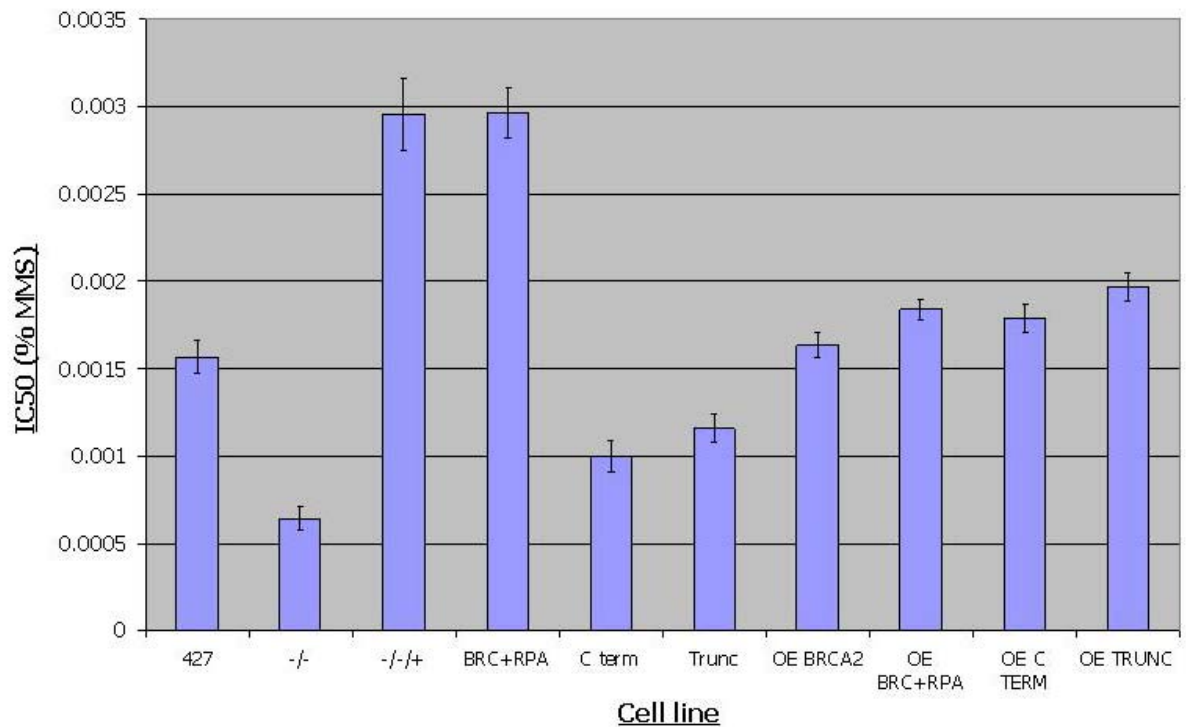
Figure 5.34 – IC50s of *T. brucei* BRCA2 variant expressers exposed to MMS. *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term), *Trunc BRCA2-/-/+* (Trunc), *OE BRC+RPA*, *OE C term* and *OE Trunc* cell lines were placed in serially decreasing amounts of MMS and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments and are compared to the previous results from wild type (427), *brca2-/-2* (-/-), *BRCA2-/-/+* (-/-/+) and *OE BRCA2* cell lines; bars indicate standard error.

Both the *C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines were more sensitive to MMS than either the *BRCA2-/-/+* or WT cell lines, but more resistant than the *brca2-/-2* mutants, findings confirmed by statistical analysis. This result is indicative that the isolated BRC repeat region or the C terminal region of BRCA2 cannot function alone to efficiently repair DNA damage induced by MMS *in vivo*, though may perhaps allow DNA damage repair to occur, albeit at a low level.

When the various truncations of *BRCA2* were over-expressed, no significant increase or decrease in sensitivity to MMS was observed, when compared to the *OE BRCA2* cell line, or indeed WT cells.

	-/-	-/-/+	BRC+RPA	C term	Trunc
WT	0.0025	0.0192	0.0021	0.0183	0.0759
-/-		0.0024	0.0007	0.0211	0.0070
-/-/+			0.7664	0.0048	0.0043
BRC+RPA				0.0002	0.0006
C term					0.0172

Table 5.15 – Statistical analysis of the Alamar Blue results for MMS. P values are shown for two sample T-tests comparing the IC50s for MMS sensitivity of wild type cells (WT), *brca2* homozygous mutant 2 (-/-), *BRCA2-/-/+* (-/-/+), *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term), and *Trunc BRCA2-/-/+* (Trunc) cell lines. Areas shaded in yellow indicate a significant difference.

When phleomycin was used as the DNA damaging agent, all of the BRCA2 variant expressers were seen to be more sensitive to this treatment than the full length BRCA2 re-expresser. This result was confirmed by the statistical analysis shown in table 5.17.
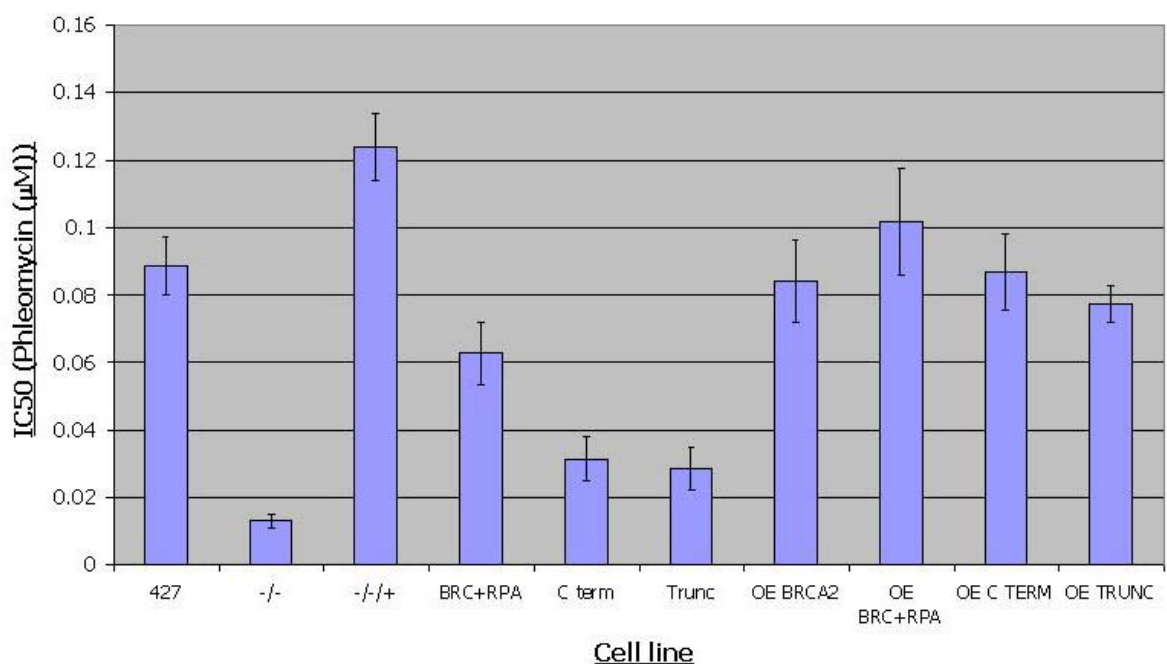
Figure 5.35 – IC50s of *T. brucei* BRCA2 variant expressers exposed to phleomycin. *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term), *Trunc BRCA2-/-/+* (Trunc), *OE BRC+RPA*, *OE C term* and *OE Trunc* cell lines were placed in serially decreasing amounts of phleomycin and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments and are compared to the previous results from wild type (427), *brca2-/-2* (-/-), *BRCA2-/-/+* (-/-/+) and *OE BRCA2* cell lines; bars indicate standard error.

The *BRC+RPA-/-/+* cell line remained more resistant to phleomycin damage than either the *C term BRCA2-/-/+*, *Trunc BRCA2-/-/+* or *brca2-/-2* cell lines, with a mean IC50 of 0.063 μM, compared to 0.032 μM , 0.029 μM and 0.013 μM, respectively. Although this sensitivity was greater than that observed for WT cells (0.099 μM), it was highly reminiscent of the putative haploinsufficiency that was observed for the *BRCA2+/-* mutants (section 4.3.4), where mean IC50s of 0.060 μM and 0.067 μM were seen. This therefore appears to indicate that the BRC-RPA fusion is functional in repairing DNA

damage induced by phleomycin when compared with *brca2-/-* mutants, but does not act as efficiently as the full length BRCA2 protein on this form of damage, unlike MMS.

Both the *C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines displayed levels of phleomycin sensitivity that were closer to the *brca2-/-* mutants than the WT or *BRCA2-/-/+* cells. This appears consistent with the MMS data, and indicates that the isolated BRC repeat region or the C terminal region of BRCA2, are largely unable to repair DNA damage.

	-/-	-/-/+	BRC+RPA	C term	Trunc
WT	0.0065	0.0840	0.2007	0.0380	0.0388
-/-		0.0002	0.0194	0.0677	0.1354
-/-/+			0.0050	0.0001	0.0002
BRC+RPA				0.0312	0.0102
C term					0.5920

Table 5.16 – Statistical analysis of the Alamar Blue results for phleomycin. P values are shown for two sample T-tests comparing the IC50s for phleomycin sensitivity of wild type cells (WT), *brca2* homozygous mutant 2 (-/-), *BRCA2-/-/+* (-/-/+), *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term), and *Trunc BRCA2-/-/+* (Trunc) mutants. Areas shaded in yellow indicate a significant difference.

As for MMS, the over-expresser cell lines again displayed no significant difference in sensitivity to phleomycin, when compared to the *OE BRCA2* cell line, or indeed WT cells.

5.5.4 Analysis of recombination efficiency

The BRCA2 variant cell lines were next subjected to analysis of their recombination efficiency following, the transformation protocol described in section 4.3.5. Three repetitions of the transformation efficiency assay were carried out for each cell line and the results are shown in figure 5.36, alongside the transformation efficiency rates for wild type, *brca2-/-2, BRCA2-/-/+* and *OE BRCA2* cell lines for comparison.

Each of the *BRC+RPA-/-/+, C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines were found to have significantly lower transformation efficiency rates than wild type or *BRCA2-/-/+* cell lines, with mean transformation efficiency rates of 0.73×10^{-6}, 0.6×10^{-6} and 0.87×10^{-6} compared to 4.53×10^{-6} and 4.27×10^{-6}, respectively. These differences were confirmed as being statistically significant by two sample T-tests displayed in table 5.18, with P values ranging from 0.003 to 0.010.
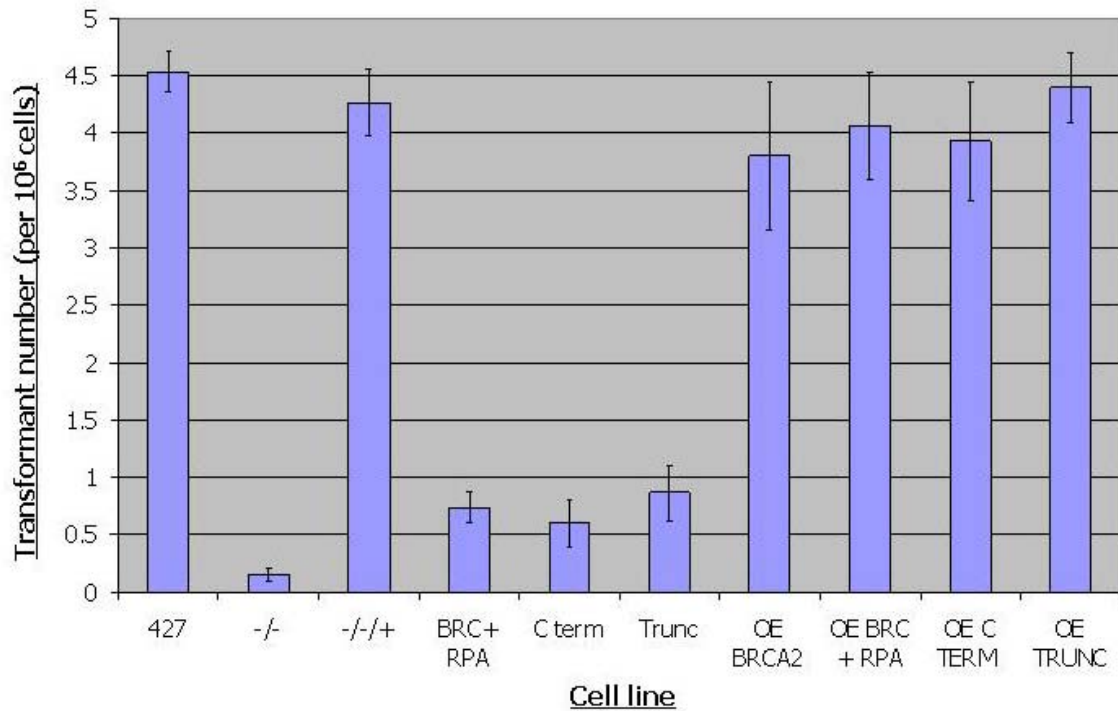
Figure 5.36 – Recombination efficiency in BRCA2 variant expressers. Values are mean numbers of transformants obtained per 10^6 cells transformed; error bars are shown from 3 repetitions. The data are presented for wild type Lister 427 cells (427), *brca2* homozygous mutant 2 (-/-), *BRCA2* re-expresser (-/-/+), BRC+RPA-/-/+ (BRC+RPA), *C term BRCA2-/-/+* (C term), *Trunc BRCA2-/-/+* (Trunc), *OE BRCA2, OE BRC+RPA*, *OE C term* and *OE Trunc* cell lines.

These results indicate that homologous recombination occurs at a low level when only the BRC repeat region or the C terminal region of BRCA2 are expressed, which may have been expected since these cell lines were found to be impaired at repairing DNA damage (section 5.5.3). Surprisingly, the BRC+RPA expresser transformation rates were also comparatively low, despite evidence that this cell line is able to repair DNA damage more effectively than the *C term BRCA2-/-/+* or *Trunc BRCA2-/-/+* mutants. This therefore infers that the DNA binding domain provided by *T. brucei* RPA cannot substitute for the C terminal region of *BRCA2* to allow homologous recombination to progress effectively. This contrasts with the findings described by Saeki *et al*., (2000), who showed that *H. sapiens* BRC repeats fused to RPA can perform recombination, albeit using a distinct assay. Nevertheless, these results appear to indicate that in *T. brucei* the full length BRCA2 protein is required for homologous recombination to occur effectively.

	-/-	-/-/+	BRC+RPA	C term	Trunc
WT	0.002	0.625	0.003	0.004	0.010
-/-		0.004	0.088	0.216	0.111
-/-/+			0.009	0.010	0.005
BRC+RPA				0.184	0.529
C term					0.270

Table 5.17 – Statistical analysis of the recombination efficiency of BRCA2 variant expressers. P values are shown for two sample T-tests comparing recombination efficiencies of wild type (WT) cells, *brca2* homozygous mutant 2 (-/-), *BRCA2* re-expresser (-/-/+), *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term) and *Trunc BRCA2-/-/+* (Trunc). Areas shaded in yellow indicate a significant difference.

The over-expresser cell lines again resulted in no significant differences in transformation efficiency rates when compared to the *OE BRCA2* cell line, or indeed WT cells.

5.5.5 Analysis of the ability of BRCA2 variants to support RAD51 foci formation

The cell lines with the various truncations of BRCA2 were next analysed for their ability to form RAD51 sub-nuclear foci following DNA damage (as described in section 4.3.6). The cells were treated with phleomycin for 18 hours and RAD51 localisation was examined by indirect immunofluorescence,

For all cells, approximately 300 cells were counted and scored for the number of foci they contained after treatment with 2 concentrations of phleomycin (0.25 μg.ml^{-1} and 1.0 μg.ml^{-1}) (see section 5.5.3). The results are displayed in table 5.19, and examples of cells and RAD51 foci are shown in figure 5.37.

Both the *C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines were found to have a greatly reduced ability to form RAD51 foci, with the majority of cells containing no foci, at any drug concentration. Indeed, the extent of this impairment is comparable to that observed for *brca2-/-* mutants, where it was not clear whether any induction of RAD51 foci occurred at all.

The results for the *BRC+RPA-/-/+* mutant demonstrated that this cell line was capable of forming RAD51 foci, albeit with some impairment. Unlike the *C term BRCA2-/-/+*, *Trunc BRCA2-/-/+* and *brca2-/-* cell lines, in the presence of phleomycin, almost half of the cells counted (44.8 %) contained 1 foci or more. This percentage of cells was, however, significantly lower than the 75 – 85 % of cells which contained foci in the WT and *BRCA2-/-/+* cell lines. These data demonstrate that the BRC-RPA fusion can support the movement of RAD51 to repair foci, but is somewhat compromised compared with WT and *BRCA2-/-/+* cell lines.

These results appear to be consistent with the DNA repair data (see above), and indicate that the full length BRCA2 is required for fully effective RAD51 localisation to DNA damage, at least caused by phleomycin. Though RPA can substitute for the C terminal region of BRCA2, suggesting that it is the BRC repeats that are primarily involved, this appears to impair the function of the protein. The reasons for this remain unclear, but it could be speculated that the RPA subunit has difficulty in removing the endogenous RPA which has coated the single stranded DNA at the sites of DNA damage. Another possibility could be that RPA fused to the BRC repeat region of BRCA2 is unable to precisely function as the C terminus, either because DNA binding is not equivalent due to sequence differences or because the C terminus of BRCA2 provides other repair functions.

		Number of foci (%)						
	BLE	0	1	2	3	4	5	6 or more
WT	0.0	96.4	3.6	0.0	0.0	0.0	0.0	0.0
	1.0	24.8	22.6	18.8	16.5	13.5	2.3	1.5
-/-	0.0	98.2	1.3	0.4	0.0	0.0	0.0	0.0
	0.25	96.9	2.0	0.5	0.5	0.0	0.0	0.0
	1.0	98.6	0.7	0.7	0.0	0.0	0.0	0.0
-/-/+	0.0	87.5	3.5	2.3	2.3	1.9	1.6	0.8
	1.0	14.8	10.5	15.6	14.8	13.7	13.3	17.2
BRC+RPA	0.0	97.6	1.5	0.9	0.0	0.0	0.0	0.0
	1.0	55.2	11.7	17.7	4.7	6.0	2.0	2.7
Trunc	0.0	99.4	0.6	0.0	0.0	0.0	0.0	0.0
	0.25	99.4	0.6	0.0	0.0	0.0	0.0	0.0
	1.0	97.7	1.3	1.0	0.0	0.0	0.0	0.0
C term	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
	0.25	99.4	0.6	0.0	0.0	0.0	0.0	0.0
	1.0	98.1	0.9	0.6	0.4	0.0	0.0	0.0

Table 5.18 – RAD51 foci formation in BRCA2 variant. The percentages of cells showing foci at given concentrations of phleomycin (BLE) are shown. Phleomycin concentrations are shown in μg.ml^{-1}. Boxes shaded in light yellow contain foci, whilst boxes shaded in bright yellow contain the highest percentage of foci.

As for the BRCA2 protein with reduced BRC repeats, to ensure that these lowered levels of RAD51 foci did not result from decreased RAD51 levels in these expresser cell lines, western analysis was carried out as described in section 5.3.5. Total protein was extracted from the *BRC+RPA-/-/+, C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines, before and after phleomycin-induced damage. Figure 5.38 demonstrates that RAD51 is still clearly expressed in the *BRC+RPA-/-/+, C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines, and there is no evidence for an increase or decrease in RAD51 levels after DNA damage.
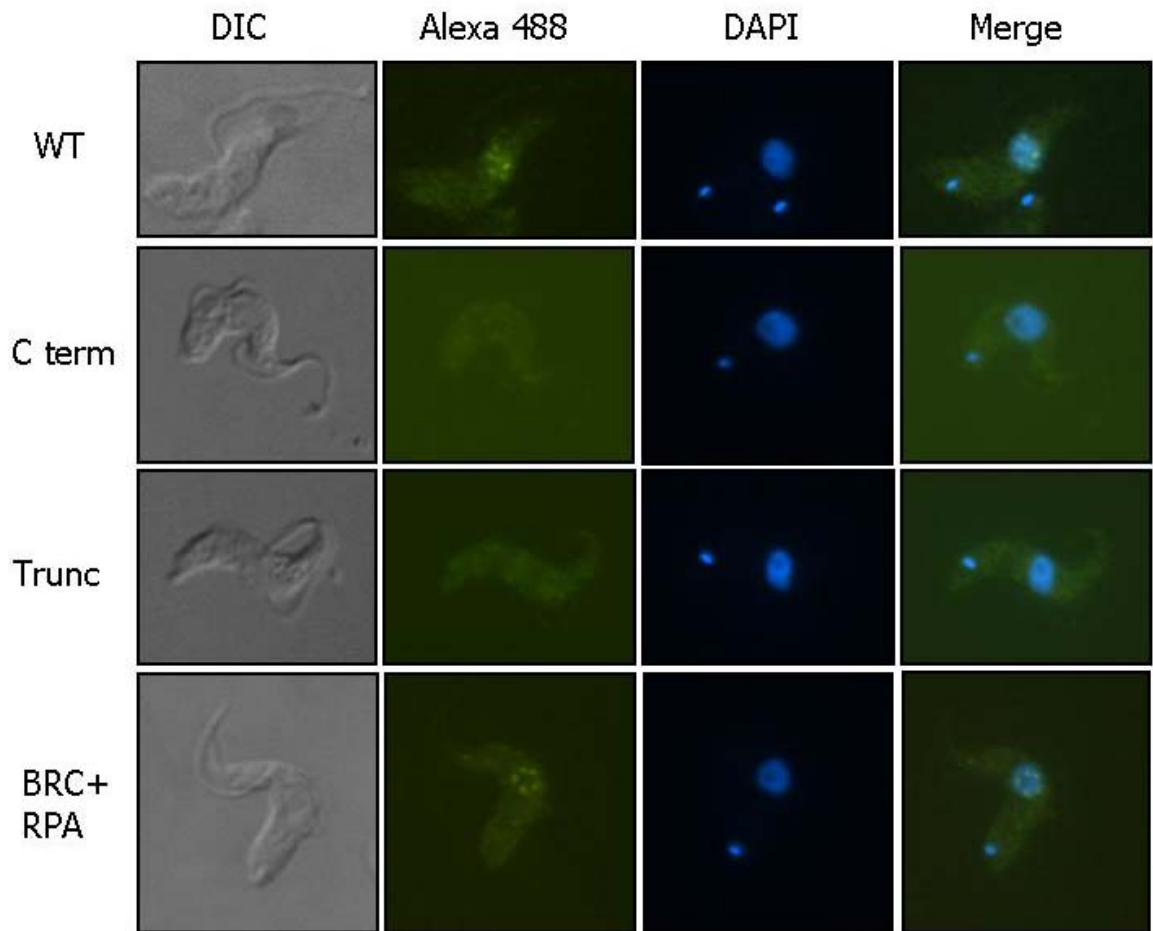
Figure 5.37 – RAD51 immunolocalisation in BRCA2 variant expressers. Representative images of *T. brucei* cells following growth in 0.25 µg.ml^{-1} and 1.0µg.ml^{-1} phleomycin for 18 hours are shown. Each cell is shown in differential interface contrast (DIC), after staining with DAPI and after hybridisation with anti-RAD51 antiserum and secondary hybridisation with Alexa Fluor 488 conjugate (Alexa 488). Merged images of DAPI and Alexa 488 cells are also shown. WT – wild type cells; C term – C term BRCA2-/-/+ cells; Trunc – Trunc BRCA2-/-/+ cells; BRC+RPA – BRC+RPA-/-/+ cells.
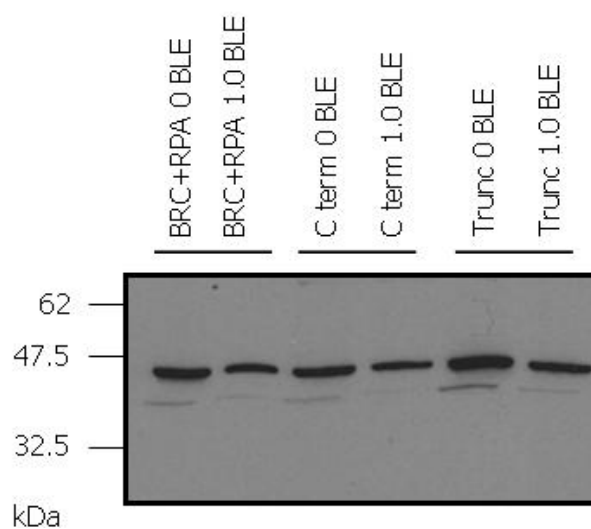
Figure 5.38 – Western blots of RAD51 in BRCA2 variant expressers. The western blots display total protein extracts from *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term) and *Trunc BRCA2-/-/+* (Trunc) cell lines probed with anti-RAD51 antiserum (RAD51). Protein extracts were prepared without damage (0µg.ml^{-1} BLE) and with damage (1.0µg.ml^{-1} BLE). Size markers are indicated. The endogenous copy of RAD51 is visible at 47kDa.

5.5.6 Analysis of VSG switching

Finally, analysis of *VSG* switching was performed in the *BRC+RPA-/-/+, C term BRCA2-/-/+* and *Trunc BRCA2-/-/+* cell lines using the same protocol described in section 5.3.6.

As before, western analysis was performed in order to determine whether VSG221 continued to be expressed in each cell line. Whole cell extracts were prepared, and electrophoresed on a 10 % SDS-PAGE gel and probed for VSG221, which resides at the active *VSG* expression site. The resulting western blots of this analysis are displayed in figure 5.39 and indicate that all cell lines were expressing VSG221.
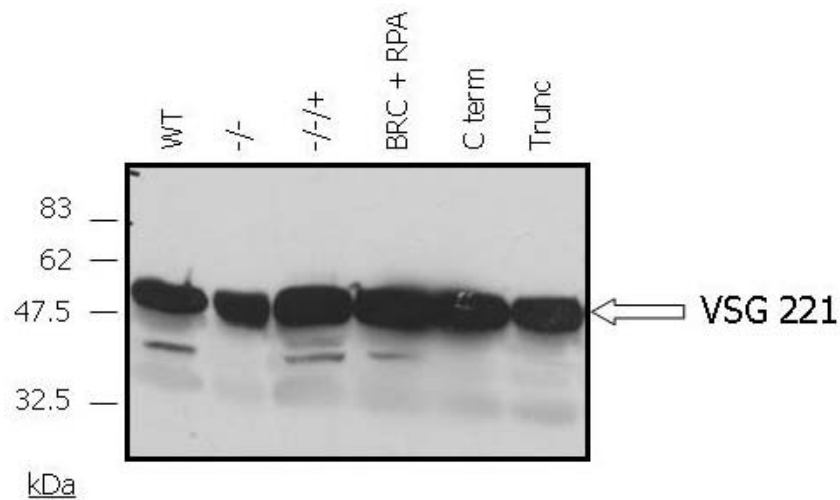
Figure 5.39 – Western blots of VSG221 in BRCA2 variant expressers. The western blots display total protein extracts from wild type (WT), *brca2-/-2* (-/-), *BRCA2-/-/+* (-/-/+), *BRC+RPA-/-/+* (BRC+RPA), *C term BRCA2-/-/+* (C term) and *Trunc BRCA2-/-/+* (Trunc) cells probed with anti-VSG 221antiserum. Size markers are indicated.

The VSG switching frequencies obtained are presented in figure 5.40, and demonstrate that the *BRC+RPA-/-/+* cell line was unaltered in its ability to switch its VSG coat relative to either the WT or *BRCA2-/-/+* cells, with a mean VSG switching frequency of 11.9 x 10^{-7} compared to 9.5 x 10^{-7} and 10 x 10^{-7}, respectively. Surprisingly, the *C term BRCA2-/-/+* cell line also appeared to be relatively unaltered in *VSG* switching frequency, with a mean VSG switching frequency of 8.2 x 10^{-7}. Only the *Trunc BRCA2-/-/+* cell line appeared to have a reduced ability to switch its VSG coat, with a mean VSG switching frequency of 4.9 x 10^{-7}, which was comparable to that found in the *brca2-/-2* mutant (3.7 x 10^{-7}).

None of these results were found to be statistically significant (table 5.20), largely due to the greater levels of variability in this assay.
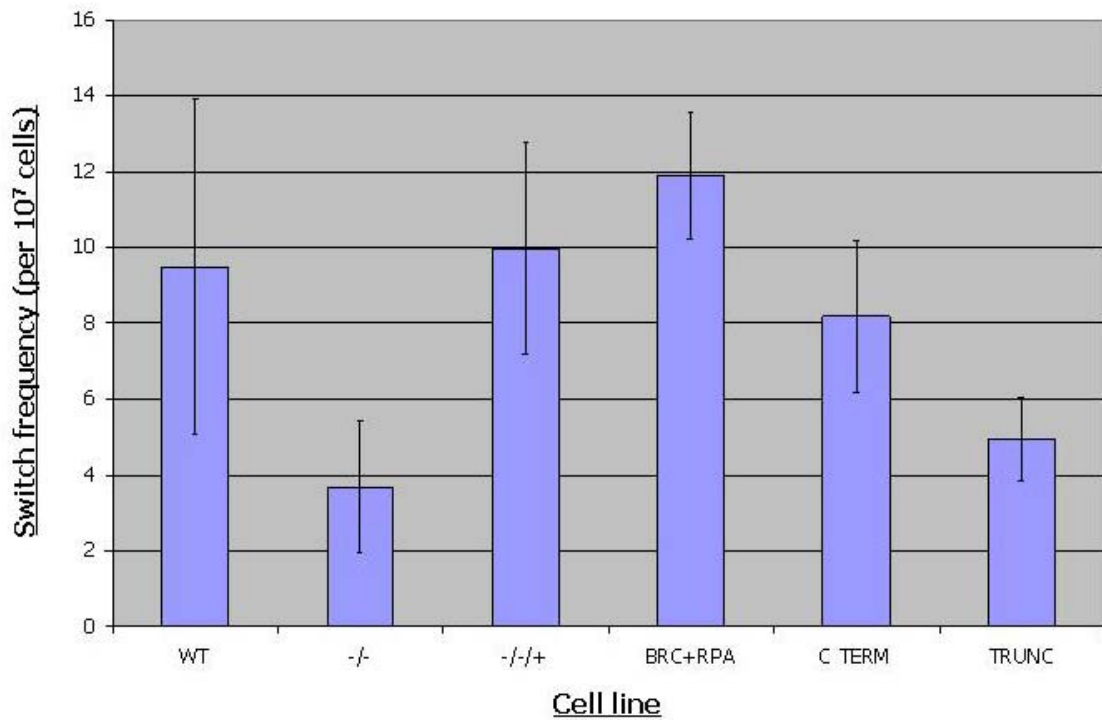
Figure 5.40 – VSG switching frequencies in BRCA2 variant expressers. Values shown are the average switching frequencies for 427 wild type cells (WT), the *brca2* homozygous mutant 2 (-/-), the *BRCA2* re-expresser (-/-/+), the *BRC+RPA-/-/+* (BRC+RPA), the *C term BRCA2-/-/+* (C term) and *Trunc BRCA2-/-/+* (Trunc). Data are from at least 3 experiments and standard error is indicated by bars.

These results appear to indicate that, in contrast to the sequence requirements for general homologous recombination, full length BRCA2 is not required for effective *VSG* switching efficiency during an acute infection. Indeed, BRCA2 can be substituted by the BRC-RPA fusion, or even the C terminal region to some extent, during this reaction.

	-/-	-/-/+	BRC+RPA	C term	Trunc
WT	0.2182	0.2743	0.0223	0.5016	0.7214
-/-		0.0675	0.1324	0.3023	0.4848
-/-/+			0.8725	0.3964	0.1202
BRC+RPA				0.4329	0.0947
C term					0.4894

Table 5.19 – Statistical analysis of the *VSG* switching frequencies in BRCA2 variant expressers. P values are shown for two sample T-tests comparing VSG switching frequencies of 427 wild type cells (WT), the *brca2* homozygous mutant 2 (-/-), the *BRCA2* re-expresser (-/-/+), the *BRC+RPA* expresser (BRC+RPA), the *C term BRCA2* expresser (C term) and the *Trunc BRCA2* expresser (Trunc) cell lines. Data are from at least 3 experiments and standard error is indicated by bars. Areas shaded in yellow indicate a significant difference (P<0.05).

5.6 Summary

The aims of this chapter were to examine the function of the BRC repeat expansion in *T. brucei* BRCA2, and to establish the functions of various motifs of BRCA2, in particular to ask if the BRC repeats and downstream C terminal part of the protein provide elements that can act in isolation or need to be present together. To do this, expresser cell lines were generated by transforming *brca2-/-* cells with constructs containing *BRCA2* genes which contained reduced numbers of BRC repeats, the BRC repeat region alone, the C terminal region alone and the BRC repeat region translationally fused to the large RPA subunit.

T. brucei brca2-/- cells, like mutants of other DNA recombination genes (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005;Robinson *et al.*, 2002), display a reduced growth rate *in vitro* and *in vivo*. *In vitro* growth rates were only seen to fully return to the levels of WT and *BRCA2-/-/+* cells in the *BRC+RPA-/-/+* expresser cell line, indicating that the RPA 50 subunit can substitute for the C terminal region in terms of population doubling times, and in the *1BRC-/-/+* cells, indicating that a single divergent BRC repeat can also function.

The generation of aberrant cell types in the *brca2-/-* cells was reverted to levels comparable with WT and *BRCA2-/-/+* cells in the expressers with reduced numbers of BRC repeats. Since these truncated BRCA2 variants were DNA repair impaired, this indicates that the BRC repeat expansion is not a component of BRCA2-related cell cycle progression. It appears likely that the C terminal region of BRCA2 might function in this role, based on 2 observations. First, the *C term BRCA2-/-/+* cell lines cell cycle distribution was distinct from *brca2-/-* cells, and was more comparable to WT and *BRCA2-/-/+* cells. Second, BRC-RPA fusion, in which the C terminus of BRCA2 was replaced with the RPA 50 subunit continued to generate aberrant cells, arguing that BRCA2, C terminal-specific functions underlie this phenotype.

Only the *BRC+RPA-/-/+* cell line was able to restore resistance against DNA damaging agents to *brca2-/-* mutants, to levels comparable with full length BRCA2, indicating that the RPA 50 subunit can substitute for the C terminal region of BRCA2 in terms of repairing DNA damage induced by MMS and to a certain degree, phleomycin. The lack of such complementation by the BRCA2 variant with 1 BRC repeat argues that the BRC repeat expansion in BRCA2 is therefore critical in terms of DNA damage repair. However, because the 1BRC and *T. vivax* BRCA2 variants provide some enhancement of DNA repair efficiency in *brca2-/-* cells, this argues that the divergent C-terminal BRC

repeat, and the *T. vivax* BRC repeat, function in *T. brucei*, most likely through RAD51 interaction.

Surprisingly, none of the cell lines expressing the BRCA2 variants were able to support efficient recombination. This indicates that the BRC repeat expansion is critical for efficient homologous recombination, and that the RPA 50 subunit cannot substitute for the C terminal BRCA2 region in order to allow homologous recombination to progress. This could be due to the RPA subunit fused to the BRC repeat region interfering with the endogenous RPA, though we have not tested this, but contrasts with findings in mammalian cells and in *U. maydis* (Kojic *et al.*, 2005;Kojic *et al.*, 2006;Saeki *et al.*, 2006).

Only the *BRC+RPA-/-/+* cell line was capable of forming phleomycin-induced sub-nuclear RAD51 foci, again implying that the BRC repeat expansion is critical for efficient interaction with RAD51 during the DNA damage response in *T. brucei*. Despite this, the BRC-RPA fusion appeared not to function as efficiently as full length BRCA2, again arguing that the RPA 50 subunit could not fully substitute for the C terminal region of BRCA2, which appears consistent with the impaired function of this protein during transformation assays.

Despite all the above findings, the most surprising result from this chapter came from the VSG switching data. The 1BRC and *T. vivax* BRCA2 variants supported this reaction, which demonstrated that the BRC repeat expansion in BRCA2 appears to be of little importance for VSG switching efficiency during an acute infection. Perhaps more surprisingly, the BRC-RPA and C-term variants of BRCA2 supported VSG switching to a greater and lesser extent respectively. This argues that VSG switching could still occur when only fragments of the BRCA2 protein were expressed. The large amount of variation in the data produced from this assay could account for these unexpected results. However, it is also legitimate to ask what this tells us about the role of BRCA2 in VSG switching.

In addition to the BRCA2 variants that were expressed in *brca2-/-* mutants, variants were also over-expressed in WT cells. However, no apparent effects were detected in any of the assays, perhaps indicating that little or no over-expression occurred.

<u>CHAPTER 6</u>

Looking for RAD51 interacting factors in *T. brucei*

6.1 Introduction

RAD51, the eukaryotic homologue of bacterial RecA, is highly conserved in all eukaryotes. It is a relatively small protein (38kDa) that is functional as a long helical polymer, made up of hundreds of monomers that wrap around DNA to form a nucleoprotein filament (Benson *et al.*, 1994), which functions in the repair of double strand breaks (DSBs) by the homologous recombination pathway (see sections 1.4.2.3 and 1.6).

Studies on human RAD51 showed its strand exchange activity to be much lower than that of RecA *in vitro* (Baumann *et al.*, 1996), suggesting the requirement of additional factors such as the single strand-binding protein RPA (replication binding protein A), RAD52 and RAD54. In both yeast and mammals, RAD52 facilitates the removal of RPA from ssDNA (Sung, 1997a;Benson *et al.*, 1998), whilst RAD54 is thought to stimulate joint-molecule formation (Petukhova *et al.*, 1998) (see section 1.4.2.3).

Following DNA damage to bacteria, RecA induction increases more than 15 fold in a 'SOS response'(Little and Mount, 1982;Walker, 1984). Whilst this up-regulation is observed in virtually all eukaryotes, including *S. cerevisiae*, *T. cruzi* and *L. major*, no such response is observed in mammalian cells (Tarsounas *et al.*, 2004). At best there is a two fold transcriptional regulation of proteins such as RAD51. Instead, RAD51 and other repair proteins that are normally diffused throughout the nucleus are rapidly relocated and concentrated into sub-nuclear complexes that are microscopically detected as foci. This creates an overall effect that increases the local concentration of repair enzymes as the cell prepares for and undergoes repair (Tarsounas *et al.*, 2004). Recombination proteins that are known to co-localise with RAD51 include RAD52 (Liu and Maizels, 2000;Lisby *et al.*, 2001;Essers *et al.*, 2002b), RAD54 (Tan *et al.*, 1999;Essers *et al.*, 2002a), RPA (Raderschall *et al.*, 1999) and the tumour suppressors BRCA1 (Scully *et al.*, 1997) and BRCA2 (Chen *et al.*, 1998b). These foci are hypothesised to exist as repair centres, though their exact composition is unknown.

Cells defective in any of the five mammalian RAD51 paralogues (RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3), which are required for normal levels of HR and resistance to ionising radiation (Thacker, 1999), either fail to form or are reduced in their ability to generate or maintain RAD51 foci in response to ionising radiation (Bishop *et al.*, 1998;Takata *et al.*, 2000;Takata *et al.*, 2001;O'Regan *et al.*, 2001;Tarsounas *et al.*, 2004). Both BRCA2 and DSS1 have also been shown to be essential in the formation of RAD51 foci in mammals (Yuan *et al.*, 1999;Tarsounas *et al.*, 2003;Gudmundsdottir *et al.*,

2004;Godthelp *et al.*, 2002). RAD51 is known to directly bind BRCA2 *in vitro*, via the BRC repeats and at a separate locus in mammalian cells, *U. maydis* and *C. elegans*. (Wong *et al.*, 1997;Chen *et al.*, 1998b;Chen *et al.*, 1999a;Esashi *et al.*, 2005;Zhou *et al.*, 2007;Petalcorin *et al.*, 2007). RAD51 is also known to interact with several other proteins (see section 1.6) including p53 (Sturzbecher *et al.*, 1996;Buchhop *et al.*, 1997;Linke *et al.*, 2003) indicating a role in genome maintenance in higher eukaryotes (Sonoda *et al.*, 1998).

RAD51 foci are also found in undamaged S-phase mammalian cells, where they are proposed to repair broken replication forks (Tashiro *et al.*, 1996;Raderschall *et al.*, 1999). The S-phase and damaged induced foci appear to be distinct from each other, as BRCA2 is not required for the formation of RAD51 foci in non-irradiated S-phase cells (Tarsounas *et al.*, 2003).

Interestingly, the presence of RAD51 foci formations following DNA damage does not appear to be restricted to mammalian cells. Indeed, these damage induced foci have been observed to form in many other eukaryotes, including *S. cerevisiae*, *C. elegans* and *U. maydis* (Kojic *et al.*, 2005;Martin *et al.*, 2005;Bishop, 1994;Gasior *et al.*, 1998).

In *T. brucei*, RAD51 is considered to be one of the most significant genes to be implicated in regulating VSG switching. It was initially hypothesised that *rad51-/-* mutants would be unable to undergo VSG switching. However, although impaired, the *T. brucei rad51-/-* mutants were still capable of switching their VSG coat (McCulloch and Barry, 1999), indicating that RAD51 is not the only protein involved in the complex process of antigenic variation in *T. brucei*. Recent work has shown that *T. brucei* RAD51 forms sub-nuclear foci in response to DNA damage (Proudfoot and McCulloch, 2005). Although such foci have been previously described in other organisms, their role and molecular composition in *T. brucei* remains unclear. The work described in section 4.3.6 appears to support the hypothesis that *T. brucei* RAD51 interacts with BRCA2. Indeed, the absence of BRCA2 causes the apparent failure of RAD51 foci formation, as does mutation of 2 RAD51 paralogues, RAD51-3 and RAD51-5 (Proudfoot and McCulloch, 2005). However, no work has examined if any of these factors interact with RAD51 in foci.

In this chapter, the hypothesis will be tested that the *T. brucei* RAD51 foci are repair centres containing multiple homologous recombination factors and specific sites of DNA lesions. Potentially, such an approach could identify *T. brucei* HR factors that have not been annotated through sequence homology with other organisms. Tandem affinity

purification will be used to identify RAD51 interacting factors, before and after induced DNA damage in both the bloodstream and procyclic stages of the *T. brucei* life cycle.

6.2 Tandem affinity purification

Affinity purification has been the method of choice for a number of years for purifying proteins (Shevchenko *et al.*, 1996;Blackstock and Weir, 1999). However, protein complexes do not always tolerate the over-expression of specific factors, as this can result in non-physiological interactions and may disrupt protein complexes (Swaffield *et al.*, 1995). Therefore, in order for protein complexes to be purified, the target protein needs to be close to its natural expression levels.

In 1999, Rigaut *et al.* (1999) developed a generic protein purification method for protein complex characterisation (Rigaut *et al.*, 1999). The tandem affinity purification (TAP) method was found to generate high yields of protein complexes from cell fractions and prior knowledge of the complex composition or function was not required. This method has similar applications to the yeast two-hybrid screen, but has the advantage that multiple interacting partners, rather than simply 2 proteins, can be identified in a single experiment (Fromont-Racine *et al.*, 1997).

The TAP tag is a fusion cassette, consisting of an IgG binding domain of *Staphylococcus aureus* protein A (Prot A) and a calmodulin binding peptide (CBP), separated by a TEV protease cleavage site. The Prot A and CBP tags were chosen, as these tags allowed efficient recovery of a fusion protein (SmX4p) that was present at a low concentration in extract from *S. cerevisiae* (Rigaut *et al.*, 1999). The conditions required for the selection and subsequent release from the IgG and calmodulin beads are contrasting, so the authors decided to insert a specific TEV protease recognition sequence between the 2 tags, allowing the release of the Prot A tag prior to CBP-bound purification (Rigaut *et al.*, 1999).

The TAP tag proved to be flexible, as the relative order of the domains can be inversed to produce N and C terminal tags. Following fusion of the TAP tag to the target protein, the construct is introduced into the host cell or organism, ideally maintaining the natural expression level of the protein. The purification method is depicted in figure 6.1 and involves the recovery of the fusion protein and possible interacting partners from the cell extract initially by affinity selection on an IgG matrix. Following washing, TEV protease releases the bound complexes. These complexes are then next bound onto calmodulin-coated beads in the presence of calcium. This second affinity step not only serves to

remove the TEV protease but also to remove furthers traces of contaminants that have been left from the single purification step. After washing, the bound material is released by EGTA and proteins analysed by mass spectrometry (MS). Rigaut *et al.* (1999) found that both purification steps were required, as purification with only the IgG beads or the calmodulin coated beads resulted in a significantly higher level of contaminants, compared to when the 2 step procedure was used. These authors also documented that TAP tagged proteins can remain functional, demonstrated by TAP tagging the small subunit of yeast cap binding complex (CBC). The resulting TAP-purified CBC was capable of forming specific complexes with radioloabelled capped RNA, indicating activity.

This methodology has proved successful for many different proteins and in many different organisms, including the Cf-9 protein in *Nicotiana benthamiana (Rivas et al., 2002),* the Ltp-28 protein in *Leishmania tarentolae* (Aphasizhev *et al.*, 2003) and the SRP-19 and RPA 12 proteins in *T. brucei* (Lustig *et al.*, 2005;Walgraffe *et al.*, 2005). Most notably, this method has proved to be extremely successful in *S. cerevisiae,* having been used recently to identify interacting factors in 2357 proteins (Krogan *et al.*, 2006).

In order to determine if this method would prove useful for identifying RAD51 interacting partners, the RAD51 sequence needed to be examined for the presence of a TEV protease cleavage site. No such sequence was identified in RAD51, but it remains a possibility that such a sequence may exist within interacting partners. However, this possibility remains low due to the high specificity of the TEV protease (Dougherty *et al.*, 1989).
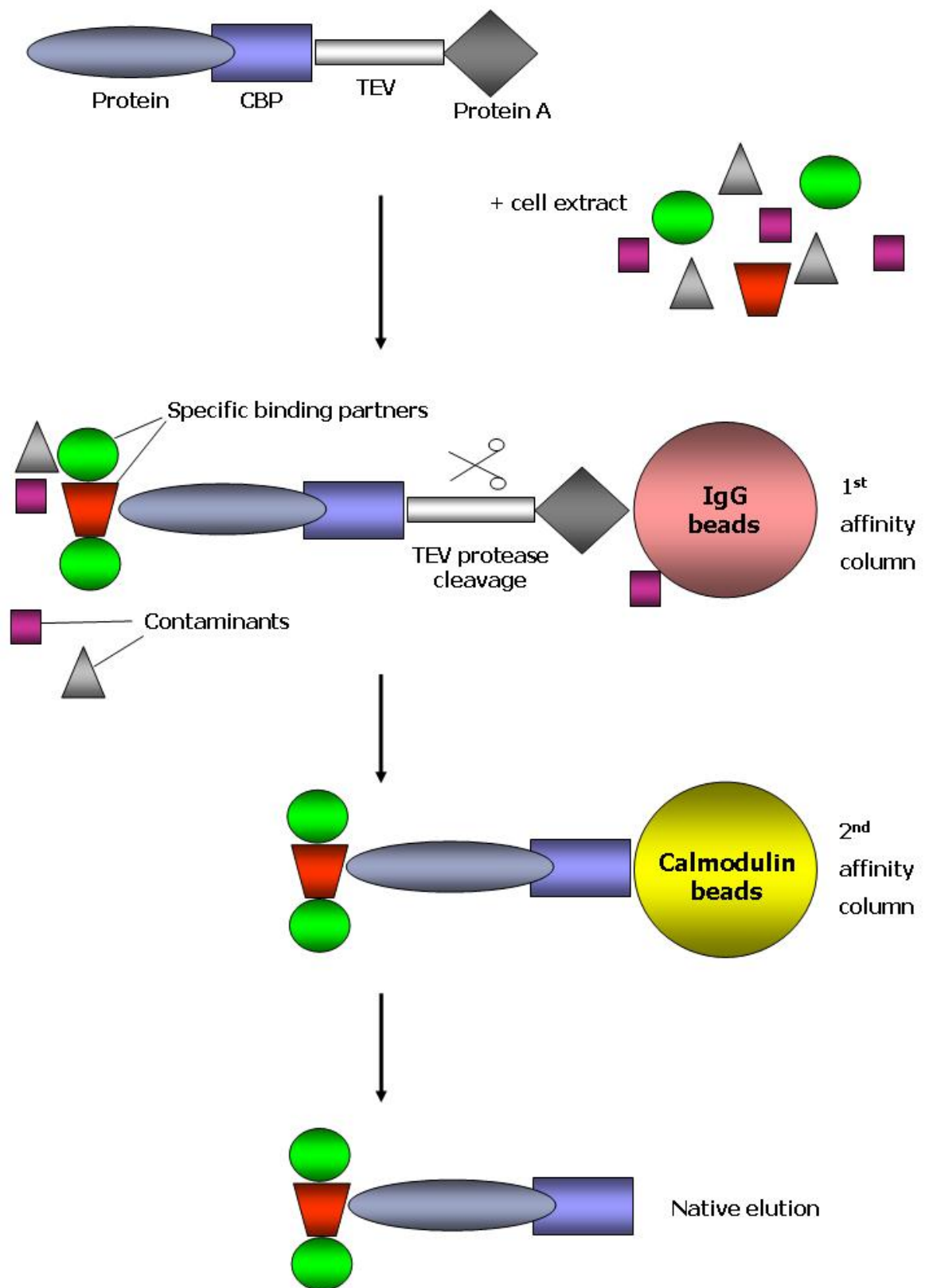
Figure 6.1 – Overview of TAP protocol. This method involves the fusion of the TAP tag to the target protein and the introduction of the construct into the organism, ideally maintaining the expression of the fusion protein at its natural level. The fusion protein and associated components are recovered from cell extracts by affinity selection on an IgG matrix. After washing, the TEV protease is added to release the bound material. The eluate is incubated with calmodulin-coated beads in the presence of calcium. This second affinity step is required to remove the TEV protease as well as traces of contaminants remaining after the first affinity selection. After washing, the bound material is released with EGTA. Figure adapted from Rigaut *et al*., 1999.

6.3 Generation of TAP tagged RAD51, both N and C terminally

It was decided to tag both the N and C termini of RAD51 in order to maximise the potential of identifying co-factors. However, through examining the three dimensional structure of a RAD51 nucleoprotein filament (Conway *et al.*, 2004), it was noted that the N terminus appeared to be exposed, whilst the C terminus appeared to be buried within the filament. This led to the assumption that tagging the N terminus of RAD51 would be less likely to impair the function of the protein compared with a C terminal TAP tag. However, the addition of a tag at either end of the protein could have caused detrimental effects on the functionality of the protein, so the addition of a tag at both termini was performed.
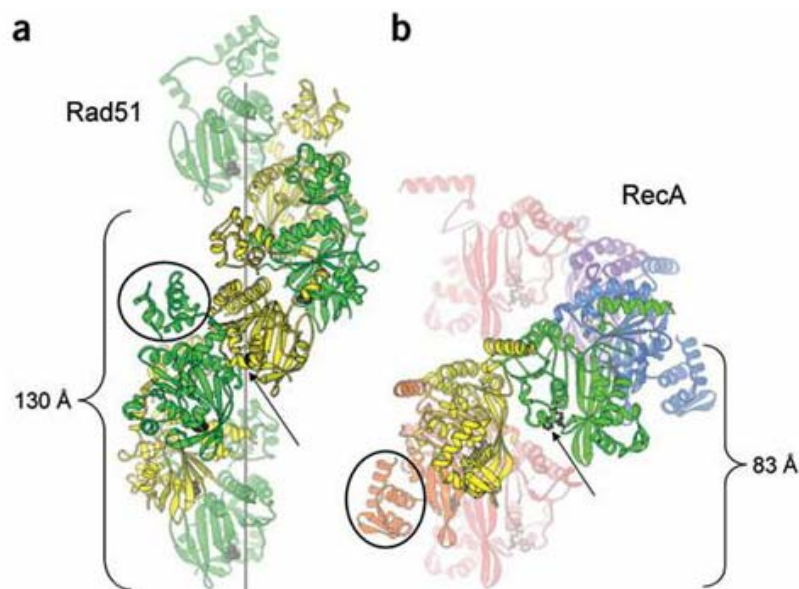
Figure 6.2 – RAD51 and RecA filaments. Filaments of (a) Rad51 and (b) RecA are shown with helical pitches of 130 Å and 83 Å respectively. Circled areas of Rad51 and RecA represent the N and C termini, respectively. Figure taken from Conway *et al.,* 2004.

The N terminal RAD51 TAP tag construct was generated by a PCR method. This method is depicted in figure 6.3 and uses PCR to amplify a region that corresponds with the 5' UTR upstream of the ORF, and a region at the very start of the ORF, rather than more conventional cloning methods. One oligonucleotide primer (*NTAP5'*) was designed that contained 99 bases of sequence that was homologous to the sequence upstream of the *RAD51* ORF start codon and 21 bases of sequence that was homologous to the neomycin antibiotic resistance cassette (NEO) in the plasmid pGL960 (a gift from M. Carrington). An equivalent primer (*NTAP3'*) was designed which contained 99 bases of sequence that was complementary to the start of the *RAD51* ORF, but not including the start codon, and 21 bases of sequence that was complementary to the calmodulin binding protein (CBP) region of pGL960.
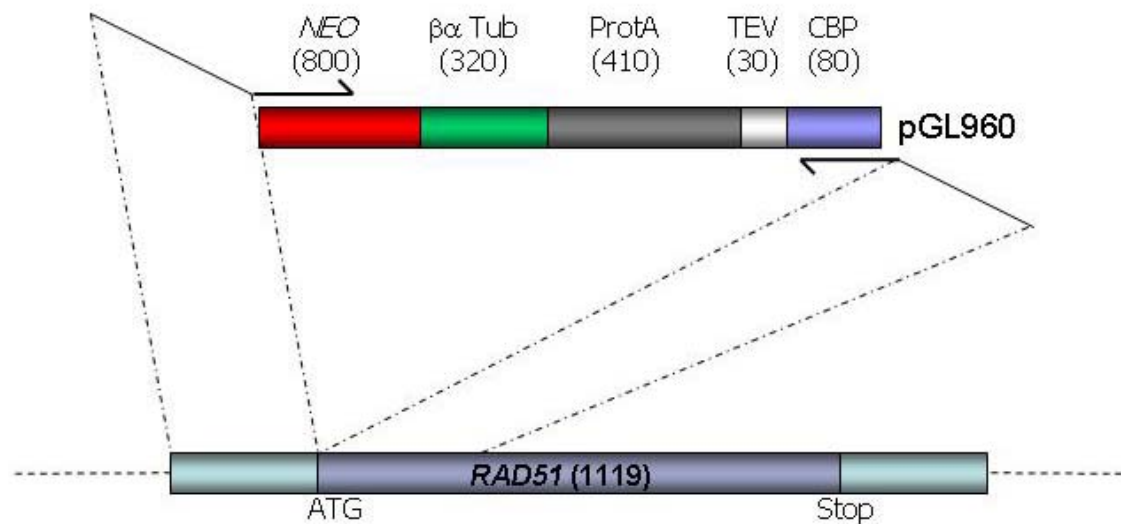
Figure 6.3 – Strategy for obtaining N terminal TAP tag constructs by PCR. PCR primers were designed, of which a forward primer contained 99 bases of sequence homologous to the 5' flank of *RAD51* and 21bases of sequence homologous to the NEO region of the plasmid pGL960. A reverse primer contained 99 bases of sequence complementary to the 5' end of the *RAD51* ORF and 21 bases of sequence complementary to the CBP region of pGL960. 5' flank corresponds to the region upstream of the RAD51 ORF, whilst the 5' end corresponds to the start of the RAD51 ORF. βα Tub: βα tubulin intergenic region (processing signal). NEO: neomycin resistance gene ORF. ProtA: protein A domain. TEV: TEV target site. CBP: calmodulin binding domain.

PCR-amplification using these primers, and pGL960 as template, generated the DNA fragments *ΔRAD51::NTAP*, which should add a TAP tag to the 5' end of *RAD51* using the 5' flanking and 5' start sequences to integrate the construct by homologous recombination (figure 6.4). The PCR generated a DNA fragment of the expected size (1840 bp), which was subsequently PCR purified (section 2.7.1.1) and approximately 5 μg used for transformations.
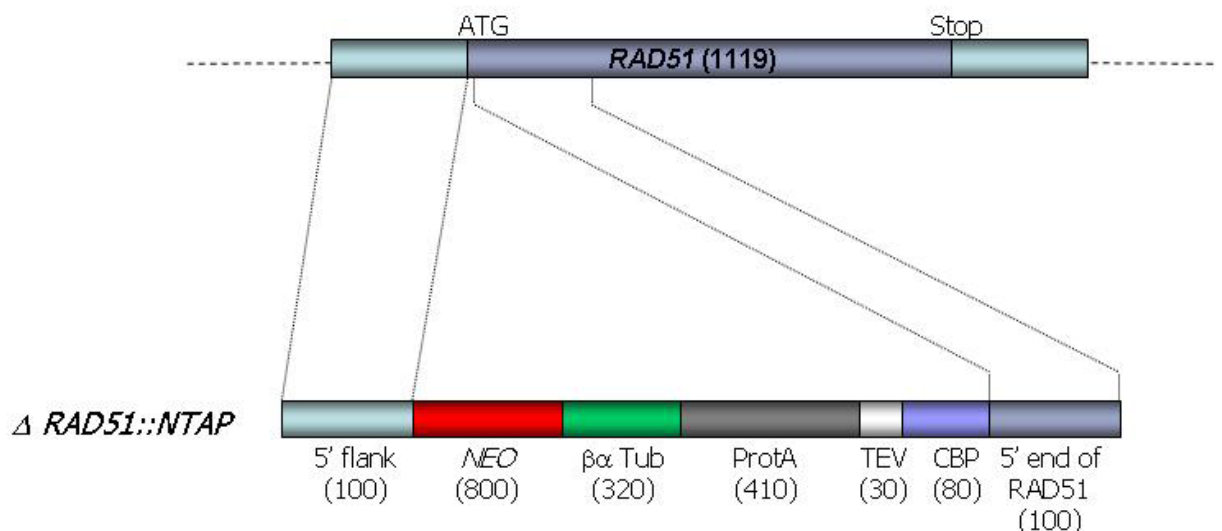
Figure 6.4 – Generation of N terminally TAP tagged RAD51. Homologous recombination allows the *Δ RAD51::NTAP* construct to integrate at the 5' end of *RAD51*. Sizes of the individual components are shown in base pairs. 5' flank and 5' end correspond to the homologous regions upstream, and the start, of the *RAD51* ORF, respectively. 5' flank corresponds to the region upstream of the *RAD51* ORF, whilst the 5' end corresponds to the start of the *RAD51* ORF. βα Tub: βα tubulin intergenic region (processing signal). NEO: neomycin resistance gene ORF. ProtA: protein A domain. TEV: TEV target site. CBP: calmodulin binding domain.

The C terminal RAD51 TAP tag construct was similarly generated using a PCR method (figure 6.5). One oligonucleotide primer (*CTAP5'*) was designed which contained 99 bases of sequence that was homologous to the end of the *RAD51* ORF, but not including the stop codon, and 21 bases of sequence that was homologous to the calmodulin binding protein (CBP) region of pGL900 (a gift from M. Carrington). An equivalent primer (*CTAP3'*) was designed which contained 99 bases of sequence that was complementary to the region downstream of the *RAD51* ORF and 21 bases of sequence that was complementary to βα Tub region of the plasmid pGL960.
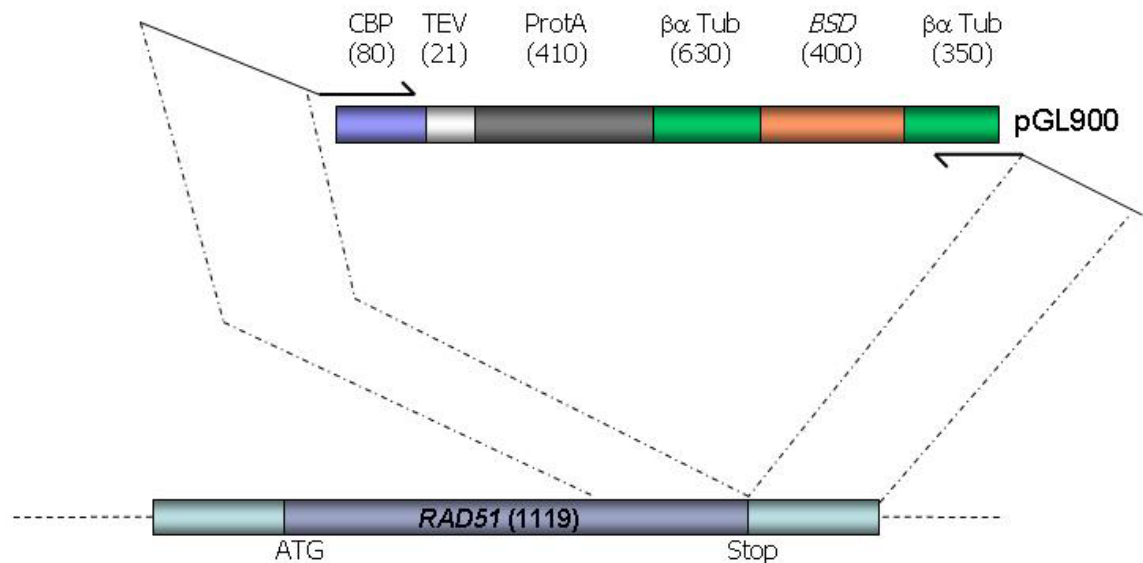
Figure 6.5 – Strategy for obtaining C terminal TAP tag constructs by PCR. PCR primers were designed, of which a forward primer contained 99 bases of sequence homologous to the 3' end of RAD51 ORF and 21bases of sequence homologous to the CBP region of pGL900. A reverse primer contained 99 bases of sequence complementary to the 3' flank of the *RAD51* ORF and 21 bases of sequence complementary to the βαTub region of pGL900. 3' flank corresponds to the region downstream of the *RAD51* ORF, whilst the 3' end corresponds to the end of the *RAD51* ORF. βα Tub: βα tubulin intergenic region (processing signal). NEO: neomycin resistance gene ORF. ProtA: protein A domain. TEV: TEV target site. CBP: calmodulin binding domain.

PCR-amplification using these primers, and pGL900 as template, generated the DNA fragment *ΔRAD51::CTAP*, which should add a TAP tag to the 3' end of *RAD51* using the 3' end and 3' flanking sequences to integrate the construct by homologous recombination (figure 6.5). This PCR generated a DNA fragment of the expected size (1240 bp), which was subsequently PCR purified (section 2.7.1.1) and approximately 5 µg used for transformations.
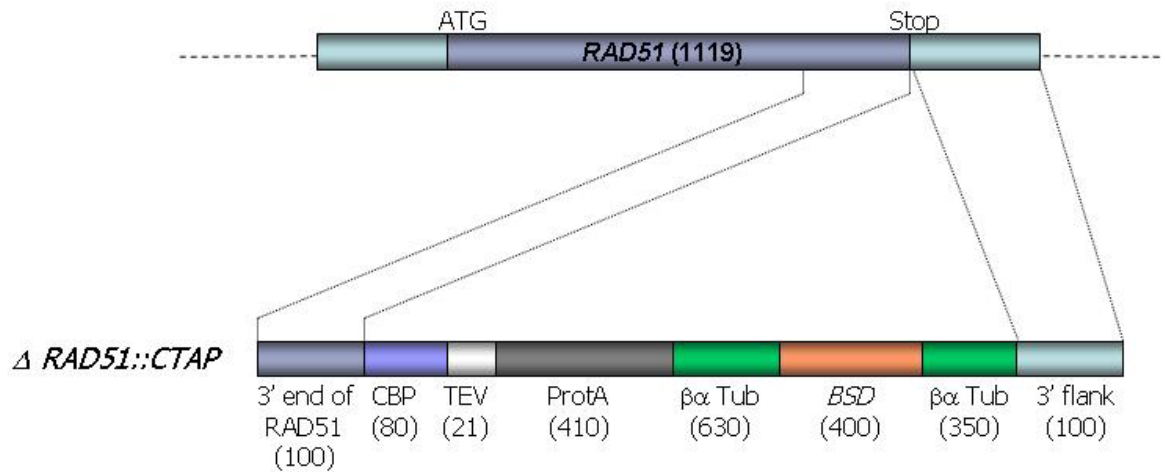
Figure 6.6 – Generation of C terminally TAP tagged RAD51. Homologous recombination allows the construct to integrate at the 3' end of *RAD51*. Sizes of the individual components are shown in base pairs. 3' flank and 3' end correspond to the homologous regions downstream and the end of the *RAD51* ORF. 3' flank corresponds to the region downstream of the *RAD51* ORF, whilst the 3' end corresponds to the end of the *RAD51* ORF. αβ Tub: αβ tubulin intergenic region (processing signal). NEO: neomycin resistance gene ORF. ProtA: protein A domain. TEV: TEV target site. CBP: calmodulin binding domain. The construct allows homologous integration to C terminally tag RAD51.

In terms of identifying RAD51 interacting factors in *T. brucei*, we were most interested in examining the bloodstream stage of the life cycle, as it is in this stage whereby *T. brucei* undergo VSG switching, which has been the main focus of my thesis, and it is conceivable that novel, RAD51-interacting factors guide this process. However, a difficulty with the TAP procedure in the bloodstream stage of *T. brucei*, is growing sufficiently large numbers of cells. From previous work, typically 1 x 10^{10} *T. brucei* cells have been used (Walgraffe *et al.*, 2005;Lustig *et al.*, 2005;Laufer *et al.*, 1999). For this reason, it was decided to TAP tag both the bloodstream and procyclic stages of *T. brucei*. Initial attempts to identify RAD51 interacting factors would therefore be performed in procyclic form cells, where it is much easier to generate larger numbers of cells, due to their ability to grow to much denser populations *in vitro*. TAP tagging in the bloodstream stage would check the viability of modifying RAD51 here, and allow further comparable purifications.

For each TAP construct, *ΔRAD51::NTAP* and *ΔRAD51::CTAP*, two separate transformations were carried out in both Lister 427 (Melville *et al.*, 2000) bloodstream stage cells and in EATRO 795 procyclic form cells, in order to generate two independent N- and C-terminal TAP tagged cell lines in both life cycle stages. To do this, Lister 427 bloodstream stage cells were transformed using the protocol described in section 2.1.3.1 and antibiotic resistant transformants were selected by placing cells on 2.5 μg.ml^{-1} G418 or 5 μg.ml^{-1} blasticidin for *ΔRAD51::NTAP* and *ΔRAD51::CTAP*, respectively. The generation of TAP tagged transformants was subsequently confirmed by PCR, western and

Southern analysis (see below). EATRO 795 procyclic form cells were transformed using the protocol described in section 2.1.3.2 and antibiotic resistant transformants were selected by placing cells on 5 µg.ml^{-1} G418 or 10 µg.ml^{-1} blasticidin for *ΔRAD51::NTAP* and *ΔRAD51::CTAP*, respectively. Polyclonal cell lines were confirmed by PCR and western blot analysis (sections 6.3.1 and 6.3.2), before generating clonal cell lines as described in section 2.1.3.2. Clonal cell lines were subsequently confirmed as having correctly TAP tagged RAD51 by PCR, western and Southern analysis (see below).

6.3.1 Confirmation of TAP tagged RAD51 by PCR

To initially confirm the generation of N and C terminal RAD51 TAP tagged cells in bloodstream and procyclic stages of *T. brucei*, PCR analysis was carried out. Two oligonucleotide primers were designed for the N terminal TAP tagged transformants, one of which (*NanalTAP5′*) contained 22 bases of sequence that was homologous to the start of the Protein A region of the tag and another (*NanalTAP3′*) which contained 22 bases of sequence that was complementary to a region in *RAD51*. Similarly, two oligonucleotide primers were designed for the C terminal TAP tagged transformants. One of these (*CanalTAP5′*) contained 22 bases of sequence that was homologous to a region in *RAD51*and another (*CanalTAP3′*) which contained 22 bases of sequence that was complementary to the end of the Protein A region of the tag. The basis of this approach is depicted in figure 6.7, which demonstrates that PCR amplification should only occur if the *ΔRAD51::NTAP* and *ΔRAD51::CTAP* constructs have correctly integrated at the 5' and 3' ends of *RAD51*, respectively.
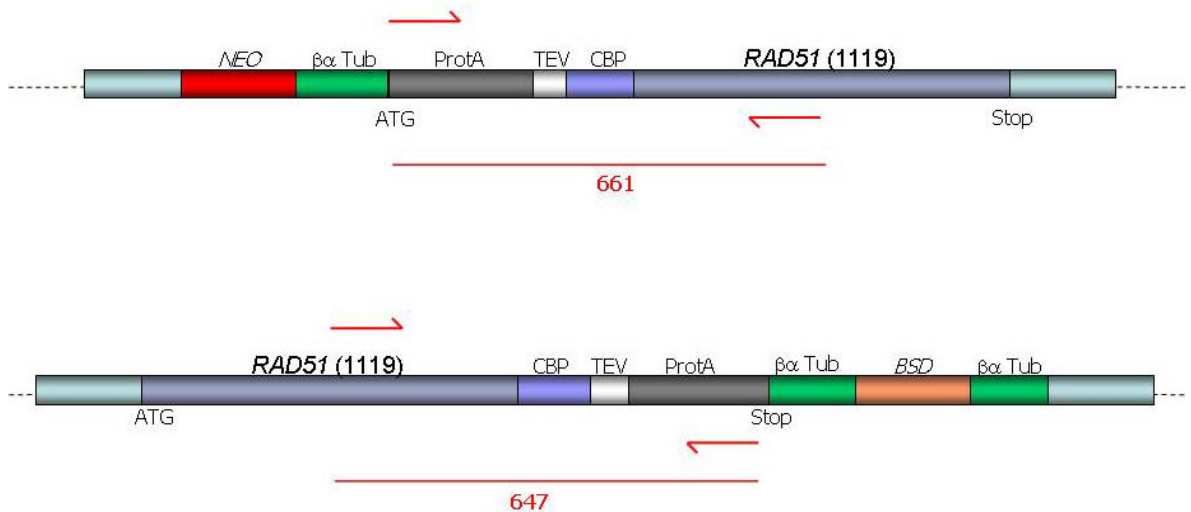
Figure 6.7 – Strategy for analysing TAP tagged transformants by PCR. Oligonucleotide primers were designed to only amplify a DNA fragment if the construct had integrated into the expected position. The upper diagram displays the correct integration of the *ΔRAD51::NTAP* construct. A forward primer was designed which contained 22 bases of sequence that was homologous to the start of the Protein A region of the tag and another which contained 22 bases of sequence that was complementary to a region in *RAD51*. The lower diagram displays the correct integration of the *ΔRAD51::CTAP* construct. A forward primer was designed which contained 22 bases of sequence that was homologous to a region in *RAD51* and another which contained 22 bases of sequence that was complementary to the end of the Protein A region of the tag. Forward and reverse primers are shown as red arrows. Predicted DNA fragment sizes are indicated in base pairs, as are the sizes of the RAD51 ORF. αβ Tub: αβ tubulin intergenic region (processing signal). NEO: neomycin resistance gene ORF. BSD: blasticidin resistance gene ORF. ProtA: protein A domain. TEV: TEV target site. CBP: calmodulin binding domain.

Genomic DNA was prepared from independent polyclonal procyclic form transformants arising from transformations using both the *ΔRAD51::NTAP* and *ΔRAD51::CTAP* constructs. PCR analysis was carried out using Taq DNA polymerase and the primers *NanalTAP5'* and *NanalTAP3'* on 2 G418 resistant populations, and *CanalTAP5'* and *CanalTAP3'* on 4 blasticidin resistant populations. The resulting PCR products were separated on a 1 % agarose gel, before being visualised under UV illumination. The results are displayed in figure 6.8 and demonstrate that both of the N terminal TAP tagged populations appeared to have correctly integrated the *ΔRAD51::NTAP* construct. However, only two of the putative C-terminal TAP tagged populations appeared to have correctly integrated the *ΔRAD51::CTAP* construct. Presumably the polyclonal populations that were resistant to blasticidin but failed to amplify a PCR product (C term 1 and 2) had integrated the construct into an ORF somewhere in the genome, though the actual integration locus was not investigated further.
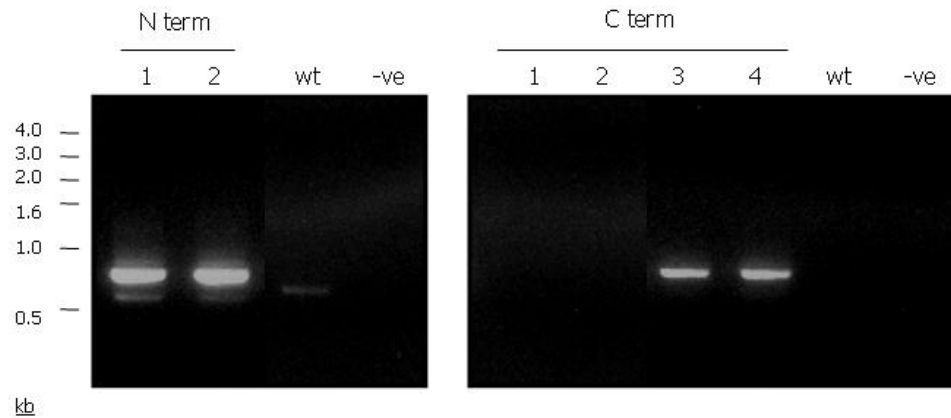
Figure 6.8 – PCR analysis of TAP tagged RAD51 transformants in polyclonal procyclic populations. The ethidium stained gel shown on the left displays PCR amplification using the primer pairs *NanalTAP5'* and *NanalTAP3'* to analyse N terminal TAP tagged RAD51. The ethidium stained gel shown on the right displays PCR amplification using the primer pairs *CanalTAP5'* and *CanalTAP3'* to analyse C terminal TAP tagged RAD51. The PCR products are displayed from 2 polyclonal G418 resistant populations (N term 1 and 2); 4 polyclonal blasticidin resistant populations (C term 1, 2, 3 and 4); wild type cells (wt) and a no DNA control (-ve). Reactions were carried out using the primer pairs depicted in figure 6.7. DNA size markers are indicated in kbp.

Genomic DNA was similarly prepared from the clonally derived bloodstream *T. brucei* transformants. Clones were examined from 2 independent transformations using both the *ΔRAD51::NTAP* and *ΔRAD51::CTAP* constructs. Two clones were examined for each independent transformation and PCR analysis was carried out as described for the procyclic form cells. The resulting PCR products are displayed in figure 6.9, and demonstrate that all of the clones examined appeared to have correctly integrated the constructs.
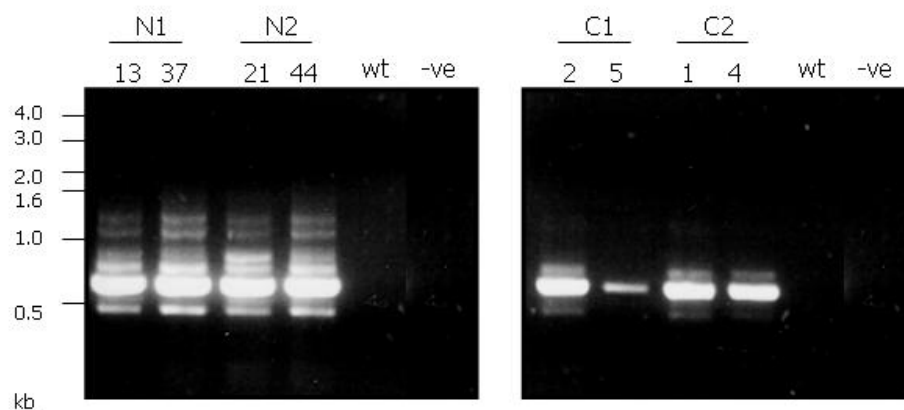
Figure 6.9 – PCR analysis of TAP tagged RAD51 transformants in bloodstream stage clones. The ethidium stained gel shown on the left displays PCR amplification using the primer pairs *NanalTAP5'* and *NanalTAP3'* to analyse N terminal TAP tagged RAD51. The ethidium stained gel shown on the right displays PCR amplification using the primer pairs *CanalTAP5'* and *CanalTAP3'* to analyse C terminal TAP tagged RAD51. The PCR products are displayed from 4 clonal G418 resistant populations (N1-13, 37 and N2-21, 44); 4 clonal blasticidin resistant populations (C1-2, 5 and C2-1, 4); wild type cells (wt) and a no DNA control (-ve). Reactions were carried out using the primer pairs depicted in figure 6.7. DNA size markers are indicated in kbp.

6.3.2 Confirmation of TAP tagged RAD51 by Western blot

Following the initial confirmation of TAP tagged transformants by PCR, western blot analysis was carried out in order to confirm if the modified versions of RAD51 were expressed. Western blot analysis was first carried out on total protein extracted from the polyclonal RAD51-TAP tagged cell lines generated in procyclic form cells. Cell extracts were separated on 10 % SDS-PAGE gels and probed with polyclonal anti-RAD51 antiserum and detected with HRP-coupled anti-rabbit IgG (figure 6.10A). These blots were subsequently stripped and re-probed with peroxidase-anti peroxidase (PAP – Sigma, P1291), which detects the protein A component of the TAP tag (figure 6.10B).
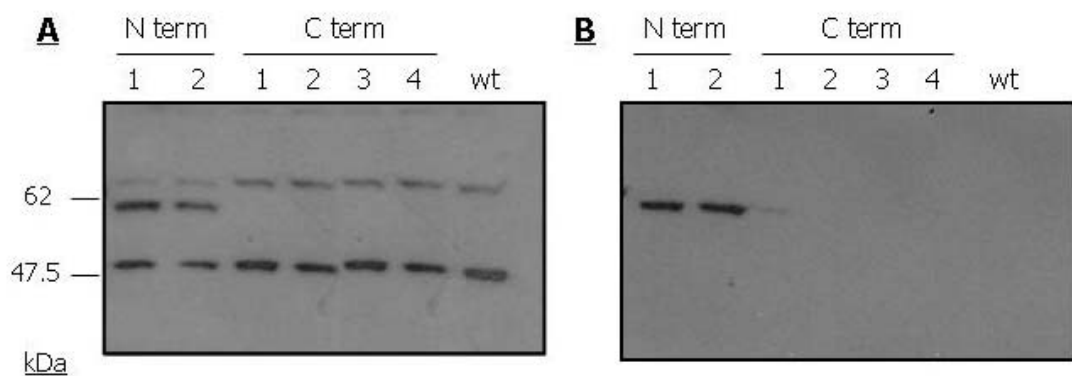
Figure 6.10 – Western blots of putative RAD51 TAP tagged polyclonal procyclic form cells. The western blots display total protein extracts probed with (A) anti-RAD51 antiserum and (B) peroxidase anti-peroxidase (PAP). The endogenous copy of RAD51 is visible at 47kDa, whilst the TAP tagged copy of RAD51 is visible at 62kDa. Size markers are indicated (kDa). N term and C term correspond to the N- or C-terminal TAP tagged versions of RAD51.

The western blot analysis displayed in figure 6.10 confirmed that the *ΔRAD51::NTAP* construct had correctly integrated into the 5' end of *RAD51*. This is established by the procyclic cell lines N term 1 and 2 possessing not only the endogenous copy of RAD51 (seen at ~ 47 kDa), but also an additional copy of RAD51 at around 60 kDa. This copy of RAD51 was assumed to be RAD51 plus a TAP tag, since the TAP tag adds approximately 15 kDa extra to the protein. This assumption was confirmed when the western blot was re-probed with PAP, as this blot displayed bands that were the same size of the higher molecular weight bands in the RAD51 blot.

Transformants obtained from the *ΔRAD51::CTAP* construct, however, displayed only endogenous copies of RAD51 when examined by western blot analysis. This result confirms that the C term 1 and 2 transformants had not correctly integrated the construct. In addition, it contradicts the results from the PCR analysis for C term 3 and 4 transformants. No additional copy of RAD51 could be observed, and when this blot was

re-probed with PAP, no bands were detected. However, on overnight exposure faint bands could be detected with PAP at approximately the expected size (62 kDa) (data not shown). This result could be explained by the fact that these analyses were performed on polyclonal populations. PCR fragments of the expected sizes would be obtained even if only a small percentage of cells in the polyclonal population had correctly integrated the construct. Detection of the protein, however, would be more difficult, explaining the absence of the expected bands. As it has already been noted that the C terminus of RAD51 appears to be buried within the nucleoprotein filament, it is possible that addition of the TAP epitope to this end of the protein could cause a growth disadvantage over other integrations.

To examine this further, the polyclonal procyclic form transformants N1, N2, C3 and C4 were cloned, following the protocol described in section 2.1.3.2. A number of clones were obtained from the *ΔRAD51::NTAP* transformants and were confirmed as having correctly integrated the construct by PCR analysis (data not shown). Despite numerous attempts to generate clones from the *ΔRAD51::CTAP* transformants, only a single clone was generated, and this was also confirmed as having correctly integrated the construct by PCR analysis (data not shown). Whole cell extracts were then prepared from each of these clones before examining protein expression as described above. The results are displayed in figure 6.11, and confirm the generation of 4 RAD51 N-terminal TAP tagged clones and 1 RAD51 C-terminal TAP tagged clone. One of the N-terminal tagged clones, N2-1, may appear unusual, in that the intensity of the binding of RAD51 antiserum was not equivalent between the endogenous and tagged alleles. Whether this is a blotting artefact, or reflects difference in expression or translation is unclear.
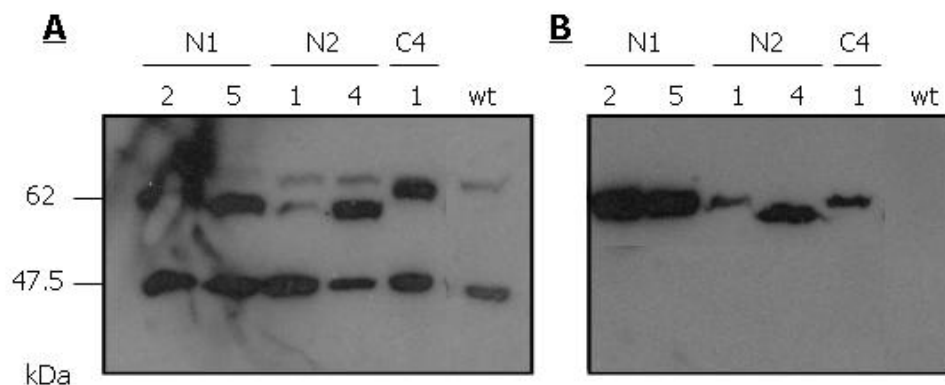
Figure 6.11 – Western blots of RAD51 TAP tagged procyclic clonal cell lines. The western blots display total protein extracts probed with (A) anti-RAD51 antiserum and (B) peroxidase anti-peroxidase (PAP). The endogenous copy of RAD51 is visible at 47kDa, whilst the TAP tagged copy of RAD51 is visible at 62kDa. Size markers are indicated (kDa). N term and C term correspond to the N- or C-terminal TAP tagged variants of RAD51.

Western blot analysis was similarly performed on total protein extracted from the cloned RAD51-TAP tagged cell lines generated in bloodstream stage cells. The resulting western blot is displayed in figure 6.12 and confirms the results from the PCR analysis. Four clones were obtained in which the TAP tag had been fused with the N-terminus of RAD51, but no C-terminally TAP tagged clones were recovered. As in procyclic form cells, the C-terminal tagging of RAD51 appeared to be selected against. The integration point of the *ΔRAD51::CTAP* constructs was not investigated further. Given the overall difficulty in adding a C-terminal TAP tag, the functionality of the single RAD51 C-terminal TAP tagged clone generated in procyclic form cells must be considered suspect. It was therefore decided to continue examination with only the N-terminal TAP tagged transformants.
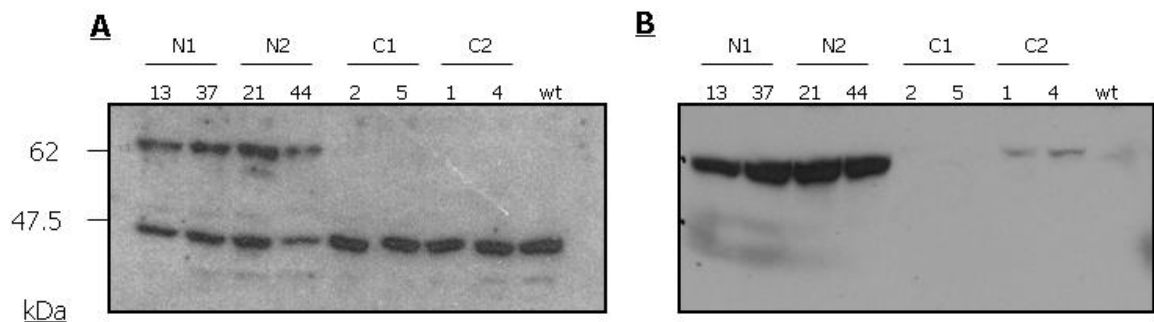
Figure 6.12 – Western blots of RAD51 TAP tagged bloodstream stage cell lines. The western blots display total protein extracts probed with (A) anti-RAD51 antiserum and (B) peroxidase anti-peroxidase (PAP). The endogenous copy of RAD51 is visible at 47kDa, whilst the TAP tagged copy of RAD51 is visible at 62kDa. Size markers are indicated (kDa). N and C correspond to the N or C terminal TAP tag of RAD51.

6.3.3 Confirmation of TAP tagged RAD51 by Southern analysis

To confirm the generation of N-terminal RAD51 TAP tagged transformants in both bloodstream stage and procyclic form cells, Southern analysis was carried out on genomic DNA and compared with wild type parental DNA. Approximately 5 μg of genomic DNA from each cell line was restriction digested with *Eco*RI and *Pst*I overnight before being electrophoresed on a 0.8 % agarose gel and Southern blotted. The blots were probed with a region of the *RAD51* ORF depicted in figure 6.13, where the expected size fragments of the transformants are also shown.
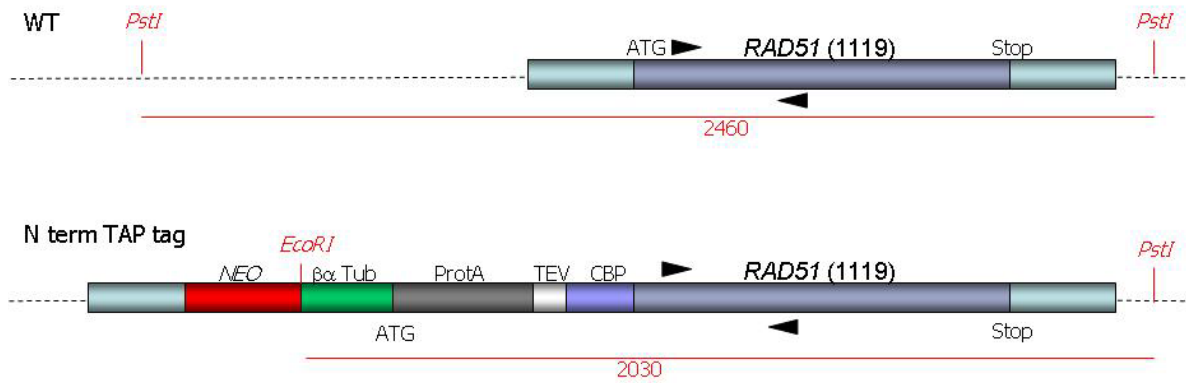
Figure 6.13 – Strategy for confirming TAP tagged cell lines by Southern analyses. The upper diagram depicts the wild type *RAD51* ORF and predicted fragment size (in bp) when digested with *Eco*RI and *Pst*I. The lower diagram depicts the ORF of *RAD51* when the construct *ΔRAD51::NTAP* had correctly integrated. The fragment size (in bp) when digested with *Eco*RI and *Pst*I is displayed. The black triangles represent primers used to generate the *RAD51* probe.

The Southern blots in figure 6.14 confirm that the construct *ΔRAD51::NTAP* had correctly integrated into the bloodstream stage clones N1-13, N1-37, N2-21 and N2-44, and in the procyclic form clones N1-2, N1-5, N2-1 and N2-4. In all cases, the blots show that when genomic DNA from wild type and transformant cells was digested with *Eco*RI and *Pst*I, a hybridising band corresponding to the wild type copy of *RAD51* was observed. For the transformants, an extra hybridising band was observed which corresponded to the TAP tagged copy of *RAD51*, as predicted in figure 6.13.
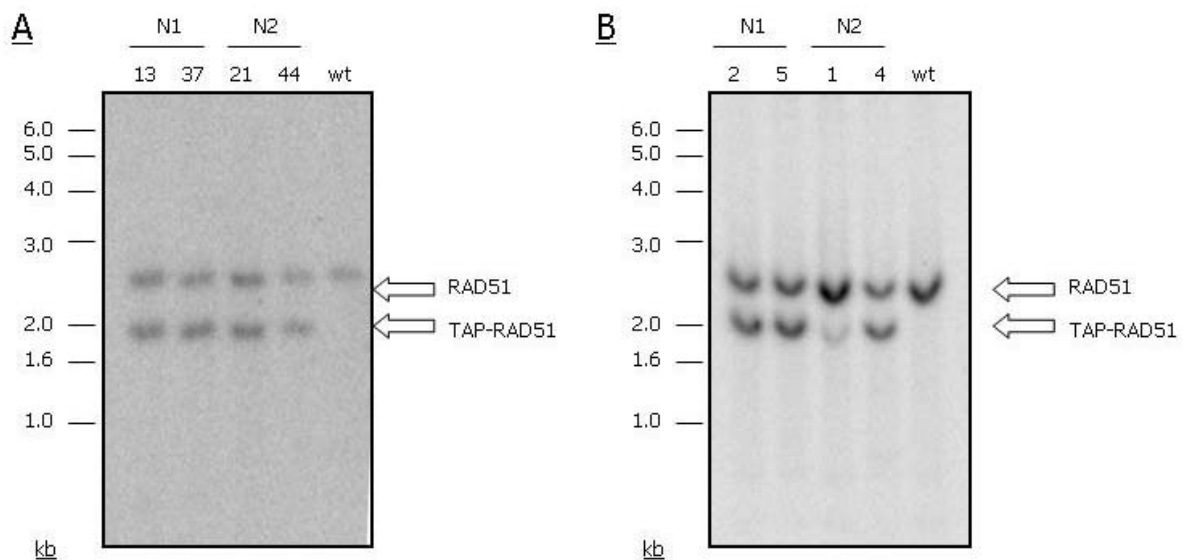
Figure 6.14 – Confirmation of RAD51 N terminally TAP-tagged transformants by Southern analysis. (A) Bloodstream cell lines and (B) procyclic cell lines were digested with *Eco*RI and *Pst*I. 5 μg of genomic DNA of each cell line was restriction digested for 12 hours before being run out on a 0.8 % agarose gel. The DNA was Southern blotted before being probed with the RAD51 open reading frame. N1 and N2 refer to the two independent N terminal transformants; WT refers to genomic DNA from untransformed cell lines. Clone numbers are indicated, as are size markers (kbp).

6.4 Generation of RAD51 heterozygous mutants in the TAP tagged cell lines

In order to determine whether or not TAP-tagged RAD51 was capable of functioning normally during DNA repair, we wanted to disrupt the unaltered copy of RAD51 in the clonal transformants and examine the ability of these mutants to form sub-nuclear RAD51 foci after DNA damage, and their sensitivity to DNA damaging agents. These parameters provide a good indication of the function of the tagged RAD51 protein since we know that *rad51-/-* mutants are unable to form foci (McCulloch, unpublished) and are sensitive to DNA damaging agents (McCulloch and Barry, 1999).
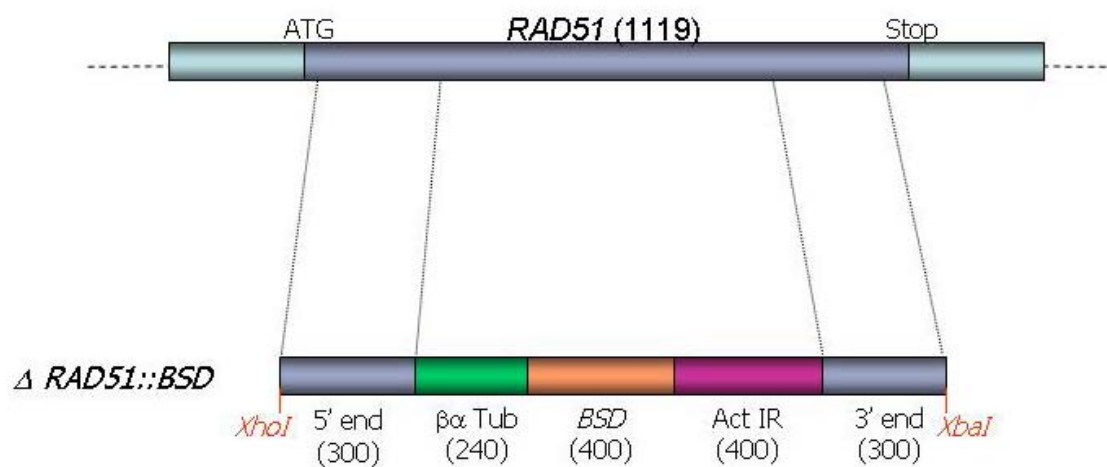
Figure 6.15 – *RAD51* gene disruption strategy. Restriction map of the construct used for disruption of *RAD51* is shown, relative to the *RAD51* ORF. The *Δ RAD51::BSD* construct was cloned into the pBC SK plasmid. 5' flank and 3' flank correspond to the homologous regions of the start and end of the RAD51 ORF. αβ Tub: αβ tubulin intergenic region (processing signal). ACT IR: Actin intergenic region (processing signal. BSD: blasticidin resistance gene ORF. Sizes of the individual components are shown in base pairs. The construct allows homologous recombination to disrupt the *RAD51* ORF.

Figure 6.15 displays the method utilised to mutate a copy of *RAD51*. In this strategy, the entire ORF was not deleted but was instead disrupted, and this method has previously proved to be successful in generating *rad51-/-*, *rad51-3-/-* and *rad51-5-/-* mutants (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005). The construct *Δ RAD51::BSD* (R. Barnes, gift) was generated by cloning 5' and 3' ends of the *RAD51* ORF into pBluescript SK, and subsequently cloning an antibiotic resistance cassette for blasticidin between these sequences. The 5' and 3' ends of the ORF allow homologous recombination following transformation, replacing the core domain of RAD51, including

the highly conserved walker A and B boxes that are needed for ATP binding and hydrolysis, with an antibiotic cassette and tubulin and actin intergenic sequences.

Two of the N-terminal TAP tagged bloodstream stage (N1-13 and N2-21) and procyclic form (N1-5 and N2-4) clones were transformed with the *Δ RAD51::BSD* construct. Antibiotic resistant transformants were selected by placing cells on either 5 μg.ml^{-1} (bloodstream) or 10 μg.ml^{-1} (procyclic) blasticidin, and the generation of heterozygous mutants was confirmed by PCR and western blot analysis (section 6.4.1, below).

6.4.1 Confirmation of heterozygous RAD51+/- mutants that retain only the TAP tagged copy of RAD51

Since the *Δ RAD51::BSD* construct could integrate into either the endogenous or the TAP tagged allele of *RAD51*, it was necessary to determine that the putative *RAD51* heterozygous mutants had retained the TAP tagged copy of *RAD51*, disrupting the endogenous allele. For this, PCR analysis was carried out using the method depicted in figure 6.16, which allows integration of the construct into each allele of *RAD51* to be differentiated by PCR product size. Oligonucleotide primers were designed, one of which (*Outside RAD51*) contained 19 bases of sequence that was homologous to a region upstream of the *RAD51* ORF and a second (*BSD 3'*), which contained 21 bases of sequence that was complementary to a region of the blasticidin ORF. If the construct integrated into the endogenous copy of *RAD51*, then a DNA fragment of approximately 1.1 kb would be amplified; integration into the N terminal TAP-tagged copy of *RAD51* would lead to PCR-amplification of a DNA fragment of approximately 2.8 kb.
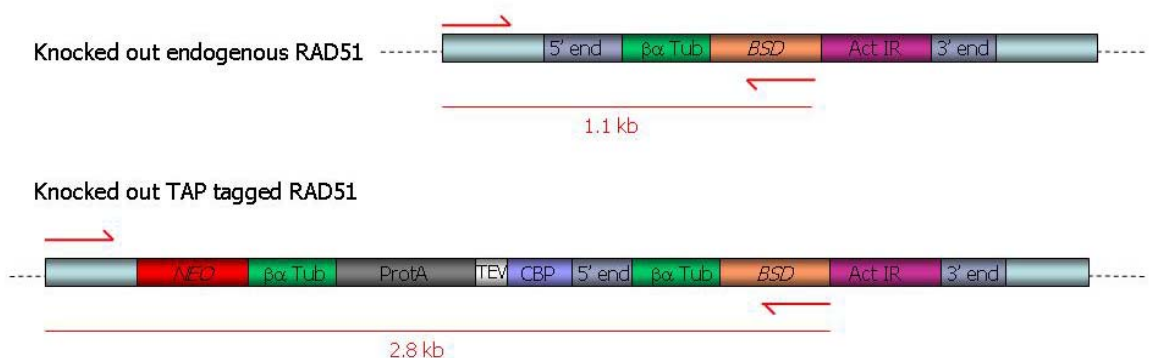
Figure 6.16 – Analysis of *RAD51* gene disruption by PCR. The *RAD51::BSD* gene disruption construct can integrate into either the endogenous or the TAP tagged allele of *RAD51*. To determine which allele the construct had integrated into, PCRs were performed. The 5' primer (*Outside RAD51*) designed was homologous to a region upstream of the *RAD51* ORF and the 3' primer (*BSD 3'*) deigned was complementary to a region in the *BSD* resistance cassette. Different size products (shown in kbp) are generated depending on which copy the construct had integrated into.

Genomic DNA was prepared from 9 blasticidin resistant bloodstream stage clones and 20

blasticidin resistant procyclic form clones. PCR analysis was performed using the primer

pairs described in figure 6.16 and Taq DNA polymerase. Three bloodstream stage (TAP 3

from N1-13 and TAP 5 and 6 from N2-21) and three procyclic form clones (1-1 and 1-5

from N1-5 and 3-3 from N2-4) were shown to have disrupted the endogenous, untagged

copy of *RAD51* (figure 6.17), whilst 6 bloodstream stage clones and 2 procyclic form

clones were found to have integrated the construct into the TAP tagged copy of *RAD51*

(data not shown). The majority of the procyclic form clones, although resistant to

blasticidin, did not yield PCR products of the expected size, suggesting integration of the

ΔRAD51::BSD construct elsewhere in the genome (data not shown). The location of this

aberrant integration, and the potential significance of such a high incidence in comparison

with bloodstream stage cells, was not investigated further.
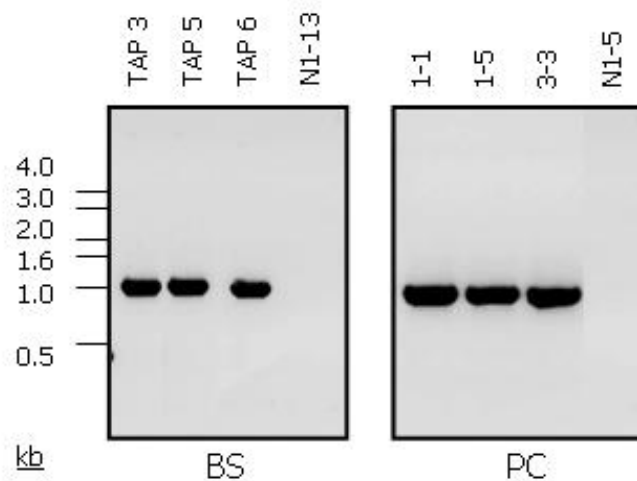
Figure 6.17 – Confirming *RAD51* gene disruption by PCR. The ethidium stained gel shows PCR analysis from RAD51-TAP tagged transformants (N1-13 and N1-5) and putative *RAD51+/-* clones. The clones obtained in both bloodstream (BS) and procyclic (PC) life cycle stages had disrupted the endogenous, untagged allele of *RAD51*, leaving the TAP-tagged allele remaining. Sizes are indicated in kbp.

The results from the PCR analysis were further confirmed by western blot analysis. Whole

cell extracts were prepared from each of the clones, and separated on 10 % SDS PAGE

gels and probed with anti-RAD51 antiserum and detected with HRP-coupled anti-rabbit

IgG as before. The results are displayed in figure 6.18, and confirm the results found from

the PCR analysis. All of the clones were found to be expressing only the TAP tagged copy

of RAD51.
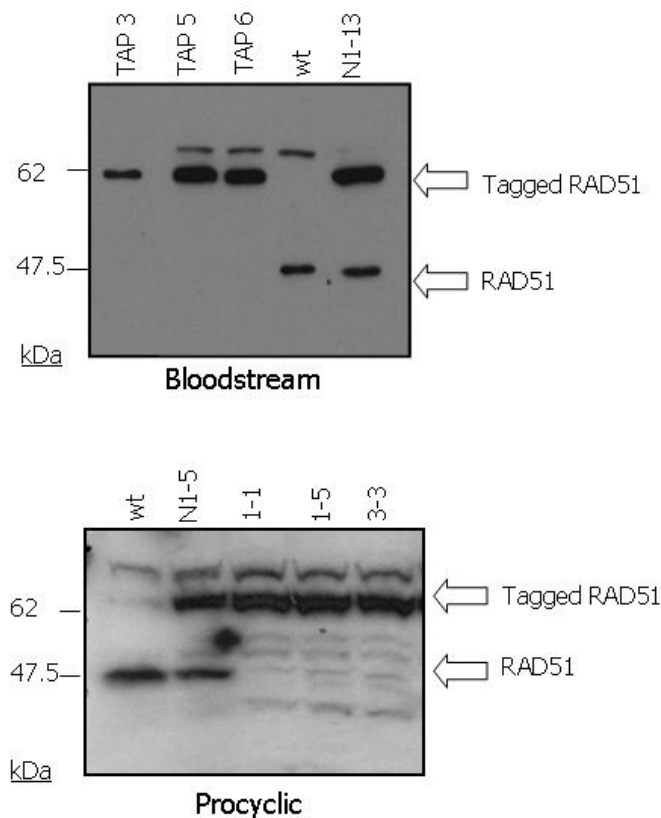
Figure 6.18 – Confirming *RAD51* gene disruption by western blot analysis. The western blots display total protein extracts from wild type cells (wt), N-terminal TAP tagged RAD51 clones (N1-13 and N1-5) and *RAD51+/-* clones, ontained in both bloodstream and procyclic life cycle stages, probed with anti-RAD51 antiserum. The endogenous copy of RAD51 is visible at 47kDa, whilst the TAP tagged copy of RAD51 is visible at 62kDa. Sizes are indicated in kDa.

6.5 Phenotypic analysis

In order to examine if the TAP tag had an adverse effect on the functionality of RAD51, specific phenotypes of the *RAD51+/-* mutants that retained only the TAP-tagged copy of RAD51 were examined. Previous work had demonstrated that *rad51-/-* mutants were impaired in their growth, were sensitive to MMS and were incapable of forming RAD51 foci after induced DNA damage, and so each of these were investigated and compared with wild type cells.

6.5.1 Analysis of in vitro growth

In vitro growth analysis was carried out on the TAP-tagged *RAD51+/-* cell lines in bloodstream and procyclic stages and compared with the wild type cells. Bloodstream stage cultures were inoculated at a cell density of 5×10^4 cells.ml^{-1}, whilst procyclic form cultures were inoculated at a cell density of 5×10^5 cells.ml^{-1}. Cell concentrations were counted using a haemocytometer (Bright-line, Sigma) at 24, 48, 72 and 96 hours

subsequently. Three repetitions of each growth experiment, for all cell lines, were carried
out and the results plotted on a semi-logarithmic scale (figures 6.19 and 6.20).
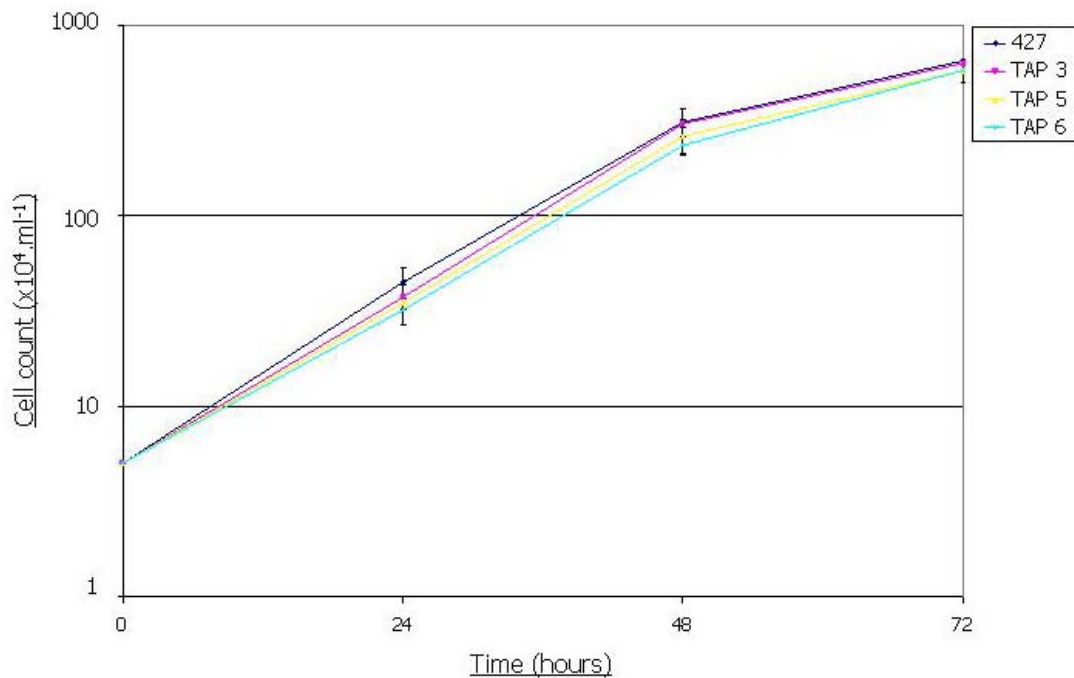
**Figure 6.19 – Analysis of *in vitro* growth of the bloodstream stage TAP-tagged *RAD51+/-*
mutants.** 5 ml cultures were set up at 5 x 10^4 cells.ml^{-1} and cell densities counted 24, 48, 72 and
96 hours subsequently. Standard errors are indicated for the counts using data from three
repetitions. 427: wild type *T. brucei* cells; TAP 3, TAP 5 and TAP 6: *RAD51* heterozygote clones
with only the TAP tagged allele of *RAD51* remaining.

The results shown in figure 6.19 demonstrate that the TAP-tagged variant of RAD51 in the
bloodstream stage of *T. brucei* was capable of supporting *in vitro* growth, as the population
doubling times of the TAP-tagged *RAD51+/-* mutants (8.34, 8.54 and 8.72 for TAP 3, TAP
5 and TAP 6, respectively), were essentially equivalent to that of the wild type bloodstream
stage cells (8.19 hours). Figure 6.20 demonstrates this was also true in the procyclic stage
of the life cycle. Again, population doubling times of the TAP-tagged *RAD51+/-* cells
(25.25, 25.05 and 25.25) appeared to be comparable with the 24.03 hours measured for the
wild type cells.
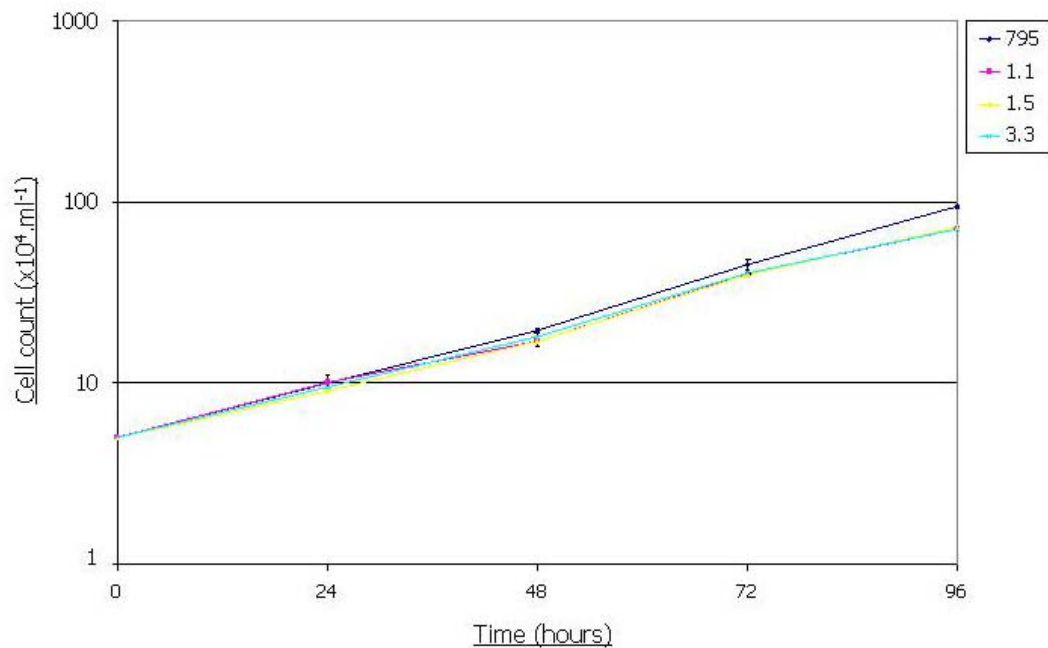
Figure 6.20 – Analysis of *in vitro* growth of procyclic form TAP-tagged *RAD51+/-* mutants. 5 ml cultures were set up at 5 x 10^5 cells.ml^{-1} and cell densities counted 24, 48, 72 and 96 hours subsequently. Standard errors are indicated for the counts using data from three repetitions. 795: wild type *T. brucei* cells; 1.1, 1.5 and 3.3: *RAD51* heterozygote clones with only the TAP tagged allele of *RAD51* remaining.

6.5.2 Analysis of DNA damage sensitivity

To analyse DNA damage sensitivities, the Alamar Blue assay, using MMS as the DNA damaging agent, was carried out for both the bloodstream stage and procyclic form cells following the same protocol as described in section 4.3.4. The only notable difference between the two life cycle stages was that the procyclic form cells were placed in conditioned media (as described in 2.1.3.2). Three repetitions were performed for all TAP-tagged *RAD51+/-* and wild type cells, the IC50s calculated and their means plotted graphically (Figures 6.21 and 6.22).

These data demonstrate that the addition of a TAP tag to the N terminus of RAD51 had no effect, in either life cycle stage, on the function of the protein in response to DNA damage repair. The mean IC50s of the TAP-tagged *RAD51+/-* cells in the bloodstream stage (0.0011 %, 0.0012 % and 0.0012 % for TAP 3, 5 and 6 respectively) were essentially equivalent compared with 0.0012 % in wild type cells. Indeed, these IC50s are significantly higher than previous results found for *rad51-/-* mutants, which displayed a mean IC50 of 0.0007 % MMS (section 4.3.4).
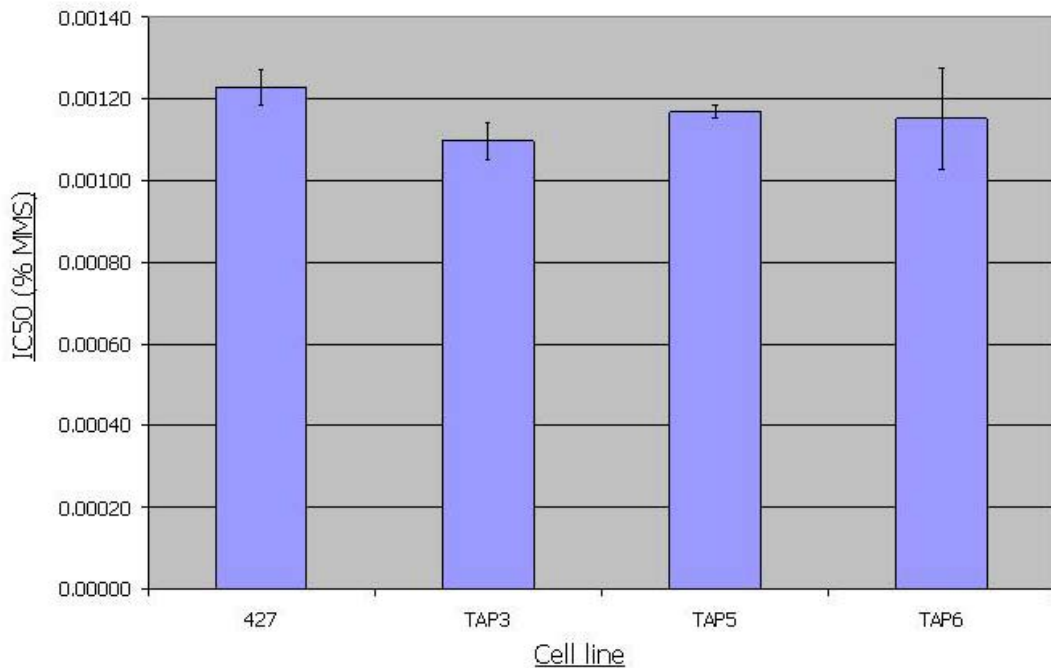
Figure 6.21 – IC50s of *T. brucei* TAP-tagged *RAD51+/-* bloodstream stage mutants exposed to MMS. Wild type and *RAD51+/-* cell lines were placed in serially decreasing amounts of MMS and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments; bars indicate standard error.
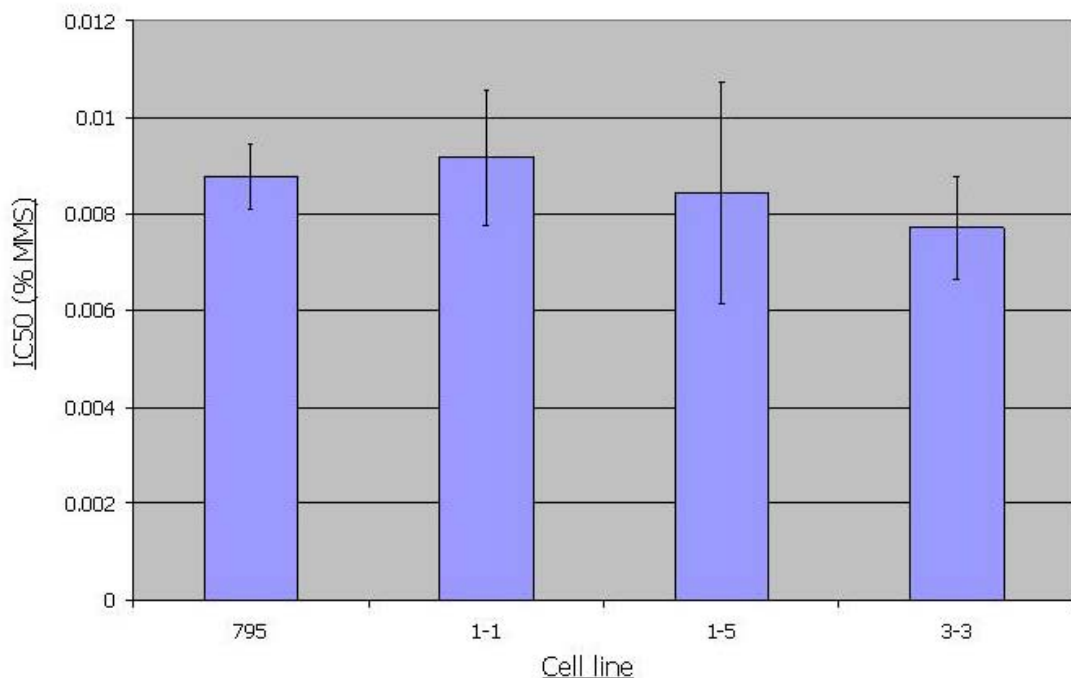
Figure 6.22 – IC50 of *T. brucei* TAP-tagged *RAD51 +/-* procyclic form mutants exposed to MMS. Wild type and *RAD51+/-* cell lines were placed in serially decreasing amounts of MMS and allowed to grow for 48 hours, before the addition of Alamar Blue. After a further 24 hours, the reduction of Alamar Blue was measured by the amount of fluorescent resorufin generated. Values are the mean IC50s from 3 experiments; bars indicate standard error.

In the procyclic stage of the life cycle, the cells appear to be intrinsically more resistant to MMS, as the wild types had IC50s of ~0.009 %, around 8-9 fold higher than WT bloodstream stage cells. Why this should be the case is unclear. Nevertheless, the mean IC50s of the TAP-tagged *RAD51+/-* cells were 0.009 %, 0.008 % and 0.009 % for 1-1, 1-5 and 3-3, respectively, which were equivalent to the wild type cells.

6.5.3 Analysis of the ability to form RAD51 foci

As a final route to determine if the addition of a TAP tag to the N terminus of RAD51 interferes with the function of the protein, the ability of the TAP-tagged *RAD51+/- T. brucei* to form RAD51 sub-nuclear foci following DNA damage was examined. Cell lines were treated with phleomycin for 18 hours and RAD51 localisation examined by indirect immunofluorescence, as described in section 4.3.6. Approximately 200 cells were then counted and scored for the number of foci they contained. The results for the bloodstream stage cells are displayed in table 6.1, and examples of these foci are shown in figure 6.23. For procyclic form cells, the results are tabulated in table 6.2 and examples of foci are shown in figure 6.24.

		Number of foci (%)						
	BLE	0	1	2	3	4	5	6 or more
WT	0.0	96.4	3.6	0.0	0.0	0.0	0.0	0.0
	1.0	24.8	22.6	18.8	16.5	13.5	2.3	1.5
TAP3	0.0	94.8	3.5	1.7	0.0	0.0	0.0	0.0
	1.0	36.1	28.2	23.3	6.2	4.0	1.3	0.9
TAP5	0.0	96.7	1.9	1.4	0.0	0.0	0.0	0.0
	1.0	33.2	33.2	23.0	5.5	2.3	1.8	0.9
TAP6	0.0	96.2	2.8	0.9	0.0	0.0	0.0	0.0
	1.0	36.0	32.6	25.3	5.1	0.0	1.1	0.0

Table 6.1 – RAD51 foci formation in wild type cells and TAP-tagged *RAD51+/-* bloodstream stage mutants. The percentages of cells showing foci at given concentrations of phleomycin (BLE) are shown. Phleomycin concentrations are shown in $\mu g.ml^{-1}$. Boxes without shading contain no foci, boxes shaded in light yellow contain foci and boxes shaded in bright yellow contain the highest percentage of foci.

As described previously (section 4.5.5), without the presence of damage RAD51 foci were rarely seen in all cell lines. However, once damage was induced, the majority of cells were found to contain one or more foci, and the TAP-tagged *RAD51+/-* mutants were found to form RAD51 foci at an approximately comparative level to wild type cells, with the percentage of cells containing foci ranging from 64-67 % compared to 75 % for wild type cells.
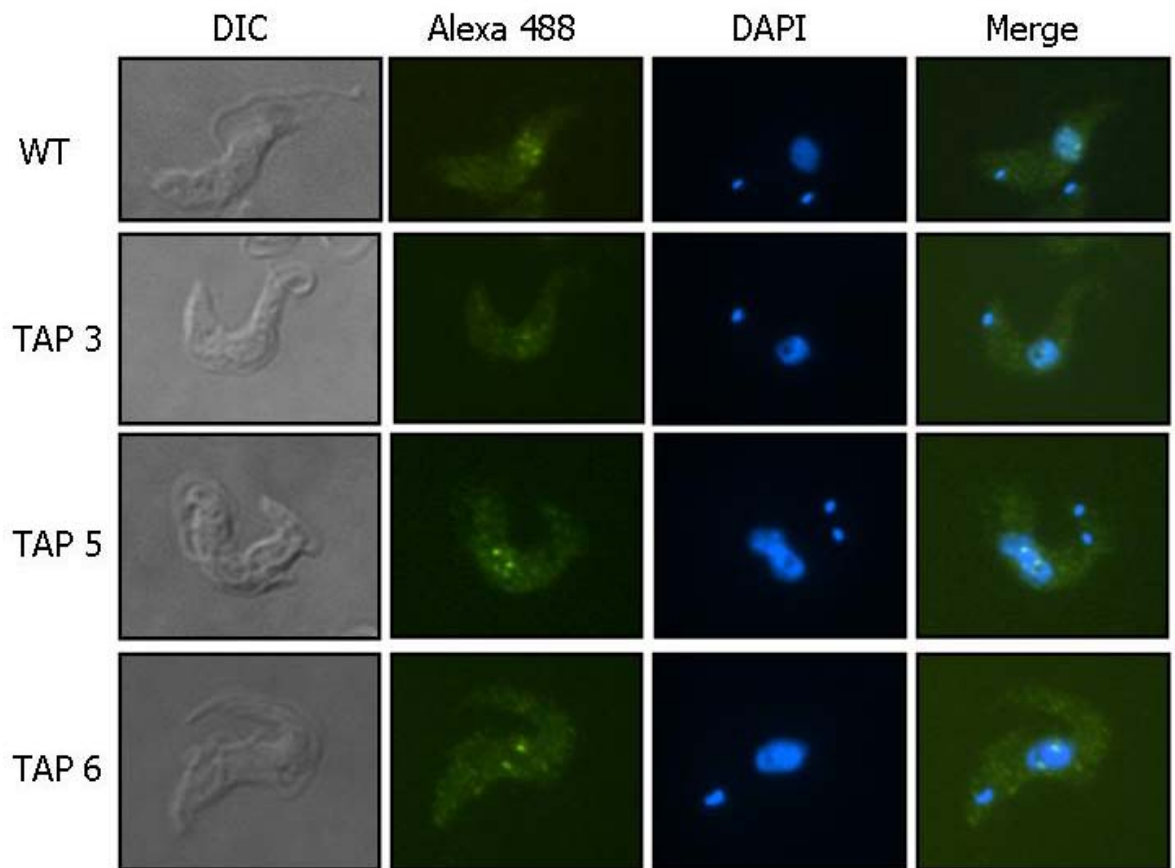
Figure 6.23 – RAD51 foci formation in TAP-tagged *T. brucei RAD51 +/-* bloodstream stage mutants. Each cell is shown in phase contrast (phase), after staining with DAPI and after hybridisation with anti-RAD51 antiserum and secondary hybridisation with Alexa Fluor 488 conjugate (Alexa 488). Merged images of DAPI and Alexa 488 cells are also shown. Wild type Lister 427 (WT) cells and *RAD51+/-* mutants with just the TAP tagged allele of RAD51 are shown (TAP 3, 5 and 6).

	BLE	Number of foci (%)						
		0	1	2	3	4	5	6 or more
WT	0.0	93.6	1.8	1.8	1.3	0.0	0.0	1.8
	1.0	27.0	8.7	7.9	7.9	10.3	12.7	25.4
1-1	0.0	92.0	6.4	1.6	0.0	0.0	0.0	0.0
	1.0	32.7	25.1	21.6	9.0	7.5	2.5	1.5
1-5	0.0	90.4	7.6	2.0	0.0	0.0	0.0	0.0
	1.0	31.5	23.4	20.3	8.1	8.1	6.1	2.5
3-3	0.0	89.3	7.7	2.0	0.5	0.0	0.5	0.0
	1.0	32.3	22.4	23.4	7.5	9.0	4.5	1.0

Table 6.2 – RAD51 foci formation in wild type cells and *RAD51+/-* procyclic form mutants. The percentages of cells showing foci at given concentrations of phleomycin (BLE) are shown. Phleomycin concentrations are shown in $\mu g.ml^{-1}$. Boxes without shading contain no foci, boxes shaded in light yellow contain foci and boxes shaded in bright yellow contain the highest percentage of foci.

RAD51 foci formation in procyclic form cells has not been examined previously, but appears to be highly comparable with bloodstream stage cells as the procyclic form TAP-tagged *RAD51+/-* cells also showed no evidence for RAD51 foci impairment. Again, without damage, the majority of cells were found to contain no foci (89-94 %). Equally, the number of WT cells that induce foci following equivalent phleomycin treatment was

comparable with the bloodstream stage (73 %). Once damage was induced, the majority of procyclic form TAP-tagged *RAD51+/-* cells also contained one or more foci, with the percentage of cells containing foci ranging from 67-69 %.
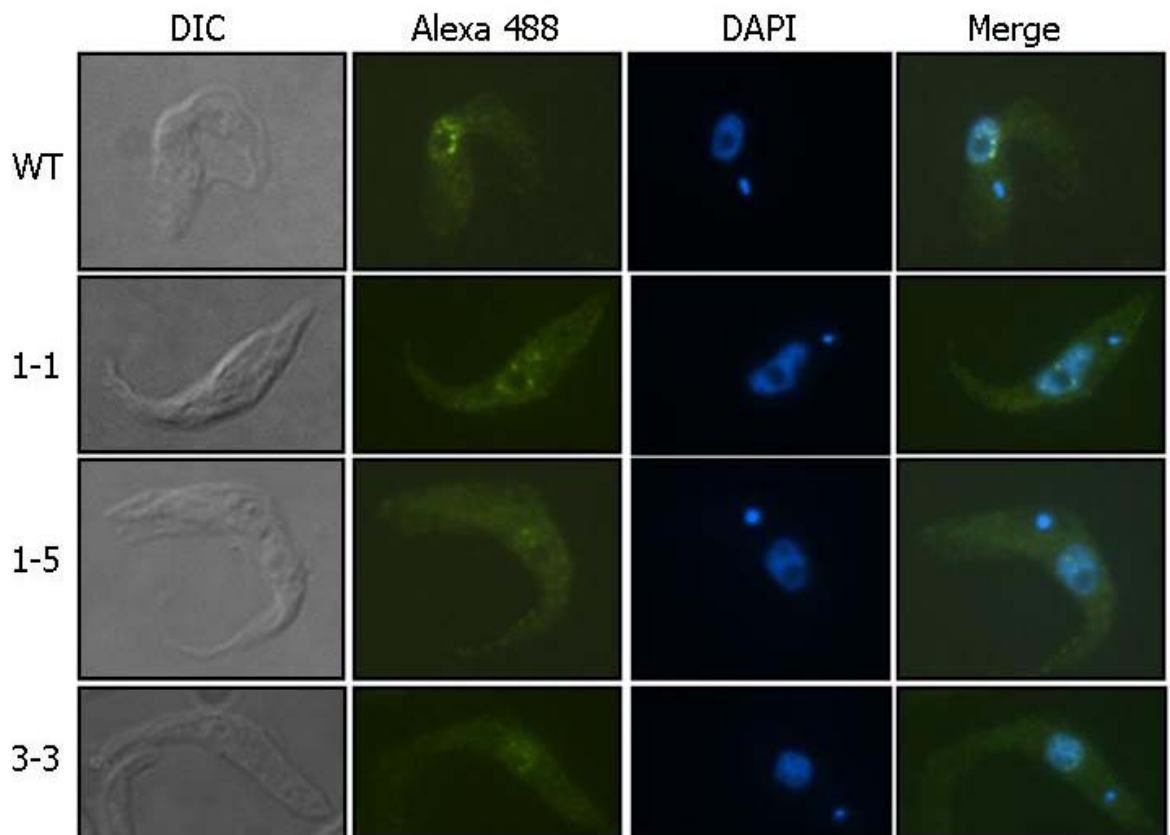
Figure 6.24 – RAD51 foci formation in TAP-tagged *T. brucei RAD51 +/-* procyclic form mutants. Each cell is shown in phase contrast (phase), after staining with DAPI and after hybridisation with anti-RAD51 antiserum and secondary hybridisation with Alexa Fluor 488 conjugate (Alexa 488). Merged images of DAPI and Alexa 488 cells are also shown. Wild type EATRO 795 (WT) cells and *RAD51+/-* mutants with just the TAP tagged allele of RAD51 are shown.

These results indicate that the addition of an N-terminal TAP tag does not interfere with RAD51's ability to reorganise its sub-cellular location into discrete sub-nuclear foci upon DNA damage in either the bloodstream or procyclic stages of the *T. brucei* life cycle. Taken in conjunction with the results from MMS sensitivity and *in vitro* growth rate assays, it appears that the TAP tagged variant of RAD51 functions in DNA repair to an extent comparable with the endogenous protein. From this, it leads to the assumption, not tested, that the TAP-tagged variant of RAD51 should be able to interact with putative RAD51-interacting partners, and that the TAP method should provide a legitimate approach to addressing this question. Of course, it is also possible that the TAP-tagged variant of RAD51, although able to function normally in these experiments, may not interact with all of the RAD51 partners, but only those sufficient for the above phenotypes that have been examined.

6.6 Attempts at RAD51 tandem affinity purification

Initial attempts at optimisation of the TAP purification protocol was carried out in the procyclic form transformants for the reasons already discussed (difficulties in growing sufficiently large numbers of bloodstream stage *T. brucei* cells). Two independent procyclic transformants (N1-5 and N2-4) were grown to a density of approximately 1×10^7 cells.ml^{-1}, before cell extracts were prepared following the protocol described by Puig *et al.*, 2001. To do this, 3×10^9 cells were harvested by centrifugation at 1600 g and washed twice in PBS before resuspending in 10 mls of IPP150 buffer (10 mM Tris-Cl (pH 8.0), 150 mM NaCl, 0.1 % NP40). 1 % Triton X-100 was then added and the cells incubated on ice for 15 minutes. Following this, the preparation was centrifuged at 10 000 g for 15 min at 4 °C (Beckman JA21 rotor) and the supernatant, representing the cell extract, subjected to affinity purification using IgG sepharose beads. This affinity purification step was achieved by placing 200 µl IgG sepharose beads (Amersham) into a 10 ml disposable polypropylene column (Pierce) and washing with 5 mls IPP150. The supernatant was subsequently placed into this column and rotated for 2 hours at 4 °C.

Following the IgG incubation period, unbound material was removed by washing the beads with 30 mls IPP150 and 10 mls TEV cleavage buffer (10 mM Tris-Cl (pH 8.0), 150 mM NaCl, 0.1 % NP40, 0.5 mM EDTA, 1 mM DTT). The IgG beads, and putative bound complexes, were then resuspended in 1 ml TEV cleavage buffer before being treated with 300 units of TEV protease (Invitrogen) for 3 hours, with rotation, at room temperature. Material that was eluted from the IgG column by TEV cleavage was recovered by gravity flow and then subjected to affinity purification using calmodulin beads. This step was performed by adding 3 volumes of calmodulin binding buffer (10 mM β–mercaptoethanol, 10 mM Tris-Cl (pH 8.0), 150 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 2 mM CaCl$_2$, 0.1 % NP40) and 3 mM CaCl$_2$ to the material eluted from the IgG beads, and rotating in the calmodulin column for 1 hour at 4 °C. The calmodulin column was prepared in the same manner as the IgG column, but using 200 µl calmodulin beads (Amersham) washed with 5 mls calmodulin binding buffer. Unbound material was removed by gravity flow and washing with 30 mls calmodulin binding buffer. Finally, five fractions of 200 µl were eluted by gravity flow using calmodulin elution buffer (10 mM β–mercaptoethanol, 10 mM Tris-Cl (pH 8.0), 150 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 2 mM EGTA, 0.1 % NP40).

To follow the purification, material from each stage of the TAP protocol was loaded onto 10 % SDS-PAGE gels and stained with Coomassie (figure 6.25). The two independent

samples looked very comparable up to the stage representing the flow through from the calmodulin column (lane E). In each case the eluate from the IgG column produced a distinct band at ~80 kDa and 2 smaller bands of ~30 and 32 kDa, which, given their size and appearance, were presumed to be the TEV protease. The identity of 80 kDa band is unknown, but it is notable that no discrete band of the size expected for RAD51 (~ 52 kDa) was visible. The absence of RAD51 was also confirmed by western blot (data not shown).

The results from the two independent samples did not produce the equivalent results from the calmodulin elution, for reasons that are unclear. In the N1-4 sample, a ~70kDa protein was visible in both the wash and elution from the calmodulin column, but was not detected for sample N2-5. Irrespective of this difference, the eluted proteins are not likely to be meaningful in the absence of RAD51. Why this should be is not clear, but could be for a number of reasons, including the failure of RAD51 to be released by TEV cleavage or failure to recover RAD51 in the cell lysis procedure.
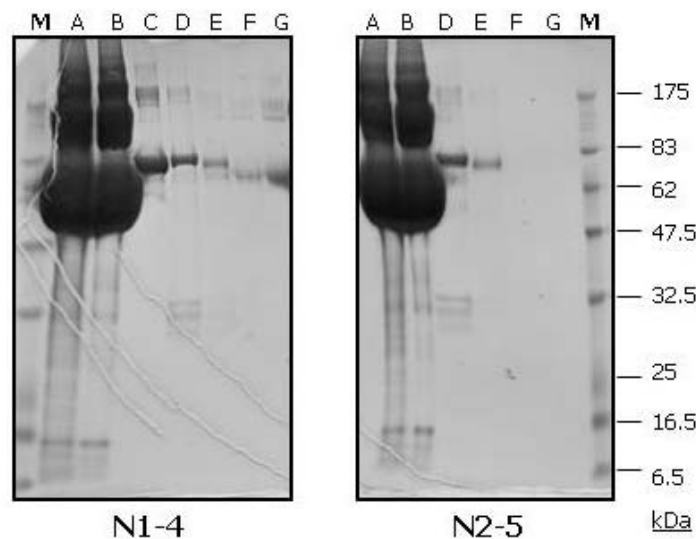
Figure 6.25 – Coomassie stained SDS-PAGE gel displaying products obtained throughout RAD51 TAP. Two independent procyclic form RAD51-TAP tagged transformants were subjected to the TAP purification procedure and samples from each step loaded onto 10 % SDS-PAGE gels. A – cell extract supernatant, B – wash with IPP150, C – wash with TEV cleavage buffer, D – eluate from IgG column, E – flow through from calmodulin column, F – wash with calmodulin binding buffer, G – eluate from calmodulin column, M – protein marker (NEB-broad range). Sizes are displayed in kDa.

Since RAD51 was not detected in any of the TAP fractions, and was notably also not detected in the supernatant from the cell extract that was subjected to affinity purification, it was decided to examine the level of RAD51 present under different cell extract preparation methods. Specifically, the lysis conditions were altered by raising the salt concentration, since it was possible that most RAD51 was bound to the DNA and was

therefore pelleted during centrifugation, excluding it from the supernatant in the previous lysis procedure.

Cell extracts were prepared on 1 x 10^9 cells using the protocol described above using 3 buffers at different salt concentrations: IPP150 (10 mM Tris-Cl (pH 8.0), 150 mM NaCl, 0.1 % NP40), IPP300 (10 mM Tris-Cl (pH 8.0), 300 mM NaCl, 0.1 % NP40) and IPP450 (10 mM Tris-Cl (pH 8.0), 450 mM NaCl, 0.1 % NP40). For each, the pellet and supernatant fractions were loaded onto 10 % SDS-PAGE gels and subjected to western blot analysis, using anti-RAD51 antiserum. Equivalent samples were also stained with Coomassie to ensure equal loading (data not shown). The results from the pellet fractions are displayed in figure 6.26.
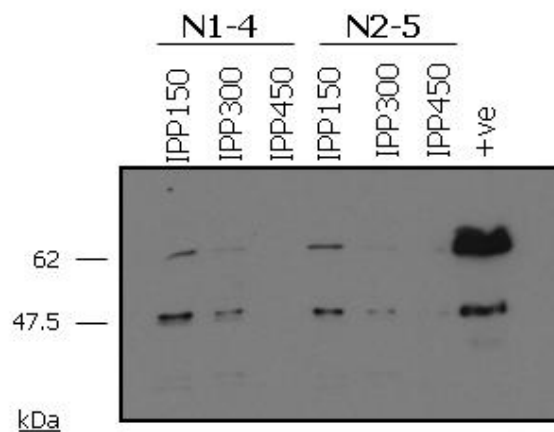
Figure 6.26 – Western blot of the pellet fractions from different lysis conditions. The western blot displays the pellet fractions of extracts prepared from two independent procyclic transformants using increasing concentrations of salt in the buffer. The samples were probed with anti-RAD51 antiserum and compared to a positive control (total protein extract from N2-5). The endogenous copy of RAD51 is visible at 47kDa, whilst the TAP tagged copy of RAD51 is visible at 62kDa. Sizes are indicated in kDa.

These data appear to suggest that in the lysis conditions used in the first TAP purification assay (IPP150), a significant quantity of RAD51 was pelleted following centrifugation of cell lysates, which may be consistent with the suggestion that RAD51 may not have been present in the supernatant that was subjected to affinity purification. Raising the salt concentration to 300 mM and to 450 mM appeared to reduce the amount of RAD51 present in the pellet fraction, leading to the assumption that RAD51 may be enriched in the supernatant fractions. Unfortunately, this assumption could not be confirmed, since RAD51 was undetectable in western blot analysis of any of the supernatant fractions (data not shown).

To determine if raising the salt concentration during cell extraction improved RAD51 yield, it was decided to repeat the TAP method, as above, but using the 450 mM NaCl lysis conditions to prepare the cell extracts. The same two independent procyclic transformants were grown to a density of approximately 1×10^7 cells.ml^{-1}, before cell extracts were prepared under IPP450 conditions. The extracts were then subjected to affinity purification using IgG sepharose beads. Putative complexes purified with IgG beads were treated with TEV protease, and then subjected to affinity purification using calmodulin beads. As before, material from each stage of the purification was loaded onto 10 % SDS-PAGE gels, before staining with Coomassie (figure 6.27).
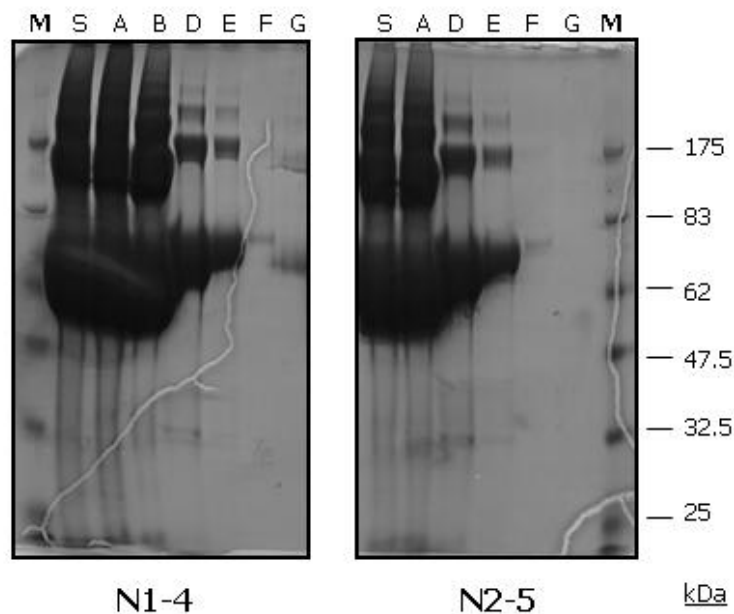
Figure 6.27 – Coomassie stained SDS-PAGE gel displaying products obtained throughout RAD51-TAP. Two independent procyclic form RAD51-TAP tagged transformants were subjected to the TAP purification procedure and samples from each step loaded onto 10 % SDS-PAGE gels. A – cell extract supernatant, B – wash with IPP150 from IgG column, C – wash with TEV cleavage buffer, D – eluate from IgG column, E – flow through from calmodulin column, F – wash with calmodulin binding buffer, G – eluate from calmodulin column, M – protein marker (NEB-broad range). Sizes are displayed in kDa.

The results from this purification method appeared to be virtually identical to those observed with the 150 mM NaCl extract. Again, bands corresponding to RAD51 could not be detected in the purification, nor indeed in the extract (as confirmed by western blot; data not shown). The reasons for this are unclear, but to attempt to improve the enrichment of RAD51, it was decided to attempt the TAP procedure using a nuclear extract instead of a whole cell extract. The rational behind this was that the different extraction procedure may more effectively remove RAD51 putatively bound to DNA, and may increase the concentration of the RAD51 protein.

Procyclic transformants were grown to a density of approximately 1×10^7 cells.ml^{-1}, before nuclear extracts were prepared, essentially according to methods described by Bell and Barry, 1995. 3×10^9 *T. brucei* cells were pelleted by centrifugation (1600 g) and washed twice in PBS. Following the final wash, the trypanosomes were resuspended in 4 mls of Buffer A (20 mM Tris-Cl (pH 8.0) 10 mM NaCl, 0.5 mM DTT). Cells were then lysed by 25 strokes of a Dounce homogeniser. Nuclei were pelleted for 5 minutes at 3700 g in a Beckman JS-7.5 rotor, and resuspended in 2 mls Buffer C (50 mM Tris-Cl (pH 8.0), 25 % glycerol, 400 mM NaCl, 0.2 mM EDTA, 0.5 mM DTT) using a Dounce homogenizer, and mixed gently at 4 °C for 30 minutes. The resulting nuclear lysate was centrifuged at 25 000 g for 30 minutes at 4 °C, and the supernatant dialysed against 50 volumes of Buffer D (50 mM Tris-Cl (pH 8.0), 20 % glycerol, 100 mM NaCl, 0.2 mM EDTA, 0.5 mM DTT) overnight at 4 °C. The dialysate was centrifuged at 25 000 g for 25 minutes at 4 °C (Beckman JA21 rotor) and the supernatant subjected to affinity purification using IgG sepharose beads, treatment with TEV protease and affinity purification with calmodulin beads, as before, except that unbound material from the IgG column were washed with 30 mls of Buffer D instead of IPP150. To attempt to concentrate any proteins released from the final elution step, this was performed using 2 mls of calmodulin elution buffer and subsequently concentrated using a Centricon centrifugal filter unit (Millipore). Samples from each stage of the purification was loaded onto 10 % SDS-PAGE gels, as were samples of the beads, before staining with Coomassie (figure 6.28).
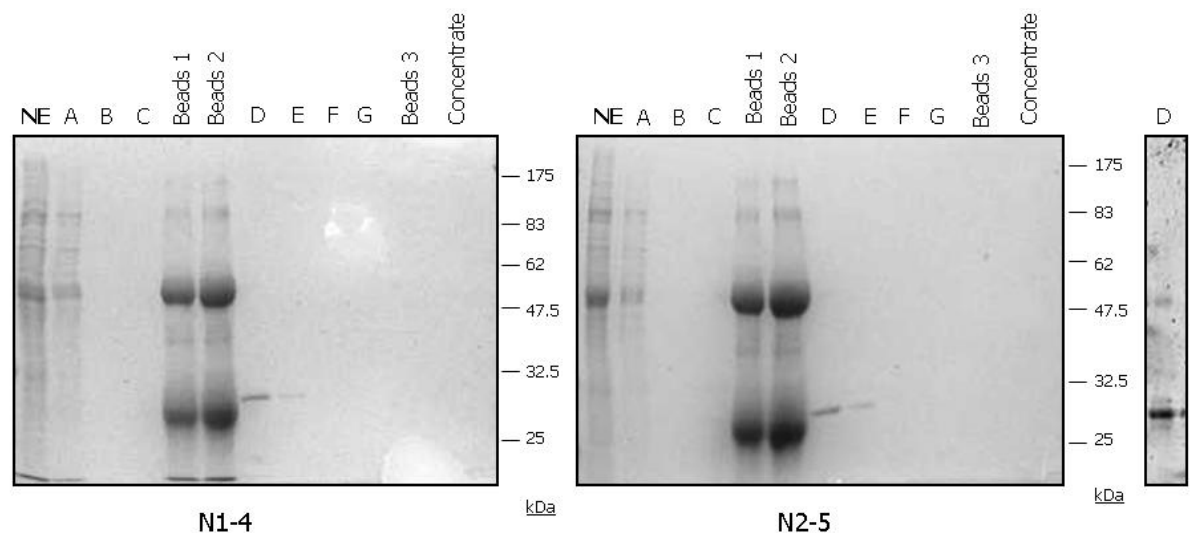
Figure 6.28 – Coomassie stained SDS-PAGE gel displaying products obtained throughout the RAD51 TAP from procyclic nuclear extracts. Two independent procyclic form RAD51-TAP tagged transformants were subjected to the TAP purification procedure and samples from each step loaded onto 10 % SDS-PAGE gels. NE – nuclear extract, A – unbound extract, B – wash with buffer D, C – wash with TEV cleavage buffer, Beads 1 – IgG beads before TEV cleavage, Beads 2 – IgG beads after TEV cleavage, D – eluate from IgG column, E – flow from calmodulin column, F – wash with calmodulin binding buffer, G – eluate from calmodulin column, Beads 3 – calmodulin beads, Concentrate – eluate after centricon concentration (Millipore). The right hand gel displays lane D from N2-5 which was stained with Sypro Ruby (Bio-RAD). Sizes are displayed in kDa.

In contrast to previous attempts, Coomassie staining of this purification did not reveal any proteins in the steps following elution from the IgG column (D). To see if proteins were present, but in smaller quantities, the samples from transformant N2-5 following elution from IgG, were loaded onto a 10 % SDS-PAGE gel and stained with Sypro Ruby (Bio-RAD), as this stain is more sensitive than Coomassie (Berggren *et al.*, 1999). Sypro Ruby staining revealed a faint band visible at ~50 kDa in addition to a more prominent ~30 kDa band. This appeared not to be RAD51, however, as it was not detected on western blot analysis (data not shown).

At this stage it was decided to repeat the tandem affinity purification, again using nuclear extract, taking the procedure only as far as IgG elution and omitting the calmodulin column. A number of conditions were then attempted to optimise binding and elution for the IgG column. This was carried out on just one transformant (N1-4).

In this experiment, supernatant from an early centrifugation step (see below) in the nuclear extract preparation was applied to the IgG column prior to the nuclear extract, in case RAD51 was primarily cytoplasmic, and was therefore being discarded during the preparation of the nuclear extract. In addition, the nuclear extract was examined, but dialysis was omitted from the preparation to reduce the potential loss of material. As before, 3×10^9 trypanosomes were pelleted by centrifugation (1600 g) and washed twice in PBS. Following the final wash, the cells were resuspended in 4 mls of Buffer A, before lysis with 25 strokes of a Dounce homogeniser. Nuclei were then pelleted for 5 minutes at 3700 g in a Beckman JS-7.5 rotor, and supernatant (NE1) subjected to affinity selection using IgG sepharose beads. NE1 was incubated with the IgG beads by rotation at 4 °C for 30 minutes, before unbound material was removed by washing with 30 mls of Buffer A and 10 mls of Buffer C. To prepare a nuclear extract, the pellet from the centrifugation step above was resuspended in 2 mls Buffer C, as in previous work, using a Dounce homogeniser and mixed gently at 4 °C for 30 min. The resulting nuclear lysate was centrifuged for 30 min at 4 °C, 25 000 g and this supernatant (NE2) subjected to affinity selection using the same IgG sepharose beads. Here NE2 was rotated with the IgG beads for 2 hours at 4 °C, before unbound material was removed by washing with 30 mls of Buffer C and 10 mls of TEV cleavage buffer. For the purification, the IgG beads and bound complexes were resuspended in 2 mls of TEV cleavage buffer before treating with 500 units of TEV protease (Invitrogen) and rotating overnight at 4 °C. The eluate was recovered by gravity flow and washed through with 2 mls of TEV cleavage buffer. Samples from each stage of the purification were loaded onto 10 % SDS-PAGE gels,

before staining with Sypro Ruby (Bio-Rad) (figure 6.29). The same samples were
subsequently analysed by western blot analysis, probing with anti-RAD51 antiserum and
peroxidase anti peroxidase (PAP) (figure 6.30).
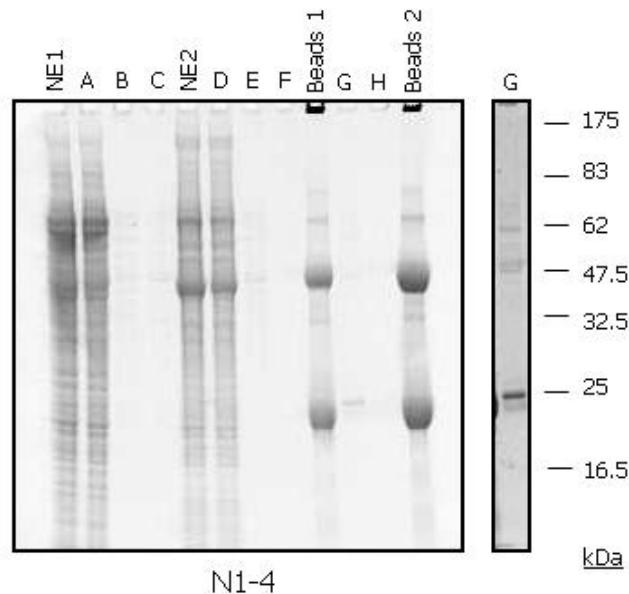
Figure 6.29 – Sypro-Ruby stained SDS-PAGE gel displaying products obtained from the IgG column in the RAD51 TAP. A procyclic RAD51-TAP tagged transformant was subjected to the first column in the TAP purification procedure and samples from each step loaded onto 10 % SDS-PAGE gels. NE1 – 'nuclear' extract 1, A – unbound extract, B – wash with buffer A, C – wash with buffer C, NE2 – nuclear extract 2, D – unbound extract, E – wash with buffer C, F – wash with TEV cleavage buffer, Beads 1 – IgG beads before TEV cleavage, G – eluate, H – wash with TEV cleavage buffer, Beads 2 – IgG beads after TEV cleavage. The right hand gel displays lane G which was exposed for a longer period of time. Sizes are displayed in kDa.

From the gel shown in figure 6.29, each extract appeared to contain a large selection of
proteins, at least some of which were distinct, which would be expected. The resulting
eluate from the IgG column (G), at least when exposed for a prolonged period, also
contained a number of proteins. Western blotting revealed the presence of RAD51 in both
the supernatant and nuclear extract preparations (figure 6.30). However, the resulting
eluate from the IgG column was found not to contain RAD51, suggesting that RAD51 had
failed to be released by TEV cleavage.
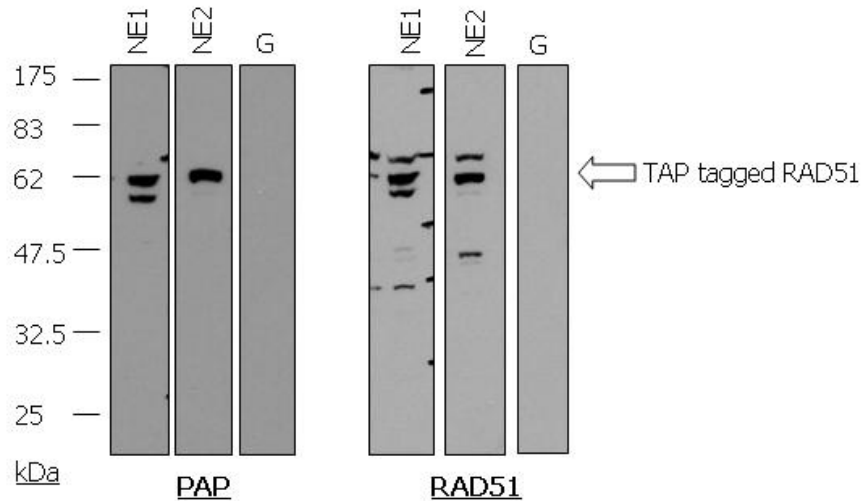
Figure 6.30 – Western blot of extracts applied to and eluted from the IgG column. The western blot displays NE1 – the cytoplasmic extract, NE2 – nuclear extract and G – eluate from the IgG column. The samples were probed with peroxidase anti peroxidase (PAP) and anti-RAD51 antiserum. The endogenous copy of RAD51 is visible at 47kDa, whilst the TAP tagged copy of RAD51 is visible at 62kDa. Sizes are indicated in kDa.

The failure of RAD51 elution from the IgG column by TEV cleavage could have been caused by a number of problems. One possibility is that the TEV protease sample used is unstable, perhaps through unfolding or degradation and would therefore be unable or less capable of cleaving TAP-tagged RAD51. To investigate this possibility, the activity of the TEV protease was examined (see below). Another possibility for the lack of detectable RAD51 in the eluted fraction from the IgG beads, could simply be that RAD51 is present in low abundance, below the threshold for detection by the RAD51 antiserum. Indeed, the PAP antiserum would not detect any cleaved products in this fraction, as it recognises the Protein A part of the TAP tag. This was investigated by probing with a CBP antibody, which recognises the calmodulin region of the TAP tag (see below). Another possibility could be that the salt concentrations in the buffers might interfere with the activity of the TEV protease. To address this, it was necessary to dialyse the nuclear extract to the recommended IPP150 salt concentrations (see below). Finally, it is conceivable that a sequence error could have been generated in the TEV cleavage site when the constructs were being produced. This option seems unlikely due to the use of a high fidelity DNA polymerase (Stratagene). Nevertheless, this option was examined through DNA sequencing (see below).

The activity of the TEV protease was investigated by incubating it with a TEV target protein (gift from B. Hunter, University of Dundee), which includes the TEV recognition site (ENLYFQS). The results of the TEV protease that was commercially purchased (AcTEV™ – Invitrogen) were also compared with the activity of TEV protease purified from *E. coli* in the McCulloch laboratory (gift from Dr. C. Stockdale). Invitrogen claim

that 1 unit of AcTEV™ enzyme will cleave > 99 % of a 3 µg control substrate in 1 hour at 30 °C. To test this, reactions were set up containing 1 X TEV buffer (50 mM Tris-Cl, pH 8.0, 0.5 mM EDTA), 1 mM DTT, 4 µg TEV target protein and different amounts of TEV protease (either AcTEV™ or TEV by Dr. C. Stockdale). The reactions were incubated at room temperature for 1 hour before being loading onto a 15 % SDS-PAGE gel and visualised by Coomassie staining (figure 6.31).

The results in figure 6.31 suggest that the AcTEV™ protease is not as active as Invitrogen claims: even with 4 units of enzyme, although some cleaved proteins were observed, substantial amounts of uncleaved substrate were visible. The activity of the 'home-made' TEV protease (Dr. C. Stockdale) was unknown, so three different concentrations were used, each of which were clearly higher than any of the AcTEV™ concentrations. At each of these concentrations, the enzyme appeared capable of cleaving all of the TEV target protein. In the attempts at TAP, 300 units of AcTEV™ protease was used at 4 °C for around 18 hours. Though it is possible that the protein is active in these conditions, it is also possible that insufficient AcTEV™ protease was used, meaning the yield of RAD51 from the IgG column was low. To address this, the 'home-made' TEV protease was used in RAD51 TAP (see below).
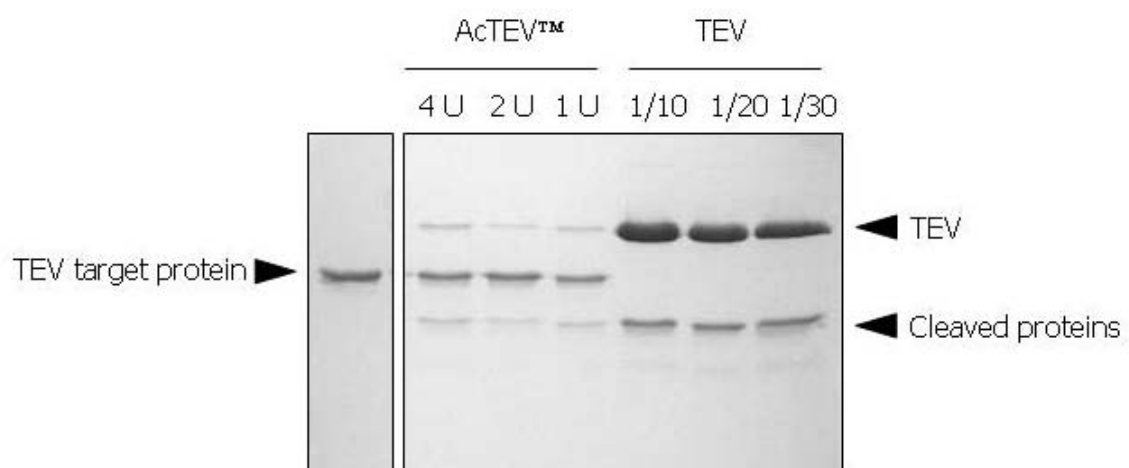
Figure 6.31 – Coomassie stained SDS-PAGE gel displaying the activity of different TEV proteases. The TEV target protein, TEV protease (TEV) and cleaved products are displayed. The AcTEV™ (Invitrogen) protease was used in 3 different reactions with 4 units (4 U), 2 units and 1 unit per 4 µg of TEV target protein. The activity of the TEV protease (Dr. C. Stockdale – gift) was unknown, and was used at 3 different dilutions from stock: 1/10, 1/20 and 1/30.

In order to detect the calmodulin epitope of the TAP tag by western analysis, a CBP TAG antibody (Santa Cruz Biotechnology – sc-33000) was used. Western blot analysis was carried out on total protein extracted from the N1-4 and N2-5 RAD51 TAP-tagged transformants generated in procyclic form cells. Cell extracts were separated on 10 %

SDS-PAGE gels and probed with CBP TAG antiserum and detected with HRP-coupled anti-goat IgG (Santa Cruz Biotechnology – sc02020). The results of this western blot, and many subsequent attempts proved, largely unsuccessful, as the CBP antibody cross reacted with whole cell extract from *T. brucei*, making specific detection of the RAD51 TAP-tagged protein impossible (data not shown).

The sequence of the TEV cleavage sites in the transformants was examined by DNA sequencing. Genomic DNA was prepared from all the transformants and PCR amplification conducted using a high fidelity DNA polymerase (Stratagene) and the primer pairs *NanalTAP5'* and *NanalTAP3'*. The resulting DNA fragments were purified and sequenced (Dundee Sequencing Service), using the same primers. In all cases no sequence errors were found in the TEV cleavage site.

To address the question surrounding TEV cleavage, another TAP procedure was conducted. This was performed on both RAD51-TAP tagged procyclic transformants (N1-5 and N2-4) and on untagged, wild type cells as a control. The main difference in this procedure compared with previous attempts, was that the nuclear extract was diluted until the NaCl concentration was 200 mM, in order to avoid potential interference with the TEV protease activity, and the TEV protease made by Dr. C. Stockdale was used instead of the Invitrogen AcTEV™ protease.

Nuclear extracts were prepared as before from 3×10^9 trypanosomes until the stage of lysis in 2 mls of Buffer C at 4 °C for 30 minutes. The resulting nuclear lysate was centrifuged for 30 min at 4 °C, 25 000 g and the supernatant diluted to 200 mM NaCl by the addition of an equal volume of Buffer C lacking salt (50 mM Tris-Cl (pH 8.0), 25 % glycerol, 0.2 mM EDTA, 0.5 mM DTT). The extract was then subjected to affinity selection using IgG sepharose beads. For this, 500 μl IgG sepharose beads (Amersham) were placed in a 10 ml disposable polypropylene column (Pierce) and washed with 5 mls Buffer C – 200mM (50 mM Tris-Cl (pH 8.0), 25 % glycerol, 200 mM NaCl, 0.2 mM EDTA, 0.5 mM DTT). Protein was allowed to bind by rotating the column for 4 hours at 4 °C. Following this, unbound material was removed by washing with 20 mls of Buffer C – 200mM and 20 mls of TEV cleavage buffer. The IgG beads and bound proteins were then resuspended in 2 mls of TEV cleavage buffer before treating with 25 μl of TEV protease (Dr. C. Stockdale) and rotating overnight at 4 °C. The eluate was recovered by gravity flow and subsequently washed through with 2 mls of TEV cleavage buffer, before being subjected to affinity purification using calmodulin beads. This step was performed by adding 3 volumes of calmodulin binding buffer (10 mM β–mercaptoethanol, 10 mM Tris-Cl (pH 8.0), 150 mM

NaCl, 1 mM Mg-acetate, 1 mM imidazole, 2 mM $CaCl_2$, 0.1 % NP40) and 3 mM $CaCl_2$ to the eluate and rotating in a calmodulin column for 4 hours at 4 °C. The calmodulin column was prepared in the same manner as the IgG column, but using 500 µl calmodulin beads (Amersham) washed with 5 mls calmodulin binding buffer. The eluate was removed by gravity flow and further unbound material was removed by washing with 30 mls calmodulin binding buffer. 1 ml of calmodulin elution buffer (10 mM β–mercaptoethanol, 10 mM Tris-Cl (pH 8.0), 150 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 2 mM EGTA, 0.1 % NP40) was incubated with the beads for 1 hour at 4 °C before eluting the fraction. This fraction was subsequently concentrated by TCA precipitation.

Samples from each stage of the purification were loaded onto 10 % SDS-PAGE gels, before staining with Sypro Ruby (Bio-Rad) (figure 6.32). The same samples were subsequently analysed by western blot analysis, probing with anti-RAD51 antiserum and peroxidase anti peroxidase (PAP) (figures 6.33 and 6.34).
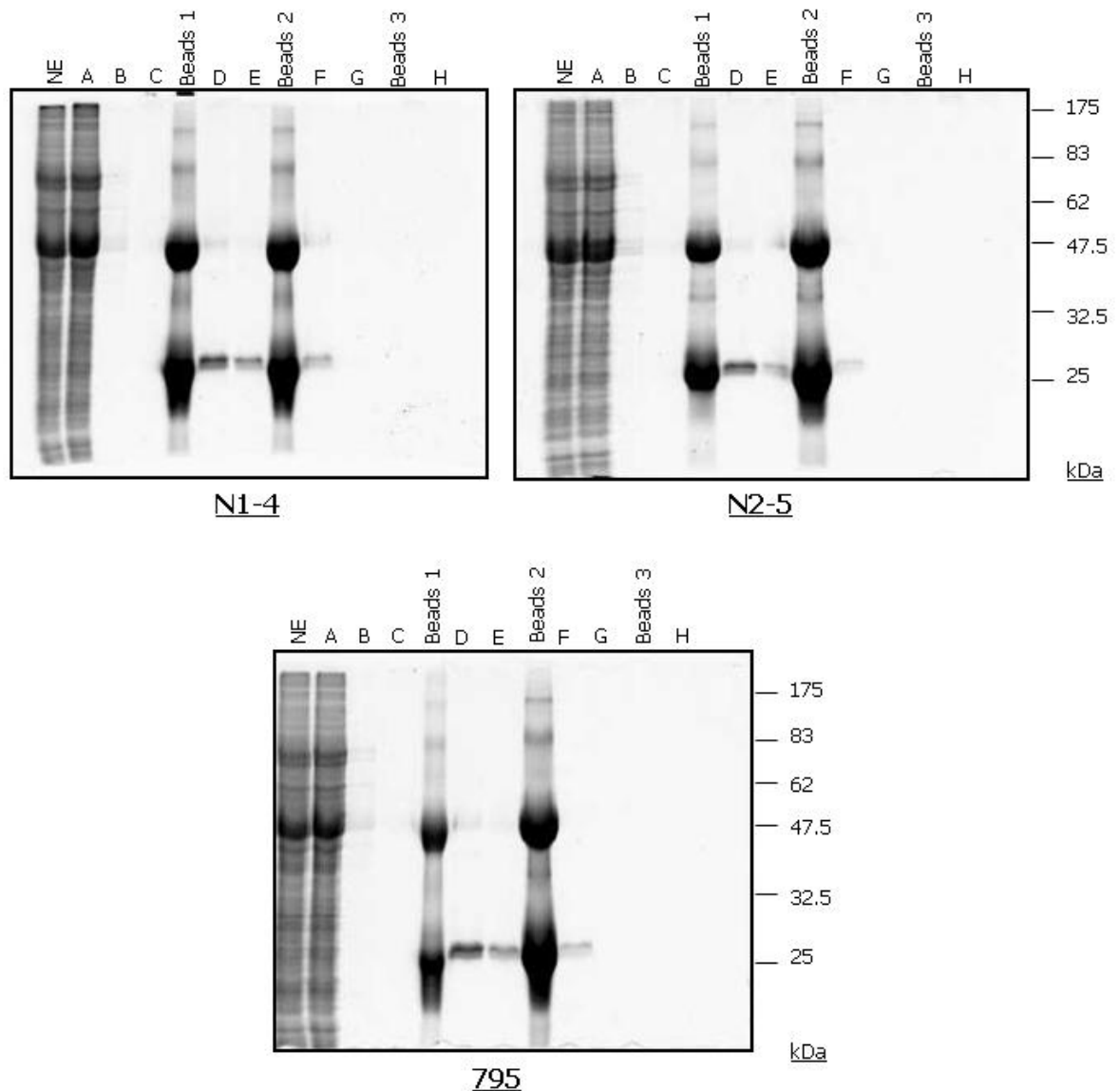
Figure 6.32 – Sypro-Ruby stained SDS-PAGE gel displaying products obtained throughout RAD51-TAP. Extracts from two independent TAP tagged procyclic transformants (N1-1 and N2-5) were subjected to the TAP purification procedure and compared with wild type extracts (from EATRO 795 cells; 795). Samples from each step were loaded onto 10 % SDS-PAGE gels. NE – nuclear extract 1, A – unbound extract from IgG beads, B – wash with buffer C (200 mM), C – wash with TEV cleavage buffer before TEV cleavage, Beads 1 – IgG beads before TEV cleavage, D – eluate, E – wash with TEV cleavage buffer after TEV cleavage, Beads 2 – IgG beads after TEV cleavage, F – unbound extract from calmodulin beads, G – wash with calmodulin binding buffer, Beads 3 – calmodulin beads, H – final eluted fraction concentrated by TCA precipitation. Size markers are displayed in kDa.

The results displayed in figure 6.32 suggest that the limited amount of protein purified from the TAP tagged transformants appear no different than the wild type samples, which suggests that no TAP tagged RAD51-specific purification has occurred. Indeed, no proteins were visible in the final sample eluted from both columns, from any extract.

The western blots displayed in figure 6.33 show the samples from the RAD51-TAP tagged transformants in figure 6.32 probed with PAP, to detect the protein A epitope of the TAP tag. These confirm that TAP tagged RAD51 was only present in the TAP tagged transformants, as it was not detected in the samples from the wild type cells. In addition,

signal was only detectable in the nuclear extract fraction and in the unbound, flow through fraction. Although bands were also detectable in the lanes where samples of the beads were loaded, this does not indicate the presence of a TAP tag, but instead results from cross reaction of the anti-serum with the heavy chain of the IgG beads (~47 kDa) (the heavy and light chains are clearly visible on the Coomassie stain; figure 6.32).

These data indicate that TAP-tagged RAD51 was present in the extracts of the transformants, but that a large portion of this RAD51 either did not bind to the IgG beads or bound weakly and was rapidly washed off. One possibility to explain this is that insufficient IgG beads were provided in the IgG column to allow all of the tagged RAD51 to bind. Alternatively, the conformation of RAD51 prevents effective binding via the TAP tag.
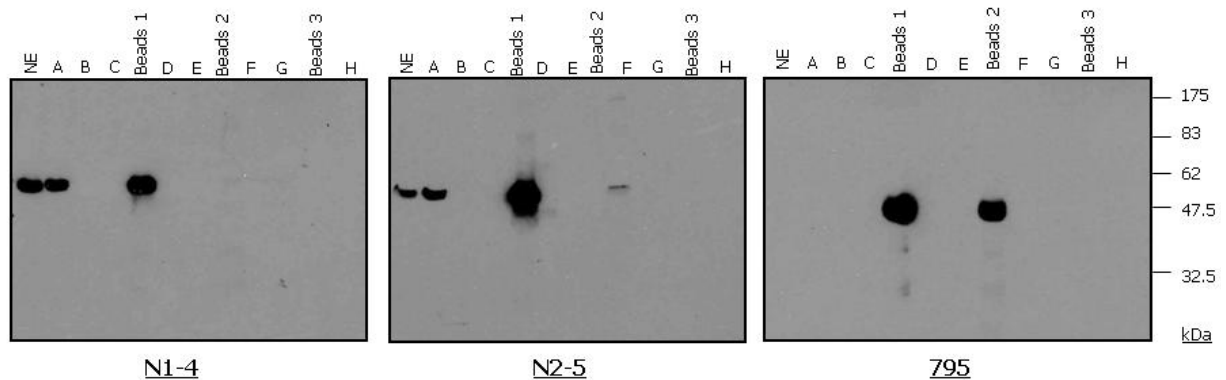
Figure 6.33 – Western blot of products obtained throughout RAD51-TAP. The western blot displays the samples from figure 6.32 probed with peroxidase anti peroxidase (PAP). The identities are as in figure 6.32. The TAP tagged copy of RAD51 is visible at 62kDa. Size markers are indicated in kDa.

The western blot displayed in figure 6.34 again shows the same samples from figure 6.32, but on this occasion probed with anti RAD51 antiserum. This confirms that endogenous, untagged RAD51 was present in the nuclear extract from all cell lines and only in the TAP tagged transformants were tagged copies of RAD51 observed. In addition, it supports the conclusion that a large proportion of tagged RAD51 was not retrieved by the IgG beads. Cross reacting bands were again visible in the lanes loaded with samples if the IgG beads. When these blots were allowed to expose overnight, faint bands of the size expected for RAD51 could be detected in the final eluted fraction (lane H) and on the calmodulin beads (Beads 3). These bands were only detected on the blots from the TAP tagged transformants, not from wild type cells (data not shown). This result appears to suggest that vary small quantities of RAD51 have been purified during the TAP procedure. However, as it has been conducted to date, this is too low a quantity to be detected by

Sypro-Ruby staining, which in turn means that any putative interacting proteins would also be present in low abundance.
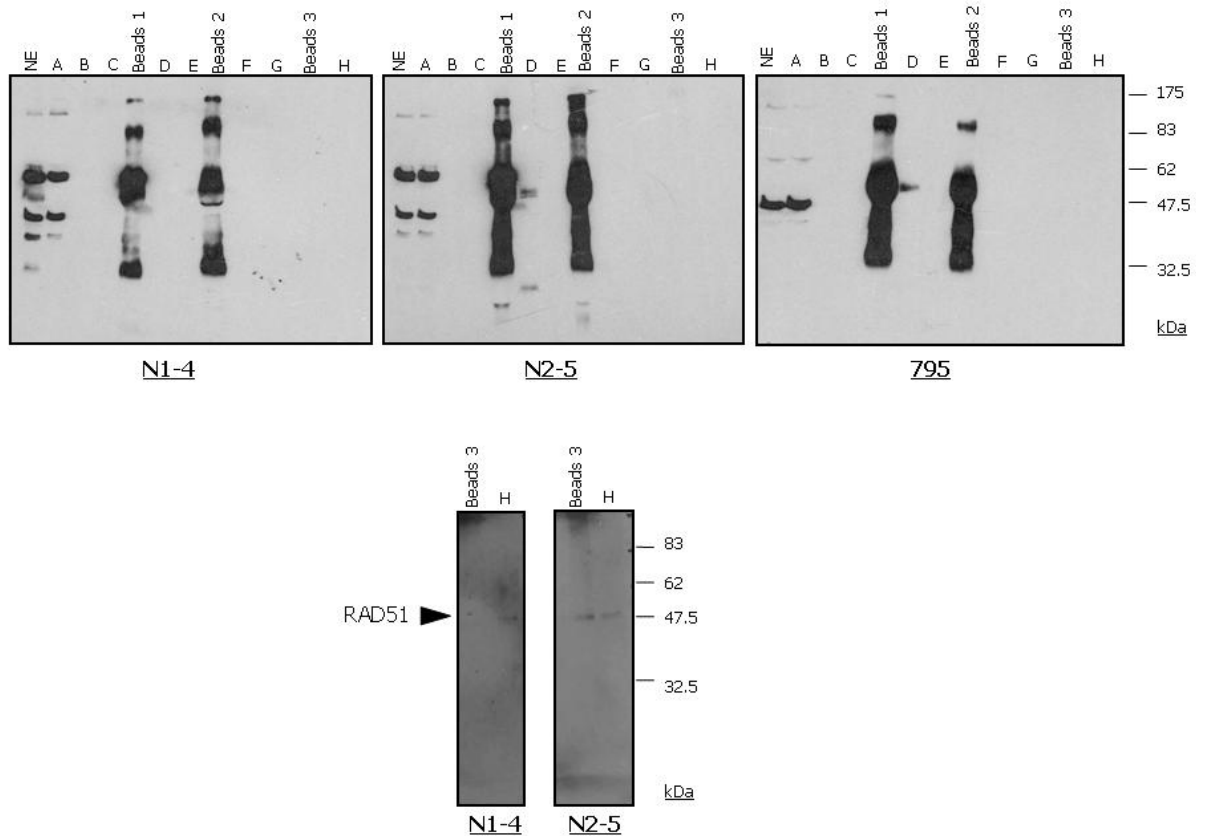
Figure 6.34 – Western blot of products obtained throughout RAD51-TAP. The western blot displays the samples from figure 6.32 probed with anti-RAD51 antiserum. The identities are as in figure 6.32. The TAP tagged copy of RAD51 is visible at 62kDa. The lower blots show the signal from samples 'Beads 3' and 'H' after overnight exposure. Size markers are indicated in kDa.

These results appear to demonstrate that the TAP tagging of RAD51 could provide useful information regarding potential interacting factors. However, the procedure would require further optimising. It is likely that the amount of nuclear extract subjected to affinity purification and the quantity of IgG beads used to produce the IgG column would need to be increased. However, due to time constraints, these experiments were unable to be performed.

6.7 Identifying RAD51 interacting factors using purified GST tagged RAD51 immobilised onto GST beads

Given the difficulties described in using RAD51 TAP, it was decided to attempt another method for identifying interacting partners. GST tagged RAD51 had recently been purified in the McCulloch laboratory (Dr. C. Stockdale) (figure 6.35). This was used to provide an alternative route towards the identification of RAD51 interacting factors by mixing nuclear extract from wild type EATRO 795 procyclic form cells with either GST protein alone or with GST tagged RAD51. The proteins that interact with either GST or GST-RAD51 were then compared following purification using a glutathione column.
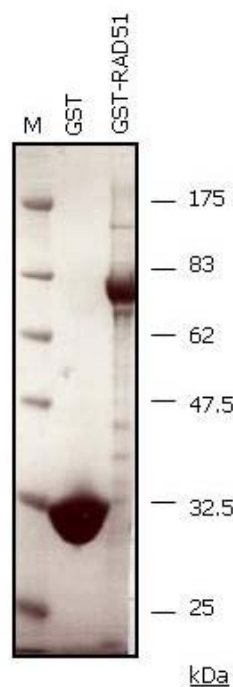
Figure 6.35 – GST and GST tagged RAD51 purified proteins. Purified proteins were run out on a 10 % SDS-PAGE gel and stained with coomassie. M indicates the size marker lane and sizes are shown in kDa.

Nuclear extracts were prepared as before from 3×10^9 trypanosomes, and the resulting nuclear lysate centrifuged for 30 min at 4 °C, 25 000 g, before the supernatant was diluted to 20 mM Tris-Cl, 10 % glycerol, 200 mM NaCl.

Reactions were set up containing 5 μg of either GST or GST-RAD51, 5 μg of nuclear extract, 20 mM Tris-Cl (pH 8.0), 1 mM DTT, 10 % glycerol, 200 mM NaCl and 0.1 % NP-40. The reactions were incubated on ice for 30 minutes to allow any interacting factors to bind. Following this incubation, 40 μl of glutathione beads (Amersham) were added and the reactions rotated for 1 hour at 4 °C. The beads were then harvested by centrifugation (500 g) and the supernatant removed. Any unbound material was removed from the beads

by washing 4 times with 750 μl binding buffer (20 mM Tris-Cl (pH 8.0), 1 mM DTT, 10 %
glycerol, 200 mM NaCl and 0.1 % NP-40). The proteins attached to the beads were finally
recovered by resuspension in 10 μl 3 x SDS PAGE buffer (150 mM Tris0Cl (pH 6.8), 30
% glycerol, 6 % SDS and 0.3 % bromophenol blue). Samples from each stage of these
purifications were separated on 10 % SDS-PAGE gels, before staining with Sypro Ruby
(Bio-Rad) (figure 6.36).
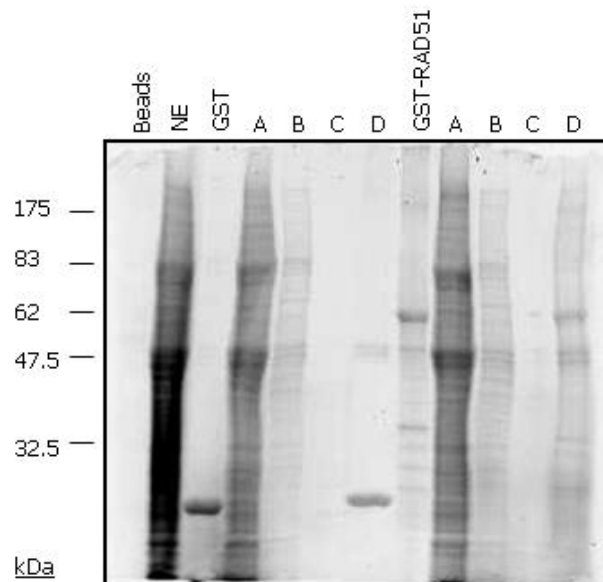
**Figure 6.36 – Sypro-Ruby stained SDS-PAGE gel displaying products obtained throughout
the GST purification method.** Beads – sample of the GST beads; NE – nuclear extract; GST –
GST purified protein; GST-RAD51 – GST tagged RAD51 purified protein; A – unbound material; B
– wash 1; C – wash 4; D – proteins which remained bound to the GST beads.

The results from the GST purification show that some proteins in the *T. brucei* nuclear
extract had the ability to bind to the GST protein alone (GST - lane D), as seen by 2
distinct bands approximately 48 kDa and 50 kDa in size. These proteins were also visible
in the sample which contained the proteins that remained bound to the GST-RAD51
purified protein (GST-RAD51 – lane D) and can therefore be excluded as being
meaningful interacting partners. The other proteins that were visible in the sample of
GST-RAD51 purified material appeared mainly to be contaminating proteins that were
already present in the GST-RAD51 purified protein sample (GST-RAD51). Given this,
whether or not any true RAD51 interacting factors were present was difficult to judge.
This method, which was only attempted on one occasion, could yet prove useful in
identifying interacting factors, but has a considerable disadvantage relative to TAP, as it
relies on only one affinity purification step.

6.8 Summary

This chapter aimed to identify RAD51 interacting factors in *T. brucei*, both in the bloodstream stage and in the procyclic form, before and after induced DNA damage. Unfortunately, putative RAD51 interacting factors were unable to be identified, largely due to lack of time to continue this project and difficulties in establishing the tandem affinity purification procedure with RAD51.

The work in this chapter has informed us that it is possible to add an epitope to the N terminus of RAD51 in *T. brucei* cells, in both the bloodstream stage and procyclic form. The addition of a TAP-tag at the N terminus of RAD51 was found not to have a detrimental effect on the parasite in terms of growth, sensitivity to MMS and its ability to form RAD51 foci, all of which are affected in *rad51-/-* mutants (McCulloch and Barry, 1999). This result supports research performed in other systems, whereby RAD51and RecA have been successfully GFP tagged at their N and C termini, respectively without causing disruption to its function (Essers *et al.*, 2002b;Yu *et al.*, 2003;Renzette *et al.*, 2005;Kojic *et al.*, 2005). In contrast to the N terminus, generating C terminally tagged RAD51 in *T. brucei* cells was found to be problematic. Despite a number of transformations in both bloodstream and procyclic cells, only one C-terminally tagged RAD51 variant was recovered. Though the basis for these problems was not explored, it was speculated that this may be due to the positioning of the C terminus in the interior of the RAD51 nucleoprotein filament (Conway *et al.*, 2004).

Attempts at the TAP procedure discovered that for purification of proteins that interact with RAD51 in this organism, a great deal of optimisation is still required. At a minimum, it is likely that the amount of nuclear extract subjected to affinity purification needs to be substantially increased, as potentially does the amount of IgG beads in order to allow TAP tagged RAD51 to bind.

These problems with TAP may, in fact, not be specific to RAD51. Recently, the TAP procedure has been identified as being inefficient for other proteins in *T. brucei* (Schimanski *et al.*, 2005). These authors suggested that, specifically, the calmodulin purification step of TAP was inefficient, and proposed that the reason behind this was that endogenous calmodulin in trypanosome extracts interacts with the calmodulin binding peptide (CBP) and therefore prevents large amounts of the TAP tagged protein from binding to the calmodulin column. To circumvent these problems, Schimanski *et al.*, developed a modified version of the TAP tag (PTP), whereby the CBP region was replaced

with a protein C epitope. They performed a direct comparison between the TAP tag and the PTP tag on the transcription factor SNAPc, and discovered that purification in the PTP was more efficient, at least for this protein. A modified version of the TAP tag has also been developed in mammalian cells (Drakas *et al.*, 2005), due to the original TAP method not yielding enough purification for protein identification in mammalian cell lines grown in monolayers. In this case, these authors added a biotinylated tag to the TAP tag in order to increase protein yield from cell extracts. It is worth noting, however, that these technical issues are not necessarily relevant to the TAP analysis of RAD51, as the work in this chapter showed that binding of the protein to the first, IgG column was inefficient, for reasons that are yet to be resolved.

The aim of this chapter was to identify RAD51 interacting factors in *T. brucei*, in particular following DNA damage, where the protein forms discrete sub-nuclear foci (Proudfoot and McCulloch, 2005). To date, only TAP was performed on undamaged procyclic form cells. It may be of interest to compare the procedure with bloodstream stage cells, and following damage. Though we know that RAD51 levels do not increase in response to damage in *T. brucei* (C. Proudfoot, Thesis), we do not know if the proteins' sub-cellular location is changed or if it becomes activated in some way to respond to damage. Potentially, such changes could alter RAD51's behaviour during the TAP procedure. In addition, we do not know if the protein behaves equivalently in the bloodstream and procyclic stages, though it seems likely. Of note, however, this chapter demonstrates that RAD51 sub-nuclear foci do form in procyclic form cells, which had not previously been tested.

CHAPTER 7

Discussion

7.1 Introduction

The main aim of this thesis was to further examine the factors that regulate antigenic variation in *Trypanosoma brucei*. This was to be achieved through two main areas of investigation. The first of these was the examination and characterisation of BRCA2 in *T. brucei*, in terms of both its structure and function. This area of investigation was to prove to be the main body of research, and stemmed initially from the suggestion that the BRCA2 homologue in *T. brucei* possessed a highly unusual organisation, containing 15 BRC repeats, a much higher number than observed in any other organism (Lo *et al.*, 2003). We hypothesised that this was a structural adaptation to account for the demands that antigenic variation places on the *T. brucei* homologous recombination reaction. This question was first examined by determining the actual number of BRC repeats in various *T. brucei* strains and subspecies. Subsequent analyses examined the function of the protein through the generation of *BRCA2* knockout mutants in Lister 427 bloodstream stage *T. brucei* and its derivative 3174.2. A series of phenotypic analyses were carried out on these cell lines in order to determine the role of BRCA2 in the repair of induced damage, homologous recombination and antigenic variation. Following on from this, a number of *T. brucei BRCA2* structural variants were generated and functionally characterised. The results of these lines of research, which had the potential to reveal why *T. brucei* BRCA2 contains so many BRC repeats, and also provide information as to the function of certain motifs within the protein, are discussed below.

The second area of investigation was to examine the role and molecular composition of *T. brucei* RAD51 sub-nuclear foci, based on the assumption that the foci are repair centres containing multiple homologous recombination factors and specific sites of DNA lesions. The main aim of this investigation was to identifying RAD51 interacting factors, before and after induced damage, through tandem affinity purification (TAP) methods (Rigaut *et al.*, 1999). However, due to complications in optimising this method for RAD51 in *T. brucei,* and to time limitations, no RAD51 interacting partners were identified (see chapter 6), so this area of investigation will not be discussed here.

7.2 T. brucei BRCA2 has undergone a recent expansion in BRC repeats

The BRC repeats of BRCA2 have been shown to be critical for the interaction with RAD51, the key enzyme of eukaryotic homologous recombination (Pellegrini and Venkitaraman, 2004;Shivji and Venkitaraman, 2004;Sung and Klein, 2006). This finding,

along with the observation that all known orthologues of BRCA2 possess at least one BRC repeat, with one or more interacting with RAD51 when multiple repeats are present (Dray *et al.*, 2006), indicates their necessity for BRCA2 function. Further evidence for this is provided by the identification of critical residues for RAD51 interaction located within the BRC repeat sequence, which are found to be conserved between several different species (Bignell *et al.*, 1997). In addition, mutations located within the BRC repeats have been shown to associate with familial ovarian cancer (Gayther *et al.*, 1997).

The number of BRC repeats differs quite significantly between BRCA2 homologues, but a general theme appears to exist, in which, the simpler the organism, the smaller the number of BRC repeats. Indeed, out of 12 BRCA2 homologues investigated in a range of unicellular organisms, 8 contain between 1 and 3 BRC repeats (Lo *et al.*, 2003) (Table 3.3). Examples of this come from the uni-cellular organisms *U. maydis*, *T. cruzi*, and *G. Lamblia* which have been shown to possess 1, 2 and 1 BRC repeats respectively (Lo *et al.*, 2003). In contrast, of 8 multi-cellular organisms, 7 have 3 or more BRC repeats (Lo *et al.*, 2003). Indeed, most vertebrate BRCA2 proteins have been shown to contain 8 BRC repeats, whilst the plant *A. thaliana* and the insect *D. melanogaster* contain 4 and 3 BRC repeats, respectively. The reason why some BRCA2 homologues contain multiple BRC repeats whilst others function with just a single repeat has not yet been investigated. However, it could be speculated that the larger genome sizes or more complex biological systems found in multi-cellular eukaryotes might exert an evolutionary pressure for an increased BRC repeat number in BRCA2, due to a greater need for homologous recombination. For example, larger numbers of BRC repeats could sequester the putatively greater amount of RAD51 needed in these organisms until it is needed for DNA repair, thereby preventing uncontrolled recombination. This hypothesis gains some support from findings that RAD51 exists in the mammalian nucleus in relatively immobile pools, one of which is bound to BRCA2 (Essers *et al.*, 2002b;Yu *et al.*, 2003). Furthermore, the BRC repeats of human BRCA2 have been shown to disrupt pre-formed RAD51 filaments and impair homologous recombination, implying that the BRC repeats interact with a monomeric form of RAD51 (Chen *et al.*, 1999a;Davies *et al.*, 2001;Pellegrini *et al.*, 2002;Shin *et al.*, 2003). An alternative hypothesis to explain increased BRC repeat number in some cells, is that high numbers of BRC repeats ensure a greater abundance of RAD51 at the sites of DSBs (Pellegrini *et al.*, 2002). This hypothesis might appear more plausible in the light of recent evidence, which show that BRCA2 functions in a more complex manner than to simply sequester RAD51: studies have demonstrated that RAD51 binding not only occurs at the BRC repeats, but also through non-BRC sequences in both mammals (Davies and Pellegrini, 2007;Esashi *et al.*, 2007)

and *C. elegans* (Petalcorin *et al.*, 2007). Moreover, the binding at these non-BRC repeat sequences has been shown to be specific for RAD51 filaments, not monomers, which occurs at the BRC repeats (Davies and Pellegrini, 2007;Esashi *et al.*, 2007;Petalcorin *et al.*, 2007). In fact, some *in vitro* studies suggest that isolated BRC repeats can actually bind RAD51 filaments without causing disruption (Galkin *et al.*, 2005), but it could be argued that the conditions under which these results were found are quite dissimilar to those *in vivo*. Further evidence pointing towards the BRC repeats playing an active role in RAD51 recombination comes from studies demonstrating that a polypeptide from *H. sapiens* BRCA2 spanning all 8 BRC repeats can promote RAD51 strand exchange (Shivji *et al.*, 2006), and a fusion of the BRC repeats with RPA, from either *H. sapiens* (Saeki *et al.*, 2006) or *U. maydis* (Kojic *et al.*, 2005) can function in DNA repair and recombination. In general, the above data suggest that BRCA2 may be an active participant in homologous recombination, co-ordinating the binding of the recombinase to damaged DNA (see figure 1.14). However, the details of this are still to be classified. A final possibility for increased BRC repeat number could lie in adaptation of some BRC repeats to distinct functions, either altering the nature of their interaction with RAD51 or allowing binding to other factors. This could be consistent with the variability in BRC repeat sequence in most organisms (Lo *et al.*, 2003), and the lack of observed RAD51 binding to 2 of the mammalian BRC repeats (Wong *et al.*, 1997;Chen *et al.*, 1998b). Indeed, it appears that one or more of the BRC repeats in *Arabidopsis thaliana* bind DMC1, the meiosis-specific homologue of RAD51 (Siaud *et al.*, 2004;Dray *et al.*, 2006). Little work, however, has explored this possibility.

Exceptions to the general theme of fewer BRC repeats for simple organisms and a larger number for more complex organisms do exist, as observed most notably in the multi-cellular eukaryote *C. elegans*, which has been shown to contain just a single BRC repeat (Lo *et al.*, 2003). This therefore demonstrates the important fact that multi-cellular organisms can and do function perfectly well with homologues of BRCA2 containing just a single BRC repeat, arguing against the hypothesis that large numbers of BRC repeats are needed for efficient homologous recombination in these organisms. Other examples that do not adhere to this theme come from the single celled organisms *Trichomonas vaginalis, Plasmodium falciparum* and *Toxoplasma gondii,* which possess 14, 6, and 8 BRC repeats respectively. The reasons for these exceptions remain unclear and no experimental work has examined these proteins, but in the case of *T. vaginalis*, it could be speculated that the large expansion of BRC repeats are due to its large genome size (Carlton *et al.*, 2007).

The BRC repeat number was investigated in a number of *T. brucei* strains and subspecies, as well as other *Trypanosome* species, through MVR-PCR and Southern analyses (sections 3.8.1.1 and 3.8.2). The *T. b. brucei* strain Lister 427, was found to possess 2 allelic variants of BRCA2, one of which contained 12 BRC repeats and the other with 10 BRC repeats. Surprisingly, both alleles of BRCA2 in the *T. b. brucei* genome sequence strain, TREU927, were found to only contain 12 BRC repeats (3 less than the expected number), as did a further *T. b. brucei* strain, EATRO795. In the *T. brucei* subspecies *T. b. rhodesiense* and *T. b. gambiense*, the BRCA2 homologues contained a lower number of BRC repeats; a larger allelic variant in both subspecies containing 8 BRC repeats and a smaller one possessing 5 and 6 BRC repeats, respectively. Sequencing the BRC repeat region in BRCA2 from 2 *Trypanosoma* species that undergo antigenic variation and belong to the same salivarian clade as *T. brucei* (Cortez *et al.*, 2006) was also performed. This revealed the presence of just a single BRC repeat in *T. vivax*, and 2 very similar BRC repeats in the *T. congolense* strain TREU1457, rather than the predicted 3 from the genome sequencing strain, IL3000. Similar analysis performed in the *T. brucei* strains and subspecies demonstrated that all but the most C terminal BRC repeat were virtually identical at the nucleotide level (< 1 bp change per repeat). Significantly, this is not observed in most other BRCA2 homologues, demonstrated most notably in mammals, which possess 8 non-identical BRC repeats (Lo *et al.*, 2003), at least 6 of which bind RAD51 in the human protein (Wong *et al.*, 1997;Chen *et al.*, 1998b;Marmorstein *et al.*, 1998). Furthermore, the most C terminal BRC repeat in all the *T. brucei* strains investigated appears to be a degenerate copy, identical in all but the last 11 amino acids, but, in common with each of the upstream repeats, is predicted to encode a BRC peptide that can bind RAD51, based on extensive sequence comparisons by Lo *et al* (Lo *et al.*, 2003). Sequencing of the *T. brucei* BRC repeats revealed another structural deviation in BRCA2: all the repeats are present in a tandem array, with each repeat separated by inter-repeat spaces of identical size and sequence. Again, this appears to be unique. In all other organisms with multiple BRC repeats, they are unevenly dispersed in the BRCA2 sequence, and do not represent such a tandem array.

Taken together, the above results display that the BRC repeat number is highly variable between the different strains and subspecies of *T. brucei*, but is notably higher than BRCA2 orthologues in closely related kinetoplastid parasites. One possible explanation for this BRC repeat expansion in *T. brucei* BRCA2 is due to an intrinsic ability of the organism to expand copies of genes and mini-satellite repeats. For example, the *T. brucei* genome is known to contain a number of multigene families, such as polymerase κ (El

Sayed *et al.*, 2005a). However, this explanation cannot be valid since multigene families are also observed in *T. cruzi* and *L. major*, both of which contain low numbers of BRC repeats in their BRCA2 homologues (El Sayed *et al.*, 2005a;Lo *et al.*, 2003). Given this result, plus the finding that the BRC repeat organisation in *T. brucei* exists in a tandem array of repeats that are virtually identical in sequence, it seems likely that the BRC repeat expansion in *T. brucei* is a result of a recent evolutionary adaptation. The variation in BRC repeat number documented here most likely occurs as a result of array expansion and contraction due to the high sequence homology of the BRC repeats. Importantly, the BRC repeat number in some of the *T. brucei* strains examined here is greater than has been described anywhere else in nature, with the possible exception of *T. vaginalis* (though this requires experimental verification).

7.3 BRCA2 regulates DNA repair and recombination in T. brucei

In order to determine if *T. brucei* BRCA2 functions in DNA repair, the sensitivity of the mutants to induced DNA damage was examined. Initially, a cloning assay was used to examine the growth of the cells in the presence of the S_N2 alkylating agent methyl methanesulphonate (MMS). Consistent with a role of such damage, *brca2-/-* mutants were found to be significantly more sensitive to MMS than either WT or *BRCA2+/-* cells, similar to findings for *T. brucei* mutants of RAD51 (McCulloch and Barry, 1999), 2 RAD51 paralogues (Proudfoot and McCulloch, 2005) and sirtuin factors (Garcia-Salcedo *et al.*, 2003;Alsford *et al.*, 2007), but unlike *T. brucei* mutants of MRE11, which displayed no such level of sensitivity to MMS (Robinson *et al.*, 2002). In order to be able to quantify these effects, the IC50s were determined by measuring the metabolic capacity of the cells over a range of MMS concentrations using Alamar blue as an indicator (Raz *et al.*, 1997). The results displayed that the *brca2-/-* cells were around 3-fold more sensitive to MMS than either the WT or *BRCA2+/-* cell lines. Importantly, when BRCA2 was re-expressed in a *brca2-/-* mutant cell line, the sensitivity to MMS was reverted, but astonishingly, this cell line demonstrated an approximate 2-fold resistance to MMS compared with either the WT cells or heterozygous mutants. The reason for this level of resistance to MMS has been unable to be established, primarily due to a failure to assess expression levels. Indeed, it could not be determined if this difference was due to an increase in BRCA2 abundance due to expression from the tubulin array, rather than the endogenous locus, or if this difference was due to a secondary mutation which could have occurred spontaneously during continuous culture of the *brca2-/-* mutant cell line.

In order to determine if BRCA2 acts to repair a range of DNA damage, the *brca2-/-* mutants' sensitivity to phleomycin, a compound that causes DNA double strand breaks (Giloni *et al.*, 1981), was also assessed. Again, the *brca2-/-* mutants were discovered to be significantly more sensitive to phleomycin than WT cells, displaying approximately 5-fold greater sensitivity, similar to findings from MRE11 mutants in *T. brucei* (Robinson *et al.*, 2002). Consistent with the MMS results, the re-expression of BRCA2 also caused an increased level of resistance to be observed. However, an unexpected finding was obtained from the *BRCA2+/-* cell lines, which displayed a small but significant increase in sensitivity compared with WT cells, indicating a level of haploinsufficiency for BRCA2, which has not previously been observed for any other *T. brucei* factor that promotes homologous recombination (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005;Robinson *et al.*, 2002). This result may indicate that BRCA2 abundance is important in *T. brucei* repair efficiency on this form of damage, which may be relevant for observations that *BRCA2+/-* and *brca2-/-* mutants display progressively increased genome rearrangements compared with WT cells following prolonged growth in culture (see below). This finding also demonstrates that distinct sensitivities are observed for mutants of BRCA2 following MMS and phleomycin-induced damage. Further evidence for this phenomenon was established with results from the BRC-RPA fusion protein, which was found to complement the *brca2-/-* deficiencies to a greater extent following MMS damage than following phleomycin damage (section 5.5.3). In addition to this, further support for distinct sensitivities to these two genotoxic agents comes from investigations in *mre11-/-* mutants, which reported no sensitivity to MMS and hypersensitivity to phleomycin (Robinson *et al.*, 2002). These findings may be due to the different mechanisms by which each agent yields lesions in DNA, and may indicate subtly different modes of repair. For example, it is known that phleomycin directly causes DNA breaks, including DSBs (Giloni *et al.*, 1981), so the importance of BRCA2 in regulating RAD51 availability or directing RAD51 strand exchange may be greater than compared with MMS, which probably indirectly leads to such lesions through the action of BER (Lindahl and Wood, 1999). The possibility also exists that MMS damage primarily affects replication fork progression, and evidence has suggested that BRCA2 functions in stabilising stalled replication forks (Lomonosov *et al.*, 2003), in part at least by controlling the mobilisation of RAD51 (Yu *et al.*, 2003).

Following the above findings, the role of the BRC repeat array was tested by examining the efficiency of DNA repair in *brca2-/-* cells expressing BRCA2 variants with a single BRC repeat (section 5.3.3). Each variant was found to cause an increased level of sensitivity to both genotoxic agents tested. Furthermore, this level of sensitivity was found

to be similar to that of the *brca2-/-* mutants. Though it has not been demonstrated directly, this is consistent with the suggestion that the role of BRCA2 in DNA damage repair is through its influence on RAD51. Moreover, this finding suggests that the BRC repeat expansion is an important determinant of DNA repair efficiency, whatever its evolutionary basis. In support of this, it is important to note that for both genotoxic agents, the level of increased sensitivity of the *brca2-/-* mutants was comparable with *rad51-/-* mutants.

In order to examine the contribution of BRCA2 to *T. brucei* homologous recombination, a transformation efficiency assay was utilised. In this assay, cell lines were electroporated with a linearised tub-*HYG*-tub plasmid, which targets a hygromycin resistance gene to the tubulin array, replacing an α-tubulin ORF by homologous recombination. The results demonstrated that *brca2-/-* mutants were 12.5 to 22.5 fold less efficient at incorporating this plasmid into its genome than either WT or heterozygous cell lines (section 4.3.5). As for the DNA repair assays, these results were highly reminiscent of those previously obtained for *rad51-/-, rad51-3-/-, rad51-5-/-* and *mre11-/-* mutants, all of which have demonstrated a role for their respective proteins in *T. brucei* homologous recombination (McCulloch and Barry, 1999;Conway *et al.*, 2002c;Robinson *et al.*, 2002;Proudfoot and McCulloch, 2005). When BRCA2 was re-expressed in a *brca2-/-* mutant cell line, transformation efficiency results were obtained that were comparable with WT and heterozygous cell lines, completely reverting the integration defect observed in the absence of BRCA2. In contrast, despite slightly higher transformation efficiency rates than the *brca2-/-* mutants, none of the variant BRCA2 proteins were able to function as efficiently as WT, *BRCA2+/-* or *BRCA2-/-/+* cell lines in this assay (sections 5.3.4 and 5.5.4), indicating the requirement of the full length BRCA2 for efficient homologous recombination in *T. brucei*, at least as measured by this assay. The impairment of recombination in BRCA2 proteins with a single BRC repeat, considered along with the data displaying that a reduction in BRC repeat number leads to a reduced ability to repair DNA damage, reinforces that the BRC repeat expansion in *T. brucei* BRCA2 is critical for both general DNA repair and homologous recombination. Again, it seems likely that this is due to an impaired interaction with RAD51, and though this was not demonstrated directly it is supported by the absence of detectable RAD51 nuclear foci following phleomycin induced damage (see below). A surprising result is that the *BRC+RPA BRCA2-/-/+* cell line could also not support efficient homologous recombination, given that similar proteins in mammalian cells (Saeki *et al.*, 2006) and in *U. maydis* (Kojic *et al.*, 2005) allowed for efficient DNA repair and recombination, and that the *T. brucei* BRC+RPA fusion functioned in DNA repair (section 5.5.3). These differences could be accounted for by a number of different factors (see below). This work also shows that the

BRC repeat domain in isolation is incapable of supporting DNA recombination or repair, presumably because it is unable to bind to DNA, at least *in vivo*.

Further analysis into the mechanisms of homologous recombination by Southern blotting demonstrated that in the clones where *brca2-/-* mutants had succeeded in incorporating the plasmid, they had done so *via* homologous recombination. The lack of aberrant integrations was unexpected, since these had been observed in *rad51-/-* mutants (Conway *et al.*, 2002c). Nevertheless, these results demonstrate that BRCA2 acts in homologous plasmid integration in *T. brucei*. Given the broad conservation of homologous recombination functions in *T. brucei* and other kinetoplastids (El Sayed *et al.*, 2005a), it seems likely that BRCA2 most likely contributes to *T. brucei* DNA repair through its role in recombination. It is interesting to note that BRCA2-independent pathways also exist. The nature of such pathways, and whether they occur *via* RAD51 in the absence of BRCA2, has not been examined.

This work has provided a clearer view of the factors involved in *T. brucei* DNA repair and recombination. Indeed, it is emerging that the machinery of *T. brucei* appears to be remarkably similar to that of higher eukaryotes, involving a number of regulating factors including RAD51, RAD51-3, RAD51-5, MRE11 and BRCA2. Conversely, it is also becoming apparent that the *T. brucei* machinery looks less like *S. cerevisiae*, which has no BRCA2, and only RAD57 and RAD55, and indeed *U. maydis*, which only has one RAD51 paralogue (Kojic *et al.*, 2006). The role of BRCA2 could be speculated to be of more importance than other regulating factors such as the RAD51-paralogues due to the fact that in the absence of BRCA2, RAD51 foci fail to form, possibly indicating an inability to deliver RAD51 to these sites or a lack of stability (see below). One important step in the homologous recombination pathway is in the removal of RPA from ssDNA, thereby allowing the RAD51 nucleoprotein filament to form. In mammals, BRCA2 can provide the role of RPA displacement from ssDNA, as can RAD52 (Martin *et al.*, 2005;Yang *et al.*, 2002;Sung, 1997a;Benson *et al.*, 1998). However, RAD52 appears to be absent from the *T. brucei* genome (El Sayed *et al.*, 2005b), implying that BRCA2 would be the only protein to perform this role, and therefore highlighting its importance. However, since a number of RAD51 paralogues have been shown to exist in *T. brucei,* and their functions have not yet been elucidated, the possibility exists that in the absence of BRCA2 one or more of these proteins could perform this role.

7.4 T. brucei BRCA2 acts in antigenic variation

The analysis of antigenic variation in *T. brucei BRCA2* mutants revealed that BRCA2 acts in VSG switching. Indeed, VSG switching frequencies were found to be 8 to 11 fold lower in the *brca2-/-* mutants compared with WT, *BRCA2+/-* or *BRCA2-/-/+* cells, a level of impairment that was highly comparable to results previously obtained for *RAD51* and *RAD51-3* mutants (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005). This confirms the importance of homologous recombination in the process of VSG switching. Further analysis of VSG switched variants revealed that gene conversion and transcriptional switching events still occurred in the absence of BRCA2. This result is again reminiscent of the findings from *RAD51* and *RAD51-3* mutants (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005), indicating that the reduced VSG switching frequencies arose due to impairment of both pathways, raising the question as to whether these reactions are enzymatically and mechanistically distinct (Proudfoot and McCulloch, 2005). Though we still do not know the details of the VSG switching mechanism, this accumulated data suggest that the strand exchange step is critical, as it is likely that each protein contributes to it.

Surprisingly, the variants of BRCA2 with just a single BRC repeat (*T. vivax BRCA2-/-/+* and *1BRC BRCA2-/-/+*) remained capable of switching their VSG coat compared with WT and *BRCA2-/-/+* cell lines. These results appear to indicate that the BRC repeat expansion in *T. brucei* BRCA2 is of little importance for VSG switching efficiency during an acute infection. This unexpected result goes against the original hypothesis that the BRC repeat expansion in *T. brucei* BRCA2 is due to the high levels of antigenic variation, which relies upon homologous recombination. Nevertheless, a number of arguments can be made in support of the original hypothesis. Firstly, the VSG switching analysis was performed in a monomorphic cell line, which undergoes VSG switching at rates of only ~1 x 10^{-6} switches per cell per generation (Lamont *et al.*, 1986), much lower than the pleomorhic cell lines where high levels of switching are observed (1 x 10^{-2} switches per cell per generation) (Turner and Barry, 1989;Turner, 1997), and may not therefore be representative of VSG switching frequencies in pleomorphic cell lines containing reduced numbers of BRC repeats. However, despite this possibility, it is worth noting that the numbers of BRC repeats in BRCA2 proteins characterised in monomorphic and pleomorphic cell lines was not found to significantly differ, as might be expected if BRC repeat number contributed to VSG switching frequency. Indeed, Lister 427 was found to contain among the highest number from those characterised, indicating that the selective pressures for a high number of BRC repeats still remained in low switching cell lines. Another argument in support of

the original hypothesis arises due to the assay used for measuring VSG switching frequency, which only examines switches that occur during a single relapse peak. This is likely, therefore, to represent only the switch mechanisms used early on in an infection, when telomere proximal and intact array genes are primarily activated (Pays, 1989;Morrison *et al.*, 2005;Marcello and Barry, 2007b). Later on in infections, *VSG* pseudogenes become the preferred substrates for VSG switching, utilising mosaic gene formation (Marcello and Barry, 2007b;Thon *et al.*, 1990). The possibility therefore exists that the BRC expansion in BRCA2 is required specifically for these later reactions, and has no bearing on activation of intact genes.

Despite the above arguments, other observations argue against the hypothesis that the BRC repeat expansion is due to a requirement for VSG switching. These arise from the closely related trypanosomatids *T. congolense* and *T. vivax*, which despite only containing low numbers of BRC repeats, are still capable of surviving in mammals by undergoing VSG switching. Quite how their homologues of BRCA2 support VSG switching is currently unknown, due largely to a lack of research regarding the switching mechanisms utilised by these trypanosomatids. However, gene duplication mechanisms have been documented to occur in *T. congolense* (Majiwa *et al.*, 1985). More perplexing, however, is how BRCA2 can support VSG switching in *T. brucei*, with just a single BRC repeat, given that those variants are demonstrated to be impaired in DNA repair and homologous recombination. One possibility is that there are differing requirements between antigenic variation and the general repair and recombination mechanisms. For example, antigenic variation might not require the extensive interactions between BRCA2 and RAD51 through the BRC repeats that the general recombination mechanisms appear to do. Quite what these mechanisms would be, or how they might function, are as yet unknown. However, this perhaps gains support from research into MRE11 and RAD51-5, which despite functioning in repair and recombination in *T. brucei*, appear not to regulate VSG switching (Robinson *et al.*, 2002;Proudfoot and McCulloch, 2005), consistent with VSG switching utilising a sub-pathway of homologous recombination. A second possibility for the functioning of BRCA2 with just a single BRC repeat in *T. brucei* VSG switching is that functional differences are provided by the large number of identical BRC repeats and the downstream, divergent repeat. For example, the upstream, identical BRC repeats may provide the interactions with RAD51 which direct general DNA repair and recombination mechanisms. The downstream BRC repeat on the other hand, could have diverged for a specific role involved in antigenic variation. However, this theory does not simply account for how *T. vivax* BRCA2, containing a single BRC repeat, would support VSG switching in *T. brucei*, and it is not clear what aspect of general DNA repair/recombination might

underlie the selection for BRC repeat expansion. Conceivably, it is possible that distinctions between DNA recombination and antigenic variation are absent in this trypanosomatid, and indeed also in *T. congolense*, since we do not know if the mechanistic pathways of antigenic variation are equivalent in all trypanosomatids.

Finally, it should be noted that a BRC repeat expansion in protists is not limited to *T. brucei*. This also appears to be a common feature to Apicomplexans, with BRCA2 homologues in all *Plasmodium* species containing 6 BRC repeats and *Toxoplasma gondii* possessing 8 BRC repeats (Lo *et al.*, 2003). Whether or not these expansions in BRC repeat number share a functional basis with *T. brucei* is unknown. However, it could be speculated that this is indeed the case, since *Plasmodium falciparum* is also known to extensively utilise antigenic variation, though it is based exclusively on transcriptional switching between *var* genes (Kyes *et al.*, 2007). This seems unlikely to be the case in *Toxoplasma gondii* however,: despite containing a BRC repeat expansion in BRCA2, and antigenically distinct strains being documented, antigenic variation has not previously been documented in this parasite (Delibas *et al.*, 2006;Ajioka *et al.*, 1998;Hettmann and Soldati, 1999).

7.5 RAD51 focus formation requires BRCA2 in T. brucei

Previous research has demonstrated that in *T. brucei*, RAD51 re-localises to sub-nuclear foci following phleomycin treatment and the induction of DNA double strand breaks (Proudfoot and McCulloch, 2005;Glover *et al.*, 2008). Indeed, this process has been shown to occur in most eukaryotes and is controlled by a number of factors (Tarsounas *et al.*, 2004;Lisby and Rothstein, 2004), amongst which BRCA2 is critical (Tarsounas *et al.*, 2003;Yu *et al.*, 2003;Kojic *et al.*, 2005;Martin *et al.*, 2005). In order to determine if this was also the case in *T. brucei*, RAD51 localisation was examined by immunofluorescence, before and after treatment with phleomycin. Consistent with previous work, very few RAD51 foci were detectable before induced DNA damage, but treatment with 1.0 µg.ml^{-1} phleomycin allowed RAD51 foci to be detected in more than 75 % of WT and *BRCA2+/-* cells. *brca2-/-* mutants, however, were either unable to induce RAD51 re-localisation or to maintain foci once established, indicating a critical role in this response. These results may support the hypothesis that RAD51 is unable to be transported to the sites of DNA damage without the presence of BRCA2 *in vivo*. This, in turn, would support the hypothesis that BRCA2 is the enzyme that sequesters RAD51 until it is required for DNA repair, and transports it to the sites of DNA damage (Tarsounas *et al.*, 2004). However, the mobility of RAD51 in *T. brucei* has not been examined, meaning that it is possible that RAD51 is

able to be transported to the sites of damage in *brca2-/-* cells, but is unable to be retained, which is quite different from the sequestration/transportation hypothesis. Though BRCA2 has not been suggested to have a role in RAD51 filament stabilisation, it cannot be discarded.

Of the variant proteins that were expressed in *brca2-/-* cells, RAD51 foci were only detectable with the BRC+RPA fusion. This result was somewhat surprising given the impairment of homologous recombination in this cell line. However, the percentage of RAD51 foci detected in this cell line was not equivalent to WT cells, indicating that these cells were less efficient at performing this process. This result correlates with the partially increased sensitivity to phleomycin, indicating that although this fusion allows for RAD51 foci formation, it does not do so as efficiently as the full length protein. However, this finding contradicts similar research in mammalian cells (Saeki *et al.*, 2006) and in *U. maydis* (Kojic *et al.*, 2005), whereby a fusion of the BRC repeats to RPA, not only allowed for efficient DNA repair and RAD51 relocalisation, but also for efficient homologous recombination. Quite why these differences have been observed is unclear, but could possibly be due to differences in the assays used, the distinct activities between the BRCA2 proteins, or could reflect that in *T. brucei*, the fusion protein may not be folded in such a way to allow it to perform these functions as effectively. A more far reaching explanation could be that DNA repair is somewhat diverged in *T. brucei* relative to characterised eukaryotes.

The variants containing a reduced number of BRC repeats displayed similar deficiencies in RAD51 foci formation to the *brca2-/-* mutants. This result was quite surprising, since in the *1BRC BRCA2-/-/+* cell line, the retained BRC repeat is predicted to retain all of the critical residues required for RAD51 interaction (Lo *et al.*, 2003). However, it is conceivable that the other changes compromise the strength of the interaction. Another possibility could be that the *T. brucei* BRCA2 protein possesses a BRC repeat expansion in order to allow it to efficiently transport RAD51 to sites of DNA damage. Indeed, this could occur by allowing the RAD51 filament to form before transported to the sites of damage or to simply transport greater quantities of RAD51. A reduction in BRC repeats could therefore lead to significant amounts of unbound RAD51 and inefficient delivery of RAD51 to sites of DNA damage. Each of these explanations could be tested by future experiments that examine the amount of BRCA2, bound and unbound to RAD51, and the mobility of RAD51 in the cell.

7.6 Loss of BRCA2 causes gross chromosomal rearrangements

Mutation of BRCA2 in mammalian cell lines has been shown to cause an accumulation of gross chromosomal rearrangements, which include breakage, translocations and chromosome loss (Patel *et al.*, 1998;Yu *et al.*, 2000). Similar findings have been demonstrated in *U. maydis* (Kojic *et al.*, 2002). In order to determine if BRCA2 is also a regulator of genomic stability in *T. brucei*, the wild type and *BRCA2* mutants were cultured *in vitro* for ~ 290 generations before being re-cloned. A number of these clones were subsequently analysed by Pulsed Field Gel electrophoresis (PFGE) and Southern blotting (section 4.3.9).

Even simply by ethidium bromide staining of the gels, a number of karyotype differences could be identified in the *brca2-/-* clones compared with the WT and *BRCA2+/-* clones. Indeed, a reduction in chromosome size was observed in the majority of cases, indicating the occurrence of GCRs. Another important observation came from the *BRCA2+/-* clones, which, although not as severe as the *brca2-/-* clones, displayed small reductions in chromosome size, possibly indicating haploinsufficiency. In fact, this result appears compatible with the increased sensitivity of *BRCA2+/-* mutants to phleomycin induced DNA damage (section 4.3.4). Probing the PFGs with *VSG121* (a five-gene family), *VSG221* (a single copy *VSG* in the active *VSG* ES) and *GPI* (a single-copy gene encoding glucose 6-phosphate isomerase) appeared to confirm these observations, with severe size changes (~ 500 kb) being found in three *brca2-/-* clones when probed with *VSG221*. Probing with this gene also revealed an increase in chromosome size for two *brca2-/-* clones. As a result of these findings, an important area to address was that the size differences observed did not result from a VSG switching event. In order to address this possibility, the VSG being expressed was investigated by western blot analysis, and revealed that all of the clones investigated were still expressing VSG221 (section 4.3.9), despite being cultured for a considerable length of time.

VSG121 was found to hybridize to two chromosomes of approximately 2.1 and 2.3 Mb, both of which appeared to be smaller in all of the *brca2-/-* clones when compared to the WT cells (up to 100 kb). In support of haploinsufficiency for BRCA2, the same chromosomes also appeared to have reduced in size in two to three of the *BRCA2+/-* clones, though to a lesser extent. Further investigations into this five gene family through a Southern blot of *Xmn*I-digested genomic DNA from all of the clones, probed with *VSG121*, revealed that the chromosomal changes observed in the PFGs was due to loss of

genetic material. Indeed, 11 of the 12 *brca2-/-* clones had lost at least one copy of *VSG121*. Interestingly, the telomeric copy of *VSG121* was never lost, indicating a level of stability for telomeric sequences. Further evidence for this comes not only from the fact that all the clones continued to express VSG221, but also from *mre11-/-* mutants, which were also seen to undergo GCRs following prolonged *in vitro* culture, but were also never found to lose their telomeric copies of *VSG121* (Robinson *et al.*, 2002).

Undoubtedly, this finding that loss of BRCA2 in *T. brucei* causes GCRs to occur in the megabase chromosomes of the genome, should not have come as a surprise due to similar results being found in vertebrates (Patel *et al.*, 1998;Yu *et al.*, 2000) and *U. maydis* (Kojic *et al.*, 2002). However, what did appear unusual was the observation that these GCRs only appeared to affect megabase chromosomes and not the intermediate- or the mini-chromosomes, which contain mainly *VSG* and *VSG* ES sequences (Wickstead *et al.*, 2004;Rudenko *et al.*, 1998;Melville *et al.*, 2000). Indeed, the karyotype appeared to be relatively stable among WT, *BRCA2+/-* and *brca2-/-* cell lines, with no notable differences being detected. However, it could be argued that rearrangements may have been occurring in the smaller chromosomes but were unable to be detected due to the lack of separation in the PFGs. However, an alternative explanation comes from the fact that the megabase chromosomes and the smaller chromosomes replicate and divide at different points of the cell cycle. Indeed, the mini chromosomes have been shown to replicate and segregate earlier in the cell cycle than the megabase chromosomes (Gull *et al.*, 1998). It is not known if this is also true of intermediate chromosomes, though they share DNA sequences with the minichromosomes (Wickstead *et al.*, 2003a). The megabase chromosomal aberrations could therefore be explained by deficiencies in separating the chromosomes during mitosis, which may be related to the observations that *brca2-/-* mutants appear to have difficulty in completing nuclear division (see below). If the minichromosomes are replicated and segregated earlier than the megabase chromosomes, they could avoid the mis-segregation that is observed in the larger chromosomes. How this would be manifest as chromosome size reduction, rather than wholescale chromosome loss, is unclear, however. Equally, if it were true, we might have expected to see cell cycle abnormalities in the *BRCA2+/-* cells, which we do not.

A simpler explanation for the megabase chromosome-specific GCRs, is that they arise due to a predominance of changes in sequences found in the megabase chromosomes and not in the smaller chromosomes. Taking the PFGE results together with the loss of *VSG121* gene copies, it appears that GCRs arising in *T. brucei brca2-/-* cells may result primarily from deletions within the sub-telomeric *VSG* arrays. Indeed, each of these phenotypes are

highly reminiscent of the GCRs displayed in *mre11-/-* mutants, where the chromosome size changes were primarily due to sequence loss and were only seen in the megabase chromosomes (Robinson *et al.*, 2002). It is possible that this indicates a shared function of the proteins in the maintenance or use of subtelomeric *VSG* arrays. Equally, however, it is possible that this simply represents their roles in genome stability. It might be argued that the phenotype observed in the *mre11-/-* mutants was more severe than is displayed here for the *brca2-/-* mutants, but it is important to note that these phenotypes should not be compared directly. The main reason for this is due to the number of generations the clones were cultured for before GCRs were investigated. The *mre11-/-* clones were investigated at 550 generations, whilst the *brca2-/-* clones were investigated at 290 generations. Therefore, in order to directly compare these phenotypes, the *brca2-/-* mutants should be cultured until 550 generations and GCRs subsequently investigated. Similar work has not been done in a broad spectrum of DNA repair genes, including *RAD51*. It would be informative to ask if such GCRs represent a specific function of BRCA2 and MRE11, or represent general activities of HR enzymes. Clearly, this would be important in understanding the mechanisms for *VSG* repertoire evolution, which may be very rapid (Callejas *et al.*, 2006).

7.7 The role of BRCA2 in cell cycle progression

Potentially, one of the most interesting findings regarding *T. brucei* BRCA2 comes from the analysis into the cell cycle, which was investigated in order to examine the reason for the level of growth impairment observed in the *T. brucei brca2-/-* mutants. This was investigated through examination of the DNA content of individual cells by DAPI staining. Despite no evidence being uncovered for a cell cycle stall, a perhaps surprising result was obtained, which suggested that the *brca2-/-* mutant population contained cell cycle abnormalities compared with WT and *BRCA2+/-* cell lines. Specifically, the *brca2-/-* mutant population was found to contain a lower percentage (~ 10 %) of cells in G1 or S phase (1N1K) of the cell cycle (McKean, 2003), which was accounted for by an accumulation of cells that did not conform to any of the 'normal' cell cycle stages, and so were described as being aberrant cell types, or 'others'. A more detailed examination of these cell types revealed approximately equal numbers of cells with raised nuclear DNA and kinetoplast DNA content, with the cell types most commonly observed containing 0N1K, 0N2K, 1N3K (raised kinetoplast DNA), 1N0K, 2N1K or 2N0K (raised nuclear DNA). Remarkably, despite the extent of GCRs that had accumulated in the *brca2-/-* mutants before the re-introduction of BRCA2, the *BRCA2-/-/+* cell line was able to progress through the cell cycle without accumulating a significant number of aberrant cell

types. This perhaps suggests that the generation of GCRs and aberrant cell types are distinct phenotypes of the *BRCA2* mutant (discussed further below).

The main reason why this phenotype observed in the *brca2-/-* mutants was regarded as being a surprising one, was due to the fact that similar findings had not previously been observed in *T. brucei* mutants of other DNA repair/recombination factors that resulted in a growth impairment of very similar magnitude. In all, mutation of MRE11 (Robinson *et al.*, 2002) the RAD51 paralogues, RAD51-3 and RAD51-5 (Proudfoot and McCulloch, 2005), and (most notably) RAD51 (this work) were all examined for cell cycle phenotypes, and none displayed them in the absence of induced DNA damage. This observation therefore indicates that this phenotype does not simply result from the DNA repair deficiency of the *brca2-/-* cells, and is consistent with the possibility that *T. brucei* BRCA2 has a role beyond the simple regulation of RAD51-catalysed recombination, in either the regulation or execution of cell division.

Further evidence comes from the investigation into DNA damage sensitivities of *T. brucei* DNA repair mutants. Were the cell cycle phenotypes of *brca2-/-* mutants to reflect a greater function in DNA repair, then it could be imagined that a more significant sensitivity to DNA damage would be observed for the *brca2-/-* mutants. In fact, highly comparable levels of sensitivity to both MMS and phleomycin were obtained between *brca2-/-* and *rad51-/-* mutants, whilst only the former displayed any cell cycle differences. Furthermore, if this was indeed a valid argument, then the induction of further DNA damage should result in amplification of this phenotype. However, a distinct phenotype was observed following phleomycin treatment of *brca2-/-* mutants. This phenotype consisted not only of an increase in the percentage of cells containing 1N2K, but the distribution of aberrant cells contained a pronounced number of cells with raised kinetoplast DNA content. In addition, similar results following phleomycin treatment were observed in the *rad51-/-*, *rad51-3-/-* and *rad51-5-/-* mutants. Taken together, it appears that the induction of DNA damage through phleomycin treatment causes a delay in nuclear DNA replication, but does not block cell division, meaning that daughter cells are generated that lack nuclear DNA.

In order to analyse the reasons for the accumulation of aberrant cell types in *brca2-/-* mutants, the cells in M phase, which were about to undergo cytokinesis (containing a DNA content of 2N2K) were analysed. In WT, *BRCA2+/-* and *BRCA2-/-/+* cell lines, the majority of cells were found to contain 2 clearly separated nuclei, with only a small percentage (10-15 %) still segregating. The *brca2-/-* mutants, however, contained a larger

percentage (30-40 %) of cells with visibly connected nuclei, which were still undergoing segregation. This phenotype is again a BRCA2-specific phenomenon since the *rad51-/-* mutants did not exhibit this phenotype. Furthermore, this discovery seems to provide an explanation for the increase in the number of aberrant cell types, since in the absence of BRCA2, cells do not undergo a cell cycle stall, but proceed into cytokinesis, often whilst nuclear segregation is still occurring, thereby resulting in daughter cells that inherit either both nuclei, or none, thus accounting for the pattern of DAPI staining in the aberrant cells.

A final, compelling, piece of evidence indicating that DNA repair deficiency is not involved in this increase in number of aberrant cell types come from the BRCA2 variant expressers. The BRC+RPA expresser, despite being quite proficient in DNA repair still displayed an accumulation of aberrant cell types equivalent to the *brca2-/-* mutants. Conversely, the cell lines expressing BRCA2 with just 1BRC repeat, which were deficient in DNA repair, no longer displayed this accumulation of aberrant cells, indicating that BRC repeat number does not affect this phenotype. Finally, a remarkable result was observed in the *C term BRCA2-/-/+* cell line, whereby the number of aberrant cell types, whilst not being as low as observed in WT or *BRCA2-/-/+* cells, was significantly lower than either the BRC+RPA cell line or the cells expressing just the BRC repeat polypeptide. This result indicates that the C terminus of BRCA2, isolated from the BRC repeat, can partially complement the replication or cell division deficiency observed in the *brca2-/-* mutants. This implies that the replication or cell division deficiency phenotype is a consequence of a function that can be separated from the BRC repeats, and appears to reside within the C terminal region of BRCA2.

The basis for GCR in *brca2-/-* mutants, and in *mre11-/-* mutants, appears to reside in loss of subtelomeric sequences. However, the rearrangements can be drastic in some clones, so could these results suggest that the GCRs found in the *brca2-/-* mutants may not simply exist due to DNA repair defects? For instance, could the accumulation of aberrant cell types, resulting from an early onset into cytokinesis, before completion of DNA repair or replication, add to the chromosomal rearrangements? The lack of observable cell cycle defects in *mre11-/-* mutants appears to argue against this. However, a detailed comparison of the types of GCR has not been conducted. It is striking that the kinetoplast DNA in the *brca2-/-* mutants appears to be unaffected in the cell cycle, since normal numbers of cells with 1N2K and 2N2K DNA content are observed. This therefore appears to suggest that loss of BRCA2 does not affect kinetoplast DNA replication and segregation, but is limited to a nuclear function. Indeed, this result is in keeping with the nuclear location of BRCA2 in other organisms (Bertwistle *et al.*, 1997;Martin *et al.*, 2005;Zhou *et al.*, 2007) and the

identification of putative nuclear localisation signals in the *T. brucei* polypeptide sequence. Further work will be needed to understand the basis of GCRs in these mutants.

Despite these results providing evidence for replication or cell division abnormalities, they do not explain the reason for the increased population doubling times of the *brca2-/-* mutants. It has been established that there is no evidence for a cell cycle stall, since no increase or decrease in 1N2K or 2N2K cells is seen. In addition, the nuclear replication/segregation defect cannot provide an adequate explanation for reduced growth or cell cycle delay, since in *rad51-/-, rad51-3-/-, rad51-5-/-* and *mre11-/-* mutants, similar levels of growth impairment were observed, but no cell cycle abnormalities were found (McCulloch and Barry, 1999;Proudfoot and McCulloch, 2005;Robinson *et al.*, 2002). Finally, it was observed that expression of the BRC+RPA fusion protein reverted the population doubling time of the cells to WT rates, whilst the accumulation of aberrant cell types persisted. Taken together, it therefore seems most likely that either the *T. brucei brca2-/-* mutants take longer to complete the cell cycle, or that the DNA repair deficiency of the *brca2-/-* mutants leads to an increased rate in cell death, as is seen in other DNA repair mutants.

All the above results allow a hypothesis to be postulated, which suggest that DNA repair/recombination and DNA replication/segregation are separate functions of *T. brucei* BRCA2. Furthermore, sequence elements located in the C-terminal domain of BRCA2 are likely to function to ensure the correct transmission of nuclear DNA during *T. brucei* cell division. Undoubtedly, further investigations will be required in order to fully understand how BRCA2 contributes to the mechanisms of *T. brucei* DNA replication/segregation. However, the pattern of aberrant cells in the *brca2-/-* mutants is consistent with cytokinesis occurring before the completion of nuclear DNA segregation, yielding initially daughter cells with 2 nuclei and 1 kinetoplast (2N1K) and lacking a nucleus (0N1K). The presence of cells with further aberrant DNA contents (e.g. 1N0K, 2N0K, 2N3K and 2N4K), suggests that further rounds of replication and cell division can and do occur, despite the fact that these cells are likely to be dying.

What might be the mechanistic basis for these phenotypes? It is possible that *T. brucei* BRCA2 may function in the timing of cytokinesis, and in its absence, cells undergo cytokinesis before the completion of DNA replication and segregation. Another possibility is that *T. brucei* BRCA2 may function in efficient DNA replication and nuclear segregation, and in its absence, the completion of mitosis is delayed. The generation of mis-segregated DNA would then only occur if cytokinesis occurred under normal timing.

Either suggestion relies upon the assumption that a cell cycle checkpoint, ensuring that cytokinesis occurs only after completion of mitotic chromosome segregation, is absent from *T. brucei*, at least when BRCA2 is mutated. However, this may not be the case, since RNAi depletion of CRK3-CYC6 in bloodstream stage cells suggested that the mitosis initiation checkpoint is intact (Hammarton *et al.*, 2003). The same experiment in procyclic form *T. brucei* suggested a different scenario, whereby cytokinesis can occur in the absence of mitotic exit, or even mitosis. These results therefore point to the presence of different cell cycle checkpoints in the different life cycle stages. Nevertheless, the above possibilities for *T. brucei* BRCA2 could still remain valid due to the observations that in the presence of DNA damage, either by treatment with phleomycin (R. McCulloch, unpublished) or by inducing a telomeric DNA DSB (Glover *et al.*, 2007), cell cycle progression is not blocked in bloodstream stage *T. brucei*, implying that there may be an absence of DNA damage cell cycle checkpoints which monitor DNA integrity in *T. brucei*. Having said this, a recent study that induced a DSB at a chromosome internal locus did appear to induce a G2/M arrest, suggesting that this may not be so straightforward (Glover *et al.*, 2008).

Research on BRCA2 in mammalian cells has suggested a role for BRCA2 in cell division. Firstly, BRCA2 has been shown to interact with BRAF35 and BUBR1, two proteins which appear to modulate the initiation of mitosis through roles in chromatin condensation (Marmorstein *et al.*, 2001) and spindle attachment (Lee *et al.*, 1999;Futamura *et al.*, 2000). Secondly, BRCA2 has been suggested to localise to centrosomes (Nakanishi *et al.*, 2007). Finally, research in HeLa cells and murine embryo fibroblasts (MEFs) have found that BRCA2 localises to the cytokinetic midbody, and its disruption by RNAi or targeted gene disruption impairs or delays cytokinesis (Daniels *et al.*, 2004). It is worth noting that this study appears to suggest a distinct phenotype from that what is described in *T. brucei brca2-/-* cells. Here, the data is most readily explained by cytokinesis appearing to occur early before nuclear division has been completed, leading to the accumulation of aberrant cell types. This may suggest that BRCA2 functions quite differently in *T. brucei* and mammals, or may reflect cell cycle checkpoint differences between the parasite and host. Nevertheless, since no work to date has reported the phenotypes that are found in *T. brucei* in other organisms, including *C. elegans*, *D. melanogaster* and *U. maydis* (Martin *et al.*, 2005;Brough *et al.*, 2008;Kojic *et al.*, 2002), these findings could therefore be indicative of an evolutionary divergent role for BRCA2 within *T. brucei*. Some evidence does, in fact, point to *T. brucei* BRCA2 being linked to DNA replication rather than cell division: an interaction between *T. brucei* BRCA2 and an orthologue of CDC45 has been described (S. Oyola, PhD thesis, University of Cambridge). CDC45 functions in both the initiation and

elongation of nuclear DNA replication (Bauerschmidt *et al.*, 2007), and it is therefore conceivable that the role of BRCA2 is to link the DNA replication and repair machineries, ensuring that replication stalls are overcome. This scenario would fit in with the results described in this thesis, whereby in the absence of BRCA2, the completion of nuclear DNA replication is impaired. Whether or not this interaction is an evolutionary conservation remains yet to be seen, but it is worth noting that since the size and sequence of BRCA2 homologues from different organisms display considerable diversity (Lo *et al.*, 2003), it could be postulated that BRCA2 can adopt different roles within different organisms. Indeed, BRCA2 has already been found to interact with a number of different proteins in different organisms (Marmorstein *et al.*, 2001;Xia *et al.*, 2006b;Lu *et al.*, 2005;Dong *et al.*, 2003).

7.8 Future experiments

Despite the work detailed in this thesis providing a wealth of information in the role of BRCA2 in terms of DNA damage repair, homologous recombination and *T. brucei* antigenic variation, a number of questions remain unanswered. In order to resolve these questions, a number of experiments could be explored.

Undoubtedly, the purification of the *T. brucei* BRCA2 homologue and a number of different motifs would allow a wealth of potentially informative biochemical analyses to be performed. Not the least of these would be the confirmation of the interaction between BRCA2 and RAD51 in *T. brucei*, which has been shown to occur indirectly in this thesis, *via* the loss of RAD51 foci formation in *brca2-/-* mutants (section 4.3.6). This area of investigation would also allow the interaction between BRCA2 and RAD51 in *T. brucei* to be localised to a region of BRCA2, through the utilisation of various purified motifs of the protein. This would therefore confirm whether the BRC repeats, the C terminus or another region are capable of binding RAD51. Furthermore, the questions of whether monomeric or multimeric forms of RAD51 bind to BRCA2 could also be answered. Indeed, this could confirm if the *T. brucei* homologue of BRCA2 functions similarly to that of the *H. sapiens* protein, in that monomeric forms of RAD51 bind to the BRC repeat region of BRCA2, whilst multimeric forms bind to the C terminus (Esashi *et al.*, 2007;Davies and Pellegrini, 2007;Lord and Ashworth, 2007;Petalcorin *et al.*, 2007).

These biochemical analyses could also prove useful in identifying other BRCA2 interacting factors. For example, it is known that BRCA2 also binds the meiotic specific recombinase, DMC1, in *Arabidopsis thaliana* and *H. sapiens* (Siaud *et al.*, 2004;Thorslund

et al., 2007). Yeast 2 hybrid analysis or co-immunoprecipitation could also prove useful tools in helping to answer these questions, and potentially identify novel interacting factors.

One of the major problems of the cell lines expressing various truncations of BRCA2, is that the expression of these proteins were unable to be confirmed. Both western blot analysis and IFA proved to be unsuccessful. In order to determine if these proteins were indeed being expressed and correctly localised to the nucleus other approaches might be considered. It should be possible to tag these with other epitopes (both HA and GFP tagging has proved unsuccessful). Alternatively, purified proteins could be introduced into the cells and their localisation followed with antibodies specifically raised against them. Finally, over-expressing the proteins in order to obtain enough protein to detect by western blot or IFA could be considered. Indeed, other researchers have found that BRCA2 is difficult to detect in *T. brucei* using peptide anti-sera, unless a substantial level of over-expression is achieved (S. Oyola *pers comms*). However, the over-expression of proteins does not come without problems, including the mis-localisation of the protein, non-physiological interactions and disruption of protein complexes (Swaffield *et al.*, 1995).

Additional mutants could also be generated containing varying numbers of BRC repeats, instead of either the complete set or just the most C terminal. Indeed, these experiments were due to be performed by myself, but due to time constraints and cloning difficulties, these were unable to be generated.

In order to confirm that the modes of VSG switching can occur *via* RAD51-unrelated pathways, *rad51-/-, brca2-/-* double knockouts could be generated and the subsequent VSG switching analysis performed. One of the major obstacles in the generation of such mutants is in the restricted number of antibiotic resistance cassettes that can be utilised in *T. brucei*. One possible method to overcome this problem would be generating these mutants *via* loss of heterozygosity.

The investigation of BRCA2 is a fast paced area of research and, undoubtedly, new interacting partners are likely to be identified, along with biochemical assays providing fresh insights into the complex and perhaps multiple functions of the protein.

Appendices

Appendix 1: A list of the oligonucleotides used in this thesis.

Primer name	Sequence	Restriction sites and tags	
3'BRCVAR	CCTCTAGATGCTACTTGCAGTGACGACTC	XbaI	
3'VARPLUSTAG	AGGCCTGGTACCTGTCTAGATGCTACTTGCAGTGACGACTC	XbaI	
3'VARTAG	AGGCCTGGTACCTGTCTAGA	XbaI	
5'BRC VAR	CCGTTAACATGTACCCCTACGACGTCCCGGACTATGCCAGCCACAAAAAAGGAAGACAA	HpaI	HA
BRC probe 3'	TCCTGGCCATTTTCAGTATTC		
BRC probe 5'	AGCACTGCGGTACAAGGAAATTCC		
BRC VAR 3'3	CCCTCGCGACTATTCTCGCATAAGATCAGCGAC	NruI	
BRC VAR 3'5	AGCACTGCGGTACAAGGAAATTCC		
BRC VAR 5'3	CTCTTGGCCATTTTCAGTCCC		
BRC VAR 5'5	CCCTCGCGAATGTACCCCTACGACGTCCCGGACTATGCCAGCCACAAAAAAGGAAGACAA	NruI	HA
BRC_RPA 3'	CGGCTTTCTTGCTAGCTTGGATG		
BRC_RPA 5'	CCCCGATATCATGTACCCCTACGACGTCCCGGACTATGCCTACGGGACTGAAAATGGCCAAGAG	EcoRV	HA
BRCA_noBRC 5'	CCCTCGCGAATGTACCCCTACGACGTCCCGGACTATGCCTCGGGAGCAGGTGCCTCCTTGTCG	NruI	HA
BRCA_TRUNC 3'	CCCCTCGCGACTACGGCTTTCTTGCTAGCTTGGATG	NruI	
BRCA_TRUNC 5'	CCCCTCGCGAATGTACCCCTACGACGTCCCGGACTATGCCTACGGGACTGAAAATGGCCAAGAG	NruI	HA
BRCA2 KO3'	GACAATGAGAGTTATGACAATGCGAGAGGGATCAAGTTTTGAATGAAACGATGAAGGTATACCCGGCGAAAGGATCATCACTAAGAACATATACTAACACTATTTTATGGCAGCAACGAG		
BRCA2 KO3' Nru_RV	CCTCGCGAGATATCTAGCCAGGGAAGGTGTGTG	NruI	EcoRV
BRCA2 KO3' XbaI	CCTCTAGATATGTCCTTAAGTACTGCC	XbaI	
BRCA2 KO5'	CCGGTCCCCTTTTTCTTCTTTTTCGGCTCCTCCTCCTCCTCCTTTTCTTCCATCCTGAAATTCCCCGTGTGTTGTTTAATTAATCTGCTCGGAAGTCTGGGTCCCATTGTTTGCCTC		
BRCA2 KO5' Bam_Nru	CCTCGCGAGGATCCATAATCAGAATTTGACTTCCG	NruI	BamHI
BRCA2 KO5' XhoI	CCCTCGAGGACATGACATTTCTTGACCC	XhoI	
BRCA2 probe 3'	GTATGAACTCACACTCCGCTGG		
BRCA2 probe 5'	TCGGGAGCAGGTGCCTCC		
BSD 3'	TTAGCCCTCCCACACATAACC		
BSD 5'	ATGGCCAAGCCTTTGTCTC		
CanalTAP3'	CCTCAGGTTGACTTCCCCGCGG		
CanalTAP5'	CCCATGCCTCCACTACACGGTT		
CTAP3'	CACGACAAAACCGAAGCCTTCGAACTGCAGCATGCACTTCTCCGCTTTCGGAGTGAATGTTTTTCTACAAATGACGAGGAAAAAAAGATGGGGCACTACCCTCTACTATTTTCTTTGAT		
CTAP5'	GGAGAACAGCGTATTATTAAGGTGTATGACTCACCTTGTCTCGCTGAAAGTGAGGCCATCTTCGGCATCTATGAGAACGGTGTGGGAGACGTTAGGGACAAGAGAAGATGGAAAAAGAAT		
Cterm probe 5'	TTCAGAGGGTGAGGGGTCGG		

EcoRV101for	CCGATATCTGGGTCCCATTGTTTGCCTC	EcoRV	
EcoRV101rev	CCGATATCGTTAACTATTTTATGGCAGCAACGAG	EcoRV	HpaI
GPI for	TGGGGAGCAAAGATGAAAAC		
GPI rev	CCGTTGCCATAACGGGTCCC		
Hygro 3'	CTATTCCTTTGCCCTCGGAC		
Hygro 5'	ATGAAAAAGCCTGAACTCACC		
Mid Actin	TGTACTCAGCCCTATGCC		
NanalTAP3'	CCACCAACTTGCTGCTGCATCT		
NanalTAP5'	CCGCAGGCCTTGCGCAACACGA		
Neo 3'	TCAAGAAGGCGATAGAAGGC		
Neo 5'	CGCATGATTGAACAAGATGG		
NruIact	CCTCGCGAGTTAACTATTTTATGGCAGCAACGAG	NruI	HpaI
NruItub	CCTCGCGATGGGTCCCATTGTTTGCCTC	NruI	
NTAP3'	CGCATCCACTGCGGCATCATCGAACGCGGTATCATCGATGTCGTGAACTTCATCTTCTATCACCTCTTTAGTGCGTTTCTTATTTTTGGTGCGAGTGTTAAGTGCCCCGGAGGATGAGAT		
NTAP5'	CTTTGAGAGGTGGTGCGCGCTTTAAAGAGAAACCAGCAACAACACCGAAAGGGCTTCTCCAGATTCCCCGTGAAGGTTTCTTAACGCGTATACCAGGGGATGATTGAACAAGATGGATTG		
Outside RAD51	CGTGAAGGTTTCTTAACGC		
Pol I 3'	CATGCGCCTGTGGTTCAGCATAGC		
Pol I 5'	CAGGAGGATCGTTCGGCACCTTGGC		
RPA 3'	CCCCGATATCTTACAAGTAGGCATTAATGC	EcoRV	
RPA 5'	CAGCAGCCATCACAACAACAG		
RPA probe 3'	TACGGTGCCAGGTTGATGTGG		
RPA probe 5'	TACGCTCATTGACGAGTCTGC		
Tb BRC rep for	GCGGTACAAGGAAATTCCAC		
Tb BRC rep rev	AGGAGGCACCTGCTCCCGAA		
Tb BRCA2 for	ATGAGCCACAAAAAAGGAAGACAAGGC		
Tb BRCA2 for2	CCCTCGCGAATGTACCCCTACGACGTCCCGGACTATGCCAGCCACAAAAAAGGAAGACAA	NruI	HA
Tb BRCA2 rev	TTCTCGCATAAGATCAGCG		
Tb BRCA2 rev2	CCCTCGCGACTATTCTCGCATAAGATCAGCGAC	NruI	
Tco BRCA2 for	ATGGTCTTTTCCCAGAAGTCTAAGGG		
Tco BRCA2 for2	CCCCGATATCATGTACCCCTACGACGTCCCGGACTATGCCGTCTTTTCGCAGAAGTCTAAG	EcoRV	HA
Tco BRCA2 rev	CGAAGCAGTTGCTAAAGGTG		
Tco BRCA2 rev2	CCCCGATATCCTTCCGCATAAGATTGACGAT	EcoRV	
Tviv BRCA2 3'	CTACGCCATCGAGCAGGC		
Tviv BRCA2 5'	ATGAAGCAGCGGCAAGTAGG		
Tviv BRCA2 for2	CCCCGATATCATGTACCCCTACGACGTCCCGGACTATGCCAAGCAGCGGCAAGTAGGTGAA	EcoRV	HA
Tviv BRCA2 rev2	CCCCGATATCCTACACTGAACTCTCCTCCTGCAT	EcoRV	
Vivax probe 3'	TACCAGCCGAAGAGGCGGTCG		
Vivax probe 5'	ACTAGTAGGCCACAGCGGGTG		
VSG 121 3'	CGCTGGCTGTGGTGCTCAGAATCATGCAGA		
VSG 121 5'	TAACCTTTACAACAGAGCGCACAAACTTAA		
VSG 221 3'	TGTATCGGCGACAACTGCAG		
VSG 221 5'	ATGCCTTCCAATCAGGAGGC		

Appendix 2: Accession numbers of BRCA2 proteins.

The accession numbers for the BRCA2 proteins used during homology and phylogenetic analysis.

Species	Protein name	Accession number	Size (amino acids)	No. of BRC repeats
Arabidopsis thaliana	BRCA2B	NP_195783	1155	4
Arabidopsis thaliana	BRCA2A	NP_191913	1150	4
Caenorhabditis elegans	potential BRCA2-like protein	AAR98640	394	1
Canis familiaris	BRCA2	BAB91245	3446	8
Dictyostelium discoideum	Hypothetical protein	EAL60741	1623	1
Drosophila melanogaster	Hypothetical protein	NP_611925	947	3
Entamoeba histolytica	Hypothetical protein	147.m00103	719	1
Felis catus	BRCA2	NP_001009858	3372	8
Gallus gallus	BRCA2	BAB83985	3397	8
Giardia lamblia	Hypothetical protein	ORF: 12059 contig_15:113922..127236	1105	1
Homo Sapiens	BRCA2	AAB07223	3418	8
Leishmania major	Hypothetical protein	LmjF20.0060	1165	2
Mus musculus	BRCA2	AAC23702	3328	8
Plasmodium falciparum	Hypothetical protein	PF13_0155	2668	6
Toxoplasma gondii	Brca2 repeat containing protein	49.m03334	2741	8
Trichomonas vaginalis	BRCA2 repeat family protein	XP_001316845	1664	1
Trypanosoma brucei	Hypothetical protein	Tb927.1.640	1648	15
Trypanosoma congolense	Hypothetical protein	congo695a05.p1k_18	1179	3
Trypanosoma cruzi	BRCA2	Tc00.1047053505999.40	1030	2
Trypanosoma vivax	Hypothetical protein	tviv192h02.q1k_9	1068	1
Ustilago maydis	Rad51-associated protein Brh2	AAM92489	1075	1

Appendix 3: Accession numbers of DSS1 proteins.

The accession numbers for the DSS1 proteins used during homology and phylogenetic analysis.

Species	Protein name	Accession number	Size (amino acids)
Arabidopsis thaliana	DSS1	NP_974090	74
Homo Sapiens	DSS1	NP_006295	70
Leishmania major	Hypothetical protein	LmjF29.1290	118
Saccharomyces cerevisiae	SEM1	NP_010651	89
Schizosaccharomyces pombe	Hypothetical protein	NP_594968	71
Trypanosoma brucei	Hypothetical protein	Tb03.28C22.546	137
Trypanosoma congolense	Hypothetical protein	congo1342c06.q1k_0	138
Trypanosoma cruzi	Hypothetical protein	Tc00.1047053509999.70	144
Trypanosoma vivax	Hypothetical protein	tviv1332g04.p1k_2	125
Ustilago maydis	DSS1	AAQ67367	119

Appendix 4: Accession numbers of RAD51 proteins.

The accession numbers for the RAD51 proteins used during homology and phylogenetic analysis.

Species	Protein name	Accession number	Size (amino acids)
Leishmania major	RAD51 protein, putative	LmjF28.0550	377
Trypanosoma brucei	RAD51	AAD51713	373
Trypanosoma cruzi	RAD51	AAZ94621	371
Trypanosoma vivax	RAD51 protein, putative	tviv626b11.p1k_13	410

Appendix 5: The gene sequence of *BRCA2.*

The ORF of *BRCA2* is highlighted in purple, whilst the BRC repeats are indicated in red. The *BRCA2* specific primers and the restriction sites of the enzymes used during single copy analysis are shown.

```
EcoRV         BRCA2 5' XhoI
===     _____\
ATCATGAAGGACATGACATTTCTTGACCCATCGGAATACATACCCCTCCGATAACAAGAA
-----------------------------------------------------------
TAGTACTTCCTGTACTGTAAAGAACTGGGTAGCCTTATGTATGGGGAGGCTATTGTTCTT

AGGTGGGCATCTGTTAGGGCCCTTTTCCCCCTTTCTGCGGTACTGTTTGCTGTCGCGGAA
-----------------------------------------------------------
TCCACCCGTAGACAATCCCGGGAAAAGGGGGAAAGACGCCATGACAAACGACAGCGCCTT

GCGAATATTCACATCTGCGCTTGTTGATACTGAGTATACATAAGTGTATACCGTTGTTGT
-----------------------------------------------------------
CGCTTATAAGTGTAGACGCGAACAACTATGACTCATATGTATTCACATATGGCAACAACA

CCCGGTGCGTGCCCTTTGCACCGACGTTTCTGCTCGCTCATTCTCACCACCTCCACCTCC
-----------------------------------------------------------
GGGCCACGCACGGGAAACGTGGCTGCAAAGACGAGCGAGTAAGAGTGGTGGAGGTGGAGG

                                        _____
ACACGCGTGCAATTGAATGCCTTTTAGTACTTTGTTTGCACTACATTCCGGTCCCCTTTT
-----------------------------------------------------------
TGTGCGCACGTTAACTTACGGAAAATCATGAAACAAACGTGATGTAAGGCCAGGGGAAAA

                                                     ApoI
                                                     ======
_____
TCTTCTTTTTCGGCTCCTCCTCCTCCTCCTCCTTTTCTTCCATCCTGAAATTCCCCGTGT
-----------------------------------------------------------
AGAAGAAAAAGCCGAGGAGGAGGAGGAGGAGGAAAAGAAGGTAGGACTTTAAGGGGCACA

                    ApoI
                    ======
BRCA2 KO5'                              BRC VAR 5'5
_____\       _____
                                       M  S  H  K  K  G  R
1 GTTGTTTAATTAATCTGCTCGGAAGTCAAATTCTGATTATTATGAGCCACAAAAAAGGAA
  -----------------------------------------------------------
  CAACAAATTAATTAGACGAGCCTTCAGTTTAAGACTAATAATACTCGGTGTTTTTTCCTT

                    _____
                    \ BRCA2 KO5' Bam_Nru
                    ApoI
                    ======
_____\
   Q  G  S  N  S  G  A  R  Q  N  S  D  T  P  Q  R  N  R  T  K
20 GACAAGGCAGCAACTCTGGAGCGCGCCAAAATTCCGATACGCCGCAACGGAACCGTACGA
   -----------------------------------------------------------
   CTGTTCCGTCGTTGAGACCTCGCGCGGTTTTAAGGCTATGCGGCGTTGCCTTGGCATGCT

   C  R  S  D  A  P  K  R  Q  R  S  R  S  G  E  S  V  Q  G  K
80 AGTGCCGCTCCGATGCCCCCAAGAGACAACGCAGTCGGTCTGGAGAAAGTGTACAAGGAA
   -----------------------------------------------------------
   TCACGGCGAGGCTACGGGGGTTCTCTGTTGCGTCAGCCAGACCTCTTTCACATGTTCCTT
```

```
                                        ApoI
                                        =====
                        PstI            EcoRI
                        ======          ======                        _____
            S  P  L  Q  E  R  E  T  R  I  Q  P  R  R  D  R  T  Y  G  T
140  AAAGTCCACTGCAGGAGAGGGAAACGAGAATTCAGCCGAGGAGGGACCGCACGTACGGGA
     ------------------------------------------------------------
     TTTCAGGTGACGTCCTCTCCCTTTGCTCTTAAGTCGGCTCCTCCCTGGCGTGCATGCCCT
                                                                _____
                                                                \

                                        ApoI
                                        ======
     BRCA Trunc5'              BRC VAR 3'5
     _____\          _____\
            E  N  G  Q  E  S  T  A  V  Q  G  N  S  T  D  V  P  T  L  F
200  CTGAAAATGGCCAAGAGAGCACTGCGGTACAAGGAAATTCCACAGATGTTCCAACGCTTT
     ------------------------------------------------------------
     GACTTTTACCGGTTCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAA
     _____
        BRC VAR 5'3                          HinfI
        _____                         ======
            V  S  A  A  G  K  P  I  T  V  S  E  S  S  L  Q  V  A  R  A
260  TTGTATCTGCTGCTGGTAAACCCATAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGGG
     ------------------------------------------------------------
     AACATAGACGACGACCATTTGGGTATTGACATAGTCTCAGCAGTGACGTTCATCGTTCCC
                                     _____
                                     \    3' BRCVAR
                                                    ApoI
                                                    ======
                              BRC VAR 3'5
                              _____\
            R  M  N  T  E  N  G  Q  E  S  T  A  V  Q  G  N  S  T  D  V
320  CACGAATGAATACTGAAAATGGCCAGGAGAGCACTGCGGTACAAGGAAATTCCACAGATG
     ------------------------------------------------------------
     GTGCTTACTTATGACTTTTACCGGTCCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTAC

                                             HinfI
                                             ======
            P  T  L  F  V  S  A  A  G  K  P  I  T  V  S  E  S  S  L  Q
380  TTCCAACGCTTTTTGTATCTGCTGCTGGTAAACCCATAACTGTATCAGAGTCGTCACTGC
     ------------------------------------------------------------
     AAGGTTGCGAAAAACATAGACGACGACCATTTGGGTATTGACATAGTCTCAGCAGTGACG
                                          _____
                                          \ 3' BRCVAR         ==
                              BRC VAR 3'5
                              _____\
            V  A  R  A  R  M  N  T  E  N  G  Q  E  S  T  A  V  Q  G  N
440  AAGTAGCAAGGGCACGAATGAATACTGAAAATGGCCAGGAGAGCACTGCGGTACAAGGAA
     ------------------------------------------------------------
     TTCATCGTTCCCGTGCTTACTTATGACTTTTACCGGTCCTCTCGTGACGCCATGTTCCTT
     _____

     ApoI
     ====
     _____\                                                        =
            S  T  D  V  P  T  L  F  V  S  A  A  G  K  P  I  T  V  S  E
500  ATTCCACAGATGTTCCAACGCTTTTTGTATCTGCCGCTGGTAAACCCATAACTGTATCAG
     ------------------------------------------------------------
     TAAGGTGTCTACAAGGTTGCGAAAAACATAGACGGCGACCATTTGGGTATTGACATAGTC
                                                                  _
                                                                  \
```

```
        HinfI
        =====
                                                                          _____
         S   S   L   Q   V   A   R   A   R   M   N   T   E   N   G   Q   E   S   T   A
560  AGTCGTCACTGCAAGTAGCAAGGGCACGAATGAATACTGAAAATGGCCAAGAGAGCACTG
     ------------------------------------------------------------
     TCAGCAGTGACGTTCATCGTTCCCGTGCTTACTTATGACTTTTACCGGTTCTCTCGTGAC
     _____
     \ 3' BRCVAR
                      ApoI
                      ======
          BRC VAR 3'5
     _____\
         V   Q   G   N   S   T   D   V   P   T   L   F   V   S   A   A   G   K   P   I
620  CGGTACAAGGAAATTCCACAGATGTTCCAACGCTTTTTGTATCTGCTGCTGGTAAGCCCA
     ------------------------------------------------------------
     GCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAAAACATAGACGACGACCATTCGGGT

                      HinfI
                      ======
         T   V   S   E   S   S   L   Q   V   A   R   A   R   M   N   T   E   N   G   Q
680  TAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGGGCACGAATGAATACTGAAAATGGCC
     ------------------------------------------------------------
     ATTGACATAGTCTCAGCAGTGACGTTCATCGTTCCCGTGCTTACTTATGACTTTTACCGG
                   _____
                   \ 3' BRCVAR
                                 ApoI
                                 ======
                    BRC VAR 3'5
     _____\
         E   S   T   A   V   Q   G   N   S   T   D   V   P   T   L   F   V   S   A   A
740  AAGAGAGCACTGCGGTACAAGGAAATTCCACAGATGTTCCAACGCTTTTTGTATCTGCCG
     ------------------------------------------------------------
     TTCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAAAACATAGACGGC

                      HinfI
                      ======
         G   K   P   I   T   V   S   E   S   S   L   Q   V   A   R   A   R   M   N   T
800  CTGGTAAACCCATAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGGGCACGAATGAATA
     ------------------------------------------------------------
     GACCATTTGGGTATTGACATAGTCTCAGCAGTGACGTTCATCGTTCCCGTGCTTACTTAT
                      _____
                      \  3' BRCVAR
                                    ApoI
                                    ======
                      BRC VAR 3'5
     _____\
         E   N   G   Q   E   S   T   A   V   Q   G   N   S   T   D   V   P   T   L   F
860  CTGAAAATGGCCAAGAGAGCACTGCGGTACAAGGAAATTCCACAGATGTTCCAACGCTTT
     ------------------------------------------------------------
     GACTTTTACCGGTTCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAA

                            HinfI
                            ======
         V   S   A   A   G   K   P   I   T   V   S   E   S   S   L   Q   V   A   R   A
920  TTGTATCTGCCGCTGGTAAACCCATAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGGG
     ------------------------------------------------------------
     AACATAGACGGCGACCATTTGGGTATTGACATAGTCTCAGCAGTGACGTTCATCGTTCCC
                            _____
                            \ 3' BRCVAR
```

```
                                                        ApoI
                                                       ======

                                      BRC VAR 3'5
                         _____\
            R  M  N  T  E  N  G  Q  E  S  T  A  V  Q  G  N  S  T  D  V
   980  CACGAATGAATACTGAAAATGGCCAGGAGAGCACTGCGGTACAAGGAAATTCCACAGATG
        -----------------------------------------------------------
        GTGCTTACTTATGACTTTTACCGGTCCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTAC


                                                        HinfI
                                                       ======
            P  T  L  F  V  S  A  A  G  K  P  I  T  V  S  E  S  S  L  Q
  1040  TTCCAACGCTTTTTGTATCTGCTGCTGGTAAACCCATAACTGTATCAGAGTCGTCACTGC
        -----------------------------------------------------------
        AAGGTTGCGAAAAACATAGACGACGACCATTTGGGTATTGACATAGTCTCAGCAGTGACG
                                                           _____
                                                           \ 3' BRCVAR


                                                 BRC VAR 3'5   ==
                                       _____
            V  A  R  A  R  M  N  T  E  N  G  Q  E  S  T  A  V  Q  G  N
  1100  AAGTAGCAAGGGCACGAATGAATACTGAAAATGGCCAGGAGAGCACTGCGGTACAAGGAA
        -----------------------------------------------------------
        TTCATCGTTCCCGTGCTTACTTATGACTTTTACCGGTCCTCTCGTGACGCCATGTTCCTT
        _____

        ApoI
        ====

        _____\                                                      =
            S  T  D  V  P  T  L  F  V  S  A  A  G  K  P  I  T  V  S  E
  1160  ATTCCACAGATGTTCCAACGCTTTTTGTATCTGCTGCTGGTAAACCCATAACTGTATCAG
        -----------------------------------------------------------
        TAAGGTGTCTACAAGGTTGCGAAAAACATAGACGACGACCATTTGGGTATTGACATAGTC
                                                                   _
        HinfI                                                      \
        =====


                                                    _____
            S  S  L  Q  V  A  R  A  R  M  N  T  E  N  G  Q  E  S  T  A
  1220  AGTCGTCACTGCAAGTAGCAAGGGCACGAATGAATACTGAAAATGGCCAAGAGAGCACTG
        -----------------------------------------------------------
        TCAGCAGTGACGTTCATCGTTCCCGTGCTTACTTATGACTTTTACCGGTTCTCTCGTGAC
        _____
           3' BRCVAR
                          ApoI
                         ======
                BRC VAR 3'5
        _____\
            V  Q  G  N  S  T  D  V  P  T  L  F  V  S  A  A  G  K  P  I
  1280  CGGTACAAGGAAATTCCACAGATGTTCCAACGCTTTTTGTATCTGCCGCTGGTAAACCCA
        -----------------------------------------------------------
        GCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAAAACATAGACGGCGACCATTTGGGT


                         HinfI
                        ======
            T  V  S  E  S  S  L  Q  V  A  R  A  R  M  N  T  E  N  G  Q
  1340  TAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGGGCACGAATGAATACTGAAAATGGCC
        -----------------------------------------------------------
        ATTGACATAGTCTCAGCAGTGACGTTCATCGTTCCCGTGCTTACTTATGACTTTTACCGG
        _____
           \  3' BRCVAR
```

```
                                        ApoI
                                        ======
                    BRC VAR 3'5
        _____\
              E  S  T  A  V  Q  G  N  S  T  D  V  P  T  L  F  V  S  A  A
1400  AAGAGAGCACTGCGGTACAAGGAAATTCCACAGATGTTCCAACGCTTTTTGTATCTGCCG
      ------------------------------------------------------------
      TTCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAAAACATAGACGGC


                                        HinfI
                                        ======
              G  K  P  I  T  V  S  E  S  S  L  Q  V  A  R  A  R  M  N  T
1460  CTGGTAAACCCATAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGGGCACGAATGAATA
      ------------------------------------------------------------
      GACCATTTGGGTATTGACATAGTCTCAGCAGTGACGTTCATCGTTCCCGTGCTTACTTAT
                              _____
                              \  3' BRCVAR
                                              ApoI
                                              ======
                    BRC VAR 3'5
        _____\
              E  N  G  Q  E  S  T  A  V  Q  G  N  S  T  D  V  P  T  L  F
1520  CTGAAAATGGCCAAGAGAGCACTGCGGTACAAGGAAATTCCACAGATGTTCCAACGCTTT
      ------------------------------------------------------------
      GACTTTTACCGGTTCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAA


                                        HinfI
                                        ======
              V  S  A  A  G  K  P  I  T  V  S  E  S  S  L  Q  V  A  R  A
1580  TTGTATCTGCTGCTGGTAAACCCATAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGGG
      ------------------------------------------------------------
      AACATAGACGACGACCATTTGGGTATTGACATAGTCTCAGCAGTGACGTTCATCGTTCCC
                              _____
                              \  3' BRCVAR
                                              ApoI
                                              ======
                    BRC VAR 3'5
        _____\
              R  M  N  T  E  N  G  Q  E  S  T  A  V  Q  G  N  S  T  D  V
1640  CACGAATGAATACTGAAAATGGCCAGGAGAGCACTGCGGTACAAGGAAATTCCACAGATG
      ------------------------------------------------------------
      GTGCTTACTTATGACTTTTACCGGTCCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTAC


                                              HinfI
                                              ======
              P  T  L  F  V  S  A  A  G  K  P  I  T  V  S  E  S  S  L  Q
1700  TTCCAACGCTTTTTGTATCTGCTGCTGGTAAACCCATAACTGTATCAGAGTCGTCACTGC
      ------------------------------------------------------------
      AAGGTTGCGAAAAACATAGACGACGACCATTTGGGTATTGACATAGTCTCAGCAGTGACG
                                    _____
                                    \  3' BRCVAR
                                                          ==
                    BRC VAR 3'5
        _____\
              V  A  R  A  R  M  N  T  E  N  G  Q  E  S  T  A  V  Q  G  N
1760  AAGTAGCAAGGGCACGAATGAATACTGAAAATGGCCAAGAGAGCACTGCGGTACAAGGAA
      ------------------------------------------------------------
      TTCATCGTTCCCGTGCTTACTTATGACTTTTACCGGTTCTCTCGTGACGCCATGTTCCTT
      _____
      ApoI
      ====
      _____\                                                    ==
              S  T  D  V  P  T  L  F  V  S  A  A  G  K  P  I  T  V  S  E
1820  ATTCCACAGATGTTCCAACGCTTTTTGTATCTGCCGCTGGTAAACCCATAACTGTATCAG
      ------------------------------------------------------------
      TAAGGTGTCTACAAGGTTGCGAAAAACATAGACGGCGACCATTTGGGTATTGACATAGTC
```

```
      HinfI
      ====                                                        _____
            S   S   L   Q   V   A   R   A   R   M   N   T   E   N   G   Q   E   S   T   A
1880  AGTCGTCACTGCAAGTAGCAAGGGCACGAATGAATACTGAAAATGGCCAAGAGAGCACTG
      ------------------------------------------------------------
      TCAGCAGTGACGTTCATCGTTCCCGTGCTTACTTATGACTTTTACCGGTTCTCTCGTGAC
      _____
      \   3' BRCVAR
                      ApoI
                      ======
            BRC VAR 3'5
      _____\
            V   Q   G   N   S   T   D   V   P   T   L   F   V   S   A   A   G   K   P   I
1940  CGGTACAAGGAAATTCCACAGATGTTCCAACGCTTTTTGTATCTGCCGCTGGTAAACCCA
      ------------------------------------------------------------
      GCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAAAACATAGACGGCGACCATTTGGGT

                HinfI
                ======
            T   V   S   E   S   S   L   Q   V   A   R   A   R   M   N   T   E   N   G   Q
2000  TAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGGGCACGAATGAATACTGAAAATGGCC
      ------------------------------------------------------------
      ATTGACATAGTCTCAGCAGTGACGTTCATCGTTCCCGTGCTTACTTATGACTTTTACCGG
              _____
              \   3' BRCVAR
                          ApoI
                          ======
              BRC VAR 3'5
        _____\
            E   S   T   A   V   Q   G   N   S   T   D   V   P   T   L   F   V   S   A   A
2060  AAGAGAGCACTGCGGTACAAGGAAATTCCACAGATGTTCCAACGCTTTTTGTATCTGCCG
      ------------------------------------------------------------
      TTCTCTCGTGACGCCATGTTCCTTTAAGGTGTCTACAAGGTTGCGAAAAACATAGACGGC

                      HinfI
                      ======
            G   K   T   V   T   V   S   E   S   S   L   Q   V   A   S   A   N   A   A   S
2120  CTGGTAAAACCGTAACTGTATCAGAGTCGTCACTGCAAGTAGCAAGTGCGAACGCAGCTT
      ------------------------------------------------------------
      GACCATTTTGGCATTGACATAGTCTCAGCAGTGACGTTCATCGTTCACGCTTGCGTCGAA
              _____
              \   3' BRCVAR
              BRCA2 probe 5'
            _____\
            S   A   K   P   I   S   G   A   G   A   S   L   S   K   R   T   P   R   T   H
2180  CATCTGCTAAACCCATTTCGGGAGCAGGTGCCTCCTTGTCGAAGAGGACGCCACGTACAC
      ------------------------------------------------------------
      GTAGACGATTTGGGTAAAGCCCTCGTCCACGGAGGAACAGCTTCTCCTGCGGTGCATGTG

            R   K   S   A   S   S   S   P   L   S   S   S   K   L   A   R   K   P   F   V
2240  ACCGTAAATCAGCATCATCATCGCCATTGTCATCATCCAAGCTAGCAAGAAAGCCGTTTG
      ------------------------------------------------------------
      TGGCATTTAGTCGTAGTAGTAGCGGTAACAGTAGTAGGTTCGATCGTTCTTTCGGCAAAC
                                            _____
                                            \   BRCA2_Trunc3'
            V   P   F   A   K   N   K   G   A   V   A   K   G   V   G   E   A   V   P   S
2300  TGGTTCCTTTTTGCTAAGAATAAAGGAGCGGTTGCGAAAGGAGTAGGGGAAGCGGTGCCAT
      ------------------------------------------------------------
      ACCAAGGAAAACGATTCTTATTTCCTCGCCAACGCTTTCCTCATCCCCTTCGCCACGGTA
              C term probe
            _____\
            A   S   H   M   P   S   S   E   G   E   G   K   E   V   G   R   T   P   R   H
2360  CGGCGTCCCACATGCCGAGTTCAGAGGGTGAGGGGTCGGAAGTAGGTCGAACCCCCCGAC
      ------------------------------------------------------------
      GCCGCAGGGTGTACGGCTCAAGTCTCCCACTCCCCAGCCTTCATCCAGCTTGGGGGGCTG
```

```
              L   S   F   D   I   F   T   F   R   S   L   S   M   T   V   P   P   S   I   D
2420 ATCTTTCGTTTGACATTTTCACGTTTCGCTCATTATCGATGACGGTACCTCCCTCAATTG
     ------------------------------------------------------------
     TAGAAAGCAAACTGTAAAAGTGCAAAGCGAGTAATAGCTACTGCCATGGAGGGAGTTAAC

              E   I   V   R   G   N   F   L   F   K   Q   F   G   C   S   P   E   L   L   K
2480 ACGAGATTGTGCGAGGAAACTTTTTGTTTAAGCAATTTGGGTGTTCACCTGAACTTCTGA
     ------------------------------------------------------------
     TGCTCTAACACGCTCCTTTGAAAAACAAATTCGTTAAACCCACAAGTGGACTTGAAGACT

                                                          ApoI
                                                          ======
              L   L   E   I   P   A   E   C   E   F   I   P   S   A   N   F   R   K   A   M
2540 AGTTACTGGAAATACCAGCGGAGTGTGAGTTCATACCATCGGCAAATTTTCGTAAGGCCA
     ------------------------------------------------------------
     TCAATGACCTTTATGGTCGCCTCACACTCAAGTATGGTAGCCGTTTAAAAGCATTCCGGT
                     _____
                     \  BRCA2 probe 3'

              L   T   L   G   A   S   P   R   G   C   P   D   A   W   C   L   Q   M   L   T
2600 TGCTTACCTTGGGGGCTTCTCCACGGGGCTGCCCAGATGCGTGGTGTCTTCAAATGTTGA
     ------------------------------------------------------------
     ACGAATGGAACCCCCGAAGAGGTGCCCCGACGGGTCTACGCACCACAGAAGTTTACAACT

                    HinfI
                    ======
              S   T   L   L   K   L   R   G   L   T   L   H   I   D   P   P   L   P   V   F
2660 CGTCCACGCTTCTGAAGTTGCGGGGACTCACATTACACATTGATCCACCCCTTCCCGTGT
     ------------------------------------------------------------
     GCAGGTGCGAAGACTTCAACGCCCCTGAGTGTAATGTGTAACTAGGTGGGGAAGGGCACA

              S   V   A   H   T   L   L   H   M   C   F   K   Y   N   H   E   Y   V   E   G
2720 TTTCTGTCGCACATACTTTGCTTCACATGTGTTTTAAATATAATCACGAGTATGTTGAGG
     ------------------------------------------------------------
     AAAGACAGCGTGTATGAAACGAAGTGTACACAAAATTTATATTAGTGCTCATACAACTCC

              K   R   P   A   L   R   L   I   A   E   G   D   V   Q   A   A   S   L   V   V
2780 GCAAACGGCCTGCTTTGCGTTTGATTGCGGAAGGGGACGTTCAAGCAGCCTCACTGGTGG
     ------------------------------------------------------------
     CGTTTGCCGGACGAAACGCAAACTAACGCCTTCCCCTGCAAGTTCGTCGGAGTGACCACC

              V   W   V   V   S   V   S   F   E   E   R   L   T   P   H   T   C   T   A   V
2840 TAGTCTGGGTAGTGTCGGTATCTTTTGAGGAGCGCCTTACTCCTCACACCTGCACGGCAG
     ------------------------------------------------------------
     ATCAGACCCATCACAGCCATAGAAAACTCCTCGCGGAATGAGGAGTGTGGACGTGCCGTC

              V   S   D   G   F   Y   H   V   K   V   S   L   D   I   P   L   T   N   L   V
2900 TGGTTTCCGATGGGTTTTACCACGTTAAAGTGTCTCTTGATATTCCATTAACGAACTTAG
     ------------------------------------------------------------
     ACCAAAGGCTACCCAAAATGGTGCAATTTCACAGAGAACTATAAGGTAATTGCTTGAATC

              R   N   G   T   L   R   C   G   Q   K   I   V   T   C   G   A   R   M   L   R
2960 TTCGTAATGGAACCCTGCGGTGTGGTCAGAAGATTGTTACTTGCGGTGCGAGGATGCTGA
     ------------------------------------------------------------
     AAGCATTACCTTGGGACGCCACACCAGTCTTCTAACAATGAACGCCACGCTCCTACGACT

              R   D   C   C   S   P   L   E   C   K   D   E   V   L   L   S   I   N   Y   N
3020 GGAGAGACTGTTGTTCTCCACTAGAATGCAAAGATGAAGTGCTCCTCTCCATTAACTACA
     ------------------------------------------------------------
     CCTCTCTGACAACAAGAGGTGATCTTACGTTTCTACTTCACGAGGAGAGGTAATTGATGT
```

```
           C   T   Q   P   V   G   P   S   S   P   L   G   L   Y   H   T   C   L   P   T
3080 ACTGCACACAACCTGTGGGACCTTCCTCACCTCTAGGTCTCTATCATACTTGTTTGCCGA
     ------------------------------------------------------------
     TGACGTGTGTTGGACACCCTGGAAGGAGTGGAGATCCAGAGATAGTATGAACAAACGGCT

           L   L   P   S   A   M   D   M   L   G   G   L   V   P   C   L   K   G   R   V
3140 CACTGCTGCCTTCCGCTATGGATATGCTTGGTGGTTTGGTACCTTGTTTGAAAGGGCGAG
     ------------------------------------------------------------
     GTGACGACGGAAGGCGATACCTATACGAACCACCAAACCATGGAACAAACTTTCCCGCTC

           E   R   V   L   P   P   F   F   L   E   K   T   F   K   G   A   R   T   G   D
3200 TGGAGCGCGTGCTTCCGCCCTTTTTCCTTGAGAAAACATTTAAGGGTGCGCGAACTGGCG
     ------------------------------------------------------------
     ACCTCGCGCACGAAGGCGGGAAAAAGGAACTCTTTTGTAAATTCCCACGCGCTTGACCGC

           T   R   G   S   T   G   G   A   L   K   I   V   R   S   L   L   A   Q   L   S
3260 ACACAAGAGGCAGCACAGGTGGTGCGTTGAAAATTGTGAGGAGTTTATTGGCGCAGCTAA
     ------------------------------------------------------------
     TGTGTTCTCCGTCGTGTCCACCACGCAACTTTTAACACTCCTCAAATAACCGCGTCGATT

           F   Q   E   C   M   A   R   G   A   V   A   P   F   E   G   K   S   D   R   Q
3320 GTTTCCAGGAATGTATGGCACGCGGAGCTGTTGCTCCGTTTGAAGGAAAAAGTGACCGAC
     ------------------------------------------------------------
     CAAAGGTCCTTACATACCGTGCGCCTCGACAACGAGGCAAACTTCCTTTTTCACTGGCTG

           L   S   R   L   T   S   F   L   L   S   C   E   R   Q   G   D   V   L   L   Q
3380 AACTTTCACGTCTGACATCGTTTTTGTTGTCTTGTGAGCGACAGGGGGACGTCCTCTTGC
     ------------------------------------------------------------
     TTGAAAGTGCAGACTGTAGCAAAAACAACAGAACACTCGCTGTCCCCCTGCAGGAGAACG

           I   W   D   D   C   G   A   N   C   P   A   G   D   L   E   E   H   S   C   D
3440 AAATATGGGATGATTGCGGTGCCAACTGTCCGGCGGGGGATTTGGAGGAACATTCGTGTG
     ------------------------------------------------------------
     TTTATACCCTACTAACGCCACGGTTGACAGGCCGCCCCCTAAACCTCCTTGTAAGCACAC

           F   P   P   E   G   A   E   I   V   V   F   S   V   T   P   S   R   F   R   P
3500 ATTTTCCACCGGAGGGAGCTGAGATTGTCGTTTTCTCCGTAACCCCTTCACGCTTCCGAC
     ------------------------------------------------------------
     TAAAAGGTGGCCTCCCTCGACTCTAACAGCAAAAGAGGCATTGGGGAAGTGCGAAGGCTG

           G   H   P   F   Q   R   T   T   V   L   Y   S   R   S   P   L   R   Y   S   I
3560 CTGGTCACCCCTTCCAGCGGACGACAGTTTTGTACTCTCGGAGCCCTCTTCGGTATAGCA
     ------------------------------------------------------------
     GACCAGTGGGGAAGGTCGCCTGCTGTCAAAACATGAGAGCCTCGGGAGAAGCCATATCGT

           V   S   P   P   R   K   G   F   V   R   Q   P   L   R   S   A   E   D   V   S
3620 TAGTCTCACCGCCGCGTAAGGGGTTTGTGAGGCAACCTTTGCGCTCAGCTGAAGATGTGT
     ------------------------------------------------------------
     ATCAGAGTGGCGGCGCATTCCCCAAACACTCCGTTGGAAACGCGAGTCGACTTCTACACA

           P   K   T   E   T   G   D   A   I   D   F   A   G   L   F   V   G   T   K   S
3680 CCCCAAAAACAGAGACAGGTGATGCCATCGATTTTGCTGGCTTGTTCGTCGGCACCAAGA
     ------------------------------------------------------------
     GGGGTTTTTGTCTCTGTCCACTACGGTAGCTAAAACGACCGAACAAGCAGCCGTGGTTCT

           V   D   T   V   N   S   H   I   I   V   A   L   N   D   G   W   K   P   G   C
3740 GTGTGGACACGGTCAACTCACATATTATCGTGGCCTTAAATGACGGATGGAAACCTGGAT
     ------------------------------------------------------------
     CACACCTGTGCCAGTTGAGTGTATAATAGCACCGGAATTTACTGCCTACCTTTGGACCTA

           V   P   A   S   Y   F   M   I   D   V   P   H   A   T   G   K   K   E   I   V
3800 GTGTTCCGGCTTCCTACTTTATGATTGATGTCCCACATGCCACGGGCTCAAAAGAGATTG
     ------------------------------------------------------------
     CACAAGGCCGAAGGATGAAATACTAACTACAGGGTGTACGGTGCCCGAGTTTTCTCTAAC
```

```
            L  A  L  P  S  I  P  F  T  P  V  I  V  Q  N  A  S  F  I  R
3860  TGCTTGCTTTGCCATCAATACCATTCACACCTGTTATTGTGCAGAATGCTTCTTTTATAC
      ------------------------------------------------------------
      ACGAACGAAACGGTAGTTATGGTAAGTGTGGACAATAACACGTCTTACGAAGAAAATATG

            C  A  E  D  L  G  P  D  C  I  H  V  L  A  N  E  Y  T  K  V
3920  GTTGCGCGGAGGACTTGGGACCCGATTGCATACACGTACTGGCGAATGAGTACACGAAGG
      ------------------------------------------------------------
      CAACGCGCCTCCTGAACCCTGGGCTAACGTATGTGCATGACCGCTTACTCATGTGCTTCC

                                              HinfI            ApoI
                                              ======           ======
            Y  S  R  P  A  E  P  L  L  R  G  V  V  E  S  L  G  K  I  R
3980  TTTACAGCCGTCCTGCTGAGCCGCTCCTTCGCGGTGTTGTTGAGTCACTTGGGAAAATTC
      ------------------------------------------------------------
      AAATGTCGGCAGGACGACTCGGCGAGGAAGCGCCACAACAACTCAGTGAACCCTTTTAAG

            G  M  A  K  S  S  R  P  I  I  A  R  S  E  E  L  L  R  M  R
4040  GTGGGATGGCGAAGTCAAGTAGGCCTATTATTGCCCGTTCTGAGGAGCTGCTACGAATGC
      ------------------------------------------------------------
      CACCCTACCGCTTCAGTTCATCCGGATAATAACGGGCAAGACTCCTCGACGATGCTTACG

            T  L  S  E  E  A  R  A  D  I  C  R  L  S  R  E  L  V  G  G
4100  GGACGTTGAGTGAGGAAGCTCGGGCCGATATATGTCGACTGTCGAGAGAACTGGTTGGAG
      ------------------------------------------------------------
      CCTGCAACTCACTCCTTCGAGCCCGGCTATATACAGCTGACAGCTCTCTTGACCAACCTC

            PstI
            ======
            D  E  L  P  N  P  A  A  T  A  Q  P  S  P  R  Y  Q  L  R  Q
4160  GGGATGAGTTACCTAATCCTGCAGCCACAGCGCAGCCGTCTCCACGGTATCAATTACGCC
      ------------------------------------------------------------
      CCCTACTCAATGGATTAGGACGTCGGTGTCGCGTCGGCAGAGGTGCCATAGTTAATGCGG

            E  A  S  T  P  V  E  Q  S  I  T  V  S  E  T  S  A  A  R  T
4220  AGGAAGCATCTACCCCTGTTGAGCAGAGTATTACGGTTTCAGAAACAAGTGCAGCACGTA
      ------------------------------------------------------------
      TCCTTCGTAGATGGGGACAACTCGTCTCATAATGCCAAAGTCTTTGTTCACGTCGTGCAT

            L  S  S  E  E  Q  V  E  D  L  R  S  S  N  V  K  A  S  P
4280  CCTTAAGCAGTGAAGAAGAACAGGTAGAGGATTTAAGGTCCTCGAATGTCAAGGCTAGCC
      ------------------------------------------------------------
      GGAATTCGTCACTTCTTCTTGTCCATCTCCTAAATTCCAGGAGCTTACAGTTCCGATCGG

            R  R  H  V  F  G  N  I  V  G  F  R  L  L  K  C  Q  G  K  D
4340  CACGGCGTCATGTATTTGGCAACATTGTGGGATTTCGGCTTCTCAAGTGCCAGGGCTCTG
      ------------------------------------------------------------
      GTGCCGCAGTACATAAACCGTTGTAACACCCTAAAGCCGAAGAGTTCACGGTCCCGAGAC

            K  P  E  C  I  E  I  L  G  G  R  P  S  T  L  V  S  G  K  G
4400  ACAAACCCGAATGCATTGAGATTTTGGGTGGCCGTCCCAGCACTCTCGTCTCTGGAAGTG
      ------------------------------------------------------------
      TGTTTGGGCTTACGTAACTCTAAAACCCACCGGCAGGGTCGTGAGAGCAGAGACCTTCAC

            K  F  V  V  S  P  S  D  F  S  Q  S  L  V  Y  F  E  A  D  I
4460  GAAAGTTTGTAGTGTCCCCCTCGGACTTTTCTCAGAGCCTTGTGTACTTTGAGGCTGATA
      ------------------------------------------------------------
      CTTTCAAACATCACAGGGGGAGCCTGAAAAGAGTCTCGGAACACATGAAACTCCGACTAT

            Q  F  G  A  T  A  K  Q  C  A  Q  T  K  V  R  S  P  S  V  L
4520  TTCAATTTGGTGCCACCGCAAAACAATGCGCTCAAACCAAGGTAAGATCTCCTTCCGTGC
      ------------------------------------------------------------
      AAGTTAAACCACGGTGGCGTTTTGTTACGCGAGTTTGGTTCCATTCTAGAGGAAGGCACG
```

```
          H  S  L  L  E  Q  C  I  P  L  K  R  A  C  A  L  T  V  D  E
4580 TTCATTCATTATTGGAGCAGTGTATTCCATTGAAACGAGCATGCGCCTTAACAGTGGACG
     ------------------------------------------------------------
     AAGTAAGTAATAACCTCGTCACATAAGGTAACTTTGCTCGTACGCGGAATTGTCACCTGC

          I  F  A  D  Y  Y  L  A  R  I  K  Q  L  E  D  W  Q  T  P  H
4640 AGATTTTTGCGGATTATTACTTGGCTCGCATTAAACAATTAGAAGATTGGCAGACACCTC
     ------------------------------------------------------------
     TCTAAAAACGCCTAATAATGAACCGAGCGTAATTTGTTAATCTTCTAACCGTCTGTGGAG

          E  E  C  W  W  R  L  L  T  Q  S  H  V  V  E  I  T  S  D  V
4700 ACGAAGAGTGCTGGTGGCGGCTTCTCACTCAATCGCACGTGGTGGAAATTACCTCCGATG
     ------------------------------------------------------------
     TGCTTCTCACGACCACCGCCGAAGAGTGAGTTAGCGTGCACCACCTTTAATGGAGGCTAC

          S  G  T  P  P  E  E  L  V  G  L  Q  W  L  S  N  E  W  K  M
4760 TCTCAGGGACCCCACCGGAGGAGTTGGTCGGTTTGCAATGGCTCTCTAATGAATGGAAGA
     ------------------------------------------------------------
     AGAGTCCCTGGGGTGGCCTCCTCAACCAGCCAAACGTTACCGAGAGATTACTTACCTTCT

          L  L  N  I  L  S  G  K  L  K  H  C  L  F  M  F  S  V  E  G
4820 TGTTGCTCAATATTCTTTCTGGCTCCTTGAAGCATTGCCTGTTCATGTTCAGTGTAGAGG
     ------------------------------------------------------------
     ACAACGAGTTATAAGAAAGACCGAGGAACTTCGTAACGGACAAGTACAAGTCACATCTCC

          S  E  M  V  R  A  T  F  I  K  E  Q  C  S  V  A  D  L  M  R
4880 GAAGTGAAATGGTTCGTGCTACTTTCATCAAGGAACAGTGTAGCGTCGCTGATCTTATGC
     ------------------------------------------------------------
     CTTCACTTTACCAAGCACGATGAAAGTAGTTCCTTGTCACATCGCAGCGACTAGAATACG
            BRCA2 KO3'Nru_RV                                    _____
            _____\                          \ BRC VAR 3'3
          E
4940 GAGAATAGCCAGGGAAGGTGTGTGTTAGTATATGTTCTTAGTGATGATCCTTTCGCCGGG
     ------------------------------------------------------------
     CTCTTATCGGTCCCTTCCACACACAATCATATACAAGAATCACTACTAGGAAAGCGGCCC
        _____                    _____
                                    \
     TATACCTTCATCGTTTCATTCAAAACTTGATCCCTCTCGCATTGTCATAACTCTCATTGT
     ------------------------------------------------------------
     ATATGGAAGTAGCAAAGTAAGTTTTGAACTAGGGAGAGCGTAACAGTATTGAGAGTAACA
     _____
        BRCA2 KO3'

                                                Hinf I
                                                ======
     CAATATTTTGGTTTCGTACTCTGCGTTCGGACAAGGTTCTCTCAAGAGTCTGGTTCTTTT
     ------------------------------------------------------------
     GTTATAAAACCAAAGCATGAGACGCAAGCCTGTTCCAAGAGAGTTCTCAGACCAAGAAAA
     _

     TTTTTCCGCTCCTCTGGTTAAGTAGTTGTGGGTCTGCGCAGCATAGAACGAAGGCCAAGA
     ------------------------------------------------------------
     AAAAAGGCGAGGAGACCAATTCATCAACACCCAGACGCGTCGTATCTTGCTTCCGGTTCT

        Hind III                                     Apo I
        ======                                       ======
     CAGAAGCTTAACCTGAAGTTTGTGTTTTTATTTATTTGTTTTGAGAATTTAAACCTTGCC
     ------------------------------------------------------------
     GTCTTCGAATTGGACTTCAAACACAAAAATAAATAAACAAAACTCTTAAATTTGGAACGG

     ACTTCCCTTTTTTTTTTTTTTGCACTTTCAATTTCAGCCAAGAAAAACGAAAGAACTGAG
     ------------------------------------------------------------
     TGAAGGGAAAAAAAAAAAAAACGTGAAAGTTAAAGTCGGTTCTTTTTGCTTTCTTGACTC
```

```
GGAAGGACGTGGTAGTAGTGTTGCACATTCGTCAGTACGGCAGTACTTAAGGACATACTC
------------------------------------------------------------
CCTTCCTGCACCATCATCACAACGTGTAAGCAGTCATGCCGTCATGAATTCCTGTATGAG
                                      _____
                                      \   BRCA2 KO3' Xba
```

Appendix 6: The gene sequence of *RAD51*.

The ORF of *RAD51* is highlighted in purple and the *RAD51* specific primers are shown.

```
      TGAGCTGCTCTGCTGCTGTCTCATTGTTTCCTTTGTACCTTCACGTTCAGCATTAGAAGT
      ------------------------------------------------------------
      ACTCGACGAGACGACGACAGAGTAACAAAGGAAACATGGAAGTGCAAGTCGTAATCTTCA
                                    _____
      TCGGTGAGTCGTGGGGAATTCTTTGAGAGGTGGTGCGCGCTTTAAAGAGAAACCAGCAAC
      ------------------------------------------------------------
      AGCCACTCAGCACCCCTTAAGAAACTCTCCACCACGCGCGAAATTTCTCTTTGGTCGTTG
                       N TAP 5'
      _____
                                                                  M
   1  AACACCGAAAGGGCTTCTCCAGATTCCCCGTGAAGGTTTCTTAACGCGTATACCAGGGGA
      ------------------------------------------------------------
      TTGTGGCTTTCCCGAAGAGGTCTAAGGGGCACTTCCAAAGAATTGCGCATATGGTCCCCT
      ___\
        N  T  R  T  K  N  K  K  R  T  K  E  V  I  E  D  E  V  H  D
   2  TGAACACTCGCACCAAAAATAAGAAACGCACTAAAGAGGTGATAGAAGATGAAGTTCACG
      ------------------------------------------------------------
      ACTTGTGAGCGTGGTTTTTATTCTTTGCGTGATTTCTCCACTATCTTCTACTTCAAGTGC
      _____
        \                           N TAP 3'
        I  D  D  T  A  F  D  D  A  A  V  D  A  V  N  D  N  T  Q  E
  62  ACATCGATGATACCGCGTTCGATGATGCCGCAGTGGATGCGGTCAACGACAACACGCAAG
      ------------------------------------------------------------
      TGTAGCTACTATGGCGCAAGCTACTACGGCGTCACCTACGCCAGTTGCTGTTGTGCGTTC
      _____
        M  Q  Q  Q  V  G  D  A  A  G  G  P  S  F  R  V  L  Q  I  M
 122  AGATGCAGCAGCAAGTTGGTGACGCTGCCGGTGGGCCTTCCTTTCGTGTCCTTCAGATAA
      ------------------------------------------------------------
      TCTACGTCGTCGTTCAACCACTGCGACGGCCACCCGGAAGGAAAGCACAGGAAGTCTATT
        E  N  Y  G  V  A  S  A  D  I  K  K  L  M  E  C  G  F  L  T
 182  TGGAAAACTATGGAGTTGCTTCTGCTGATATCAAAAAGTTGATGGAGTGTGGCTTTCTCA
      ------------------------------------------------------------
      ACCTTTTGATACCTCAACGAAGACGACTATAGTTTTTCAACTACCTCACACCGAAAGAGT
        V  E  S  V  A  Y  A  P  K  K  S  I  L  A  V  K  G  I  S  E
 242  CCGTTGAGTCTGTCGCGTATGCACCGAAGAAATCAATTTTAGCAGTGAAGGGCATAAGTG
      ------------------------------------------------------------
      GGCAACTCAGACAGCGCATACGTGGCTTCTTTAGTTAAAATCGTCACTTCCCGTATTCAC
        A  K  A  E  K  I  M  A  E  C  C  R  L  T  P  M  G  F  T  R
 302  AGGCAAAGGCTGAGAAGATAATGGCGGAGTGTTGTAGACTCACTCCGATGGGCTTCACGC
      ------------------------------------------------------------
      TCCGTTTCCGACTCTTCTATTACCGCCTCACAACATCTGAGTGAGGCTACCCGAAGTGCG
        A  T  V  F  Q  E  Q  R  K  E  T  I  M  V  T  T  G  S  R  E
 362  GCGCTACGGTTTTCCAAGAGCAACGGAAAGAAACTATTATGGTCACGACAGGCAGCCGTG
      ------------------------------------------------------------
      CGCGATGCCAAAAGGTTCTCGTTGCCTTTCTTTGATAATACCAGTGCTGTCCGTCGGCAC
        V  D  K  L  L  G  G  G  I  E  V  G  S  I  T  E  L  F  G  E
 422  AGGTGGACAAACTCCTTGGAGGTGGCATTGAAGTTGGTAGCATCACGGAACTTTTCGGTG
      ------------------------------------------------------------
      TCCACCTGTTTGAGGAACCTCCACCGTAACTTCAACCATCGTAGTGCCTTGAAAAGCCAC
```

```
           F  R  T  G  K  T  Q  L  C  H  T  L  C  V  T  C  Q  L  P  L
482  AGTTTCGCACAGGGAAGACGCAGCTCTGCCATACCCTTTGTGTTACATGCCAACTTCCAC
     ------------------------------------------------------------
     TCAAAGCGTGTCCCTTCTGCGTCGAGACGGTATGGGAAACACAATGTACGGTTGAAGGTG

           S  Q  G  G  G  E  G  M  A  L  Y  I  D  T  E  G  T  F  R  P
542  TTTCGCAAGGTGGTGGTGAGGGAATGGCGCTTTATATTGACACTGAGGGAACATTCCGTC
     ------------------------------------------------------------
     AAAGCGTTCCACCACCACTCCCTTACCGCGAAATATAACTGTGACTCCCTTGTAAGGCAG

           E  R  L  V  A  V  A  E  R  Y  S  L  D  P  E  A  V  L  E  N
602  CTGAGCGCTTGGTAGCTGTGGCGGAACGGTACAGCCTGGACCCAGAGGCCGTTCTTGAAA
     ------------------------------------------------------------
     GACTCGCGAACCATCGACACCGCCTTGCCATGTCGGACCTGGGTCTCCGGCAAGAACTTT

           V  A  C  A  R  A  Y  N  T  D  H  Q  Q  Q  L  L  L  Q  A  S
662  ATGTGGCTTGCGCTCGTGCTTACAACACGGACCATCAGCAGCAGTTGTTATTGCAAGCGT
     ------------------------------------------------------------
     TACACCGAACGCGAGCACGAATGTTGTGCCTGGTAGTCGTCGTCAACAATAACGTTCGCA

           A  T  M  A  E  H  R  V  A  I  I  V  V  D  S  A  T  A  L  Y
722  CCGCCACTATGGCTGAGCACCGTGTCGCAATAATCGTTGTTGATTCCGCCACAGCTCTTT
     ------------------------------------------------------------
     GGCGGTGATACCGACTCGTGGCACAGCGTTATTAGCAACAACTAAGGCGGTGTCGAGAAA

           R  T  D  Y  N  G  R  G  E  L  A  A  R  Q  M  H  L  G  K  F
782  ACCGTACTGATTACAATGGACGGGGTGAGTTGGCAGCACGGCAGATGCATCTTGGAAAGT
     ------------------------------------------------------------
     TGGCATGACTAATGTTACCTGCCCCACTCAACCGTCGTGCCGTCTACGTAGAACCTTTCA

           L  R  S  L  R  N  L  A  N  E  Y  N  V  A  V  V  V  T  N  Q
842  TCCTCCGCTCTTTGCGCAATCTTGCTAATGAGTACAACGTGGCCGTCGTTGTTACCAATC
     ------------------------------------------------------------
     AGGAGGCGAGAAACGCGTTAGAACGATTACTCATGTTGCACCGGCAGCAACAATGGTTAG

           V  V  A  N  V  D  G  A  A  P  T  F  Q  A  D  S  K  K  P  I
902  AGGTTGTTGCCAATGTGGATGGTGCTGCCCCCACATTCCAAGCGGATTCTAAGAAACCCA
     ------------------------------------------------------------
     TCCAACAACGGTTACACCTACCACGACGGGGGTGTAAGGTTCGCCTAAGATTCTTTGGGT
                                                                  _
           G  G  H  I  M  A  H  A  S  T  T  R  L  S  L  R  K  G  R  G
962  TTGGGGGCCACATCATGGCACATGCCTCCACTACACGGTTGAGCTTACGGAAGGGAAGGG
     ------------------------------------------------------------
     AACCCCCGGTGTAGTACCGTGTACGGAGGTGATGTGCCAACTCGAATGCCTTCCCTTCCC
     _____
           E  Q  R  I  I  K  V  Y  D  S  P  C  L  A  E  S  E  A  I  F
1022 GAGAACAGCGTATTATTAAGGTGTATGACTCACCTTGTCTCGCTGAAAGTGAGGCCATCT
     ------------------------------------------------------------
     CTCTTGTCGCATAATAATTCCACATACTGAGTGGAACAGAGCGACTTTCACTCCGGTAGA
              C TAP 5'
     _____\
           G  I  Y  E  N  G  V  G  D  V  R  D
1082 TCGGCATCTATGAGAACGGTGTGGGAGACGTTAGGGACTAGTGCCCCATCTTTTTTTTCC
     ------------------------------------------------------------
     AGCCGTAGATACTCTTGCCACACCCTCTGCAATCCCTGATCACGGGGTAGAAAAAAAAGG
                                                         _____
                                                         \
     TCGTCATTTGTAGAAAAACATTCACTCCGAAAGCGGAGAAGTGCATGCTGCAGTTCGAAG
     ------------------------------------------------------------
     AGCAGTAAACATCTTTTTGTAAGTGAGGCTTTCGCCTCTTCACGTACGACGTCAAGCTTC
     _____
              C TAP 3'
```

```
GCTTCGGTTTTGTCGTGTTTATGGTACACGTTATTCTGTCCTCCTTCCGGGCAGCAGGTC
------------------------------------------------------------
CGAAGCCAAAACAGCACAAATACCATGTGCAATAAGACAGGAGGAAGGCCCGTCGTCCAG
```

```
ACTTTTCGATTTGCCTACGACTTCACTCTTTTTTTTTTCTCTTTCCTCCTGTTTAGTGCA
------------------------------------------------------------
TGAAAAGCTAAACGGATGCTGAAGTGAGAAAAAAAAAGAGAAAGGAGGACAAATCACGT
```

Appendix 7: Pair-wise comparison of the BRC repeats.

The BRC repeat sequences from the *T. brucei* BRCA2 polypeptide (Tb1 and Tb15) were compared against the BRC repeat sequences from putative Brca2 homologues from *T. congolense* (Tcon), *T. cruzi* (Tc1 and Tc2), *T. vivax* (Tv) and *H. sapiens* (H1-8). Pair-wise alignments were performed using AlignX (Vector NTI) and the percentage identities and similarities calculated. The percentage identities are displayed in bold.

	Tb	Tb15	Tcon	Tv	Tcr1	Tcr2	Lm1	Lm2	H1	H2	H3	H4	H5	H6	H7	H8
Tb		**65.7** 71.4	**48.6** 65.7	**31.4** 51.4	**48.6** 65.7	**48.6** 62.9	**42.9** 60.0	**40.0** 57.1	**20.0** 37.1	**25.7** 31.4	**22.9** 37.1	**17.1** 34.3	**17.1** 37.1	**20.0** 34.3	**31.4** 48.6	**22.9** 48.6
Tb15			**48.6** 57.1	**28.6** 40.0	**42.9** 51.4	**40.0** 48.6	**45.7** 57.1	**45.7** 54.3	**14.3** 31.4	**25.7** 34.3	**20.0** 40.0	**20.0** 31.4	**17.1** 31.4	**25.7** 37.1	**31.4** 51.4	**25.7** 42.9
Tcon				**42.9** 54.3	**54.3** 62.9	**57.1** 71.4	**45.7** 54.3	**37.1** 48.6	**20.0** 42.9	**16.7** 38.9	**28.6** 37.1	**22.9** 37.1	**28.6** 42.9	**20.0** 34.3	**37.1** 54.3	**37.1** 48.6
Tv					**34.3** 54.3	**31.4** 51.4	**34.3** 57.1	**31.4** 51.4	**25.7** 40.0	**45.7** 54.3	**25.7** 40.0	**28.6** 54.3	**34.3** 42.9	**22.9** 40.0	**34.3** 48.6	**34.3** 51.4
Tcr1						**82.9** 85.7	**54.3** 54.3	**37.1** 51.4	**20.0** 40.0	**22.9** 34.3	**25.7** 37.1	**25.7** 37.1	**25.7** 31.4	**20.0** 37.1	**40.0** 48.6	**34.3** 51.4
Tcr2							**48.6** 51.4	**34.3** 48.6	**17.1** 37.1	**22.9** 34.3	**25.7** 34.3	**28.6** 40.0	**22.9** 31.4	**20.0** 31.4	**42.9** 51.4	**28.6** 45.7
Lm1								**45.7** 57.1	**20.0** 31.4	**25.7** 40.0	**31.4** 45.7	**34.3** 42.9	**17.1** 28.6	**34.3** 48.6	**34.3** 42.9	**37.1** 48.6
Lm2									**14.3** 31.4	**17.1** 31.4	**28.6** 40.0	**31.4** 37.1	**25.7** 37.1	**22.9** 34.3	**25.7** 40.0	**25.7** 45.7
H1										**22.9** 37.1	**28.6** 37.1	**28.6** 42.9	**11.4** 40.0	**34.3** 42.9	**28.6** 51.4	**25.7** 40.0
H2											**22.9** 37.1	**31.4** 57.1	**22.9** 37.1	**31.4** 45.7	**34.3** 54.3	**22.9** 45.7
H3												**42.9** 54.3	**17.1** 31.4	**25.7** 40.0	**25.7** 42.9	**25.7** 48.6
H4													**22.9** 42.9	**31.4** 51.4	**40.0** 54.3	**42.9** 51.4
H5														**14.3** 28.6	**25.7** 37.1	**22.9** 40.0
H6															**25.7** 42.9	**25.7** 45.7
H7																**34.3** 48.6

List of References

Reference List

Acosta-Serrano,A., Vassella,E., Liniger,M., Kunz,R.C., Brun,R., Roditi,I., and Englund,P.T. (2001) The surface coat of procyclic Trypanosoma brucei: programmed expression and proteolytic cleavage of procyclin in the tsetse fly. *Proc Natl Acad Sci U S A* **98**: 1513-1518.

Adam,R.D. (2000) The Giardia lamblia genome. *Int J Parasitol* **30**: 475-484.

Aitcheson,N., Talbot,S., Shapiro,J., Hughes,K., Adkin,C., Butt,T. *et al.* (2005) VSG switching in Trypanosoma brucei: antigenic variation analysed using RNAi in the absence of immune selection. *Mol Microbiol* **57**: 1608-1622.

Ajioka,J.W., Boothroyd,J.C., Brunk,B.P., Hehl,A., Hillier,L., Manger,I.D. *et al.* (1998) Gene Discovery by EST Sequencing in Toxoplasma gondii Reveals Sequences Restricted to theáApicomplexa. *Genome Research* **8**: 18-28.

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.

Alexiadis,V., and Kadonaga,J.T. (2002) Strand pairing by Rad54 and Rad51 is enhanced by chromatin. *Genes Dev* **16**: 2767-2771.

Allen,D.J., Makhov,A., Grilley,M., Taylor,J., Thresher,R., Modrich,P., and Griffith,J.D. (1997) MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J* **16**: 4467-4476.

Alsford,S., Kawahara,T., Isamah,C., and Horn,D. (2007) A sirtuin in the African trypanosome is involved in both DNA repair and telomeric gene silencing but is not required for antigenic variation. *Mol Microbiol* **63**: 724-736.

Aphasizhev,R., Aphasizheva,I., Nelson,R.E., and Simpson,L. (2003) A 100-kD complex of two RNA-binding proteins from mitochondria of Leishmania tarentolae catalyzes RNA annealing and interacts with several RNA editing components. *RNA* **9**: 62-76.

Arkhipova,I.R., and Morrison,H.G. (2001) Three retrotransposon families in the genome of Giardia lamblia: two telomeric, one dead. *Proc Natl Acad Sci U S A* **98**: 14497-14502.

Arnot,D.E., Roper,C., and Bayoumi,R.A. (1993) Digital codes from hypervariable tandemly repeated DNA sequences in the Plasmodium falciparum circumsporozoite gene can genetically barcode isolates. *Mol Biochem Parasitol* **61**: 15-24.

Au,K.G., Welsh,K., and Modrich,P. (1992) Initiation of methyl-directed mismatch repair. *J Biol Chem* **267**: 12142-12148.

Auffret,C.A., and Turner,M.J. (1981) Variant specific antigens of Trypanosoma brucei exist in solution as glycoprotein dimers. *Biochem J* **193**: 647-650.

Bangs,J.D., Crain,P.F., Hashizume,T., McCloskey,J.A., and Boothroyd,J.C. (1992) Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. *J Biol Chem* **267**: 9805-9815.

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

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. (1986) Antigenic variation during Trypanosoma vivax infections of different host species. *Parasitology* **92 (Pt 1)**: 51-65.

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

Barry,J.D., Ginger,M.L., Burton,P., and McCulloch,R. (2003) Why are parasite contingency genes often associated with telomeres? *Int J Parasitol* **33**: 29-45.

Barry,J.D., Marcello,L., Morrison,L.J., Read,A.F., Lythgoe,K., Jones,N. *et al.* (2005) What the genome sequence is revealing about trypanosome antigenic variation. *Biochem Soc Trans* **33**: 986-989.

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

Bauerschmidt,C., Pollok,S., Kremmer,E., Nasheuer,H.P., and Grosse,F. (2007) Interactions of human Cdc45 with the Mcm2-7 complex, the GINS complex, and DNA polymerases delta and epsilon during S phase. *Genes Cells* **12**: 745-758.

Baumann,P., Benson,F.E., and West,S.C. (1996) Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. *Cell* **87**: 757-766.

Becker,M., Aitcheson,N., Byles,E., Wickstead,B., Louis,E., and Rudenko,G. (2004) Isolation of the repertoire of VSG expression site containing telomeres of Trypanosoma brucei 427 using transformation-associated recombination in yeast. *Genome Res* **14**: 2319-2329.

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.

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.

Benson,F.E., Stasiak,A., and West,S.C. (1994) Purification and characterization of the human Rad51 protein, an analogue of E. coli RecA. *EMBO J* **13**: 5764-5771.

Beranek,D.T. (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **231**: 11-30.

Berggren,K., Steinberg,T.H., Lauber,W.M., Carroll,J.A., Lopez,M.F., Chernokalskaya,E. *et al.* (1999) A luminescent ruthenium complex for ultrasensitive detection of proteins immobilized on membrane supports. *Anal Biochem* **276**: 129-143.

Berriman,M., Ghedin,E., Hertz-Fowler,C., Blandin,G., Renauld,H., Bartholomeu,D.C. *et al.* (2005) The genome of the African trypanosome Trypanosoma brucei. *Science* **309**: 416-422.

Berriman,M., Hall,N., Sheader,K., Bringaud,F., Tiwari,B., Isobe,T. *et al.* (2002) The architecture of variant surface glycoprotein gene expression sites in Trypanosoma brucei. *Mol Biochem Parasitol* **122**: 131-140.

Bertwistle,D., Swift,S., Marston,N.J., Jackson,L.E., Crossland,S., Crompton,M.R. *et al.* (1997) Nuclear Location and Cell Cycle Regulation of the BRCA2 Protein. *Cancer Res* **57**: 5485-5488.

Bignell,G., Micklem,G., Stratton,M.R., Ashworth,A., and Wooster,R. (1997) The BRC repeats are conserved in mammalian BRCA2 proteins. *Hum Mol Genet* **6**: 53-58.

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.

Bitter,W., Gerrits,H., Kieft,R., and Borst,P. (1998) The role of transferrin-receptor variation in the host range of Trypanosoma brucei. *Nature* **391**: 499-502.

Blackstock,W.P., and Weir,M.P. (1999) Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol* **17**: 121-127.

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.

Bochkarev,A., Bochkareva,E., Frappier,L., and Edwards,A.M. (1999) The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding. *EMBO J* **18**: 4498-4504.

Bochkareva,E., Korolev,S., Lees-Miller,S.P., and Bochkarev,A. (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J* **21**: 1855-1863.

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.

Bork,P., Blomberg,N., and Nilges,M. (1996) Internal repeats in the BRCA2 protein sequence. *Nat Genet* **13**: 22-23.

Borst,P. (2002) Antigenic variation and allelic exclusion. *Cell* **109**: 5-8.

Borst,P. (1986) Discontinuous transcription and antigenic variation in trypanosomes. *Annu Rev Biochem* **55**: 701-732.

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

Borst,P., and Greaves,D.R. (1987) Programmed gene rearrangements altering gene expression. *Science* **235**: 658-667.

Borst,P., and Rudenko,G. (1994) Antigenic variation in African trypanosomes. *Science* **264**: 1872-1873.

Borst,P., and Ulbert,S. (2001) Control of VSG gene expression sites. *Mol Biochem Parasitol* **114**: 17-27.

Borst,P., and van Leeuwen,F. (1997) beta-D-glucosyl-hydroxymethyluracil, a novel base in African trypanosomes and other Kinetoplastida. *Mol Biochem Parasitol* **90**: 1-8.

Bosco,G., and Haber,J.E. (1998) Chromosome break-induced DNA replication leads to nonreciprocal translocations and telomere capture. *Genetics* **150**: 1037-1047.

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.

Bringaud,F., Biteau,N., Donelson,J.E., and Baltz,T. (2001) Conservation of metacyclic variant surface glycoprotein expression sites among different trypanosome isolates. *Mol Biochem Parasitol* **113**: 67-78.

Brough,R., Wei,D., Leulier,S., Lord,C.J., Rong,Y.S., and Ashworth,A. (2008) Functional analysis of Drosophila melanogaster BRCA2 in DNA repair. *DNA Repair* **7**: 10-19.

Browning,C.H., and Gulbransen,R. (1936) Immunity following cure of experimental Trypanosoma brucei infection by a chemotherapeutic agent. *The Journal of Pathology and Bacteriology* **43**: 479-486.

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

Buchhop,S., Gibson,M.K., Wang,X.W., Wagner,P., Sturzbecher,H.W., and Harris,C.C. (1997) Interaction of p53 with the human Rad51 protein. *Nucleic Acids Res* **25**: 3868-3874.

Burton,P., McBride,D.J., Wilkes,J.M., Barry,J.D., and McCulloch,R. (2007) Ku heterodimer-independent end joining in Trypanosoma brucei cell extracts relies upon sequence microhomology. *Eukaryot Cell* **6**: 1773-1781.

Callejas,S., Leech,V., Reitter,C., and Melville,S. (2006) Hemizygous subtelomeres of an African trypanosome chromosome may account for over 75% of chromosome length. *Genome Res* **16**: 1109-1118.

Capbern,A., Giroud,C., Baltz,T., and Mattern,P. (1977) [Trypanosoma equiperdum: antigenic variations in experimental trypanosomiasis of rabbits]. *Exp Parasitol* **42**: 6-13.

Carlton,J.M., Hirt,R.P., Silva,J.C., Delcher,A.L., Schatz,M., Zhao,Q. *et al.* (2007) Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. *Science* **315**: 207-212.

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.

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.

Chen,C.F., Chen,P.L., Zhong,Q., Sharp,Z.D., and Lee,W.H. (1999a) Expression of BRC repeats in breast cancer cells disrupts the BRCA2-Rad51 complex and leads to radiation hypersensitivity and loss of G(2)/M checkpoint control. *J Biol Chem* **274**: 32931-32935.

Chen,G., Yuan,S.S., Liu,W., Xu,Y., Trujillo,K., Song,B. *et al.* (1999b) Radiation-induced assembly of Rad51 and Rad52 recombination complex requires ATM and c-Abl. *J Biol Chem* **274**: 12748-12752.

Chen,J., Silver,D.P., Walpita,D., Cantor,S.B., Gazdar,A.F., Tomlinson,G. *et al.* (1998a) Stable interaction between the products of the BRCA1 and BRCA2 tumor suppressor genes in mitotic and meiotic cells. *Mol Cell* **2**: 317-328.

Chen,J.M., Cooper,D.N., Chuzhanova,N., Ferec,C., and Patrinos,G.P. (2007) Gene conversion: mechanisms, evolution and human disease. *Nat Rev Genet* **8**: 762-775.

Chen,P.L., Chen,C.F., Chen,Y., Xiao,J., Sharp,Z.D., and Lee,W.H. (1998b) 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.

Chen,X.B., Melchionna,R., Denis,C.M., Gaillard,P.H., Blasina,A., Van,d.W., I *et al.* (2001) Human Mus81-associated endonuclease cleaves Holliday junctions in vitro. *Mol Cell* **8**: 1117-1127.

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.

Choy,J.S., and Kron,S.J. (2002) NuA4 Subunit Yng2 Function in Intra-S-Phase DNA Damage Response. *Mol Cell Biol* **22**: 8215-8225.

Chu,G. (1997) Double strand break repair. *J Biol Chem* **272**: 24097-24100.

Ciccia,A., Ling,C., Coulthard,R., Yan,Z., Xue,Y., Meetei,A.R. *et al.* (2007) Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. *Mol Cell* **25**: 331-343.

Clayton,C.E. (2002) Life without transcriptional control? From fly to man and back again. *EMBO J* **21**: 1881-1888.

Connelly,J.C., and Leach,D.R. (2002) Tethering on the brink: the evolutionarily conserved Mre11-Rad50 complex. *Trends Biochem Sci* **27**: 410-418.

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. (2002b) 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., 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. (2002c) Two pathways of homologous recombination in Trypanosoma brucei. *Mol Microbiol* **45**: 1687-1700.

Cortez,A.P., Ventura,R.M., Rodrigues,A.C., Batista,J.S., Paiva,F., Anez,N. *et al.* (2006) The taxonomic and phylogenetic relationships of Trypanosoma vivax from South America and Africa. *Parasitology* **133**: 159-169.

Crackower,M.A., Scherer,S.W., Rommens,J.M., Hui,C.C., Poorkaj,P., Soder,S. *et al.* (1996) Characterization of the split hand/split foot malformation locus SHFM1 at 7q21.3-q22.1 and analysis of a candidate gene for its expression during limb development. *Human Molecular Genetics* **5**: 571-579.

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

Cramer,P. (2002) Multisubunit RNA polymerases. *Curr Opin Struct Biol* **12**: 89-97.

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

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,G.A. (1990) Cellular and genetic aspects of antigenic variation in trypanosomes. *Annu Rev Immunol* **8**: 83-110.

Cross,G.A. (1996) Antigenic variation in trypanosomes: secrets surface slowly. *Bioessays* **18**: 283-291.

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

Cross,G.A., Wirtz,L.E., and Navarro,M. (1998) Regulation of vsg expression site transcription and switching in Trypanosoma brucei. *Mol Biochem Parasitol* **91**: 77-91.

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. *et al.* (1999) The modified base J is the target for a novel DNA-binding protein in kinetoplastid protozoans. *EMBO J* **18**: 6573-6581.

D'Andrea,A.D., and Grompe,M. (2003) The Fanconi anaemia/BRCA pathway. *Nat Rev Cancer* **3**: 23-34.

Dai,Q., Restrepo,B.I., Porcella,S.F., Raffel,S.J., Schwan,T.G., and Barbour,A.G. (2006) Antigenic variation by Borrelia hermsii occurs through recombination between extragenic repetitive elements on linear plasmids. *Mol Microbiol* **60**: 1329-1343.

Daniels,M.J., Wang,Y., Lee,M., and Venkitaraman,A.R. (2004) Abnormal cytokinesis in cells deficient in the breast cancer susceptibility protein BRCA2. *Science* **306**: 876-879.

Dao,V., and Modrich,P. (1998) Mismatch-, MutS-, MutL-, and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex. *J Biol Chem* **273**: 9202-9207.

Datta,A., Adjiri,A., New,L., Crouse,G.F., and Jinks,R.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.

Davies,O.R., and Pellegrini,L. (2007) Interaction with the BRCA2 C terminus protects RAD51-DNA filaments from disassembly by BRC repeats. *Nat Struct Mol Biol*.

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

de Jager,M., Dronkert,M.L., Modesti,M., Beerens,C.E., Kanaar,R., and van Gent,D.C. (2001) DNA-binding and strand-annealing activities of human Mre11: implications for its roles in DNA double-strand break repair pathways. *Nucleic Acids Res* **29**: 1317-1325.

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

Deans,B., Griffin,C.S., Maconochie,M., and Thacker,J. (2000) Xrcc2 is required for genetic stability, embryonic neurogenesis and viability in mice. *EMBO J* **19**: 6675-6685.

Deitsch,K.W., Moxon,E.R., and Wellems,T.E. (1997) Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol Mol Biol Rev* **61**: 281-293.

Delibas,S.B., Ertabaklar,H., and Ertug,S. (2006) Evaluation of antigenic variations between two virulent toxoplasma strains. *J Med Microbiol* **55**: 1333-1335.

Della,M., Palmbos,P.L., Tseng,H.M., Tonkin,L.M., Daley,J.M., Topper,L.M. *et al.* (2004) Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* **306**: 683-685.

Devaux,S., Lecordier,L., Uzureau,P., Walgraffe,D., Dierick,J.F., Poelvoorde,P. *et al.* (2006) Characterization of RNA polymerase II subunits of Trypanosoma brucei. *Mol Biochem Parasitol* **148**: 60-68.

Dingwall,C., Robbins,J., Dilworth,S.M., Roberts,B., and Richardson,W.D. (1988) The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *The Journal of Cell Biology* **107**: 841-849.

Dipaolo,C., Kieft,R., Cross,M., and Sabatini,R. (2005) Regulation of Trypanosome DNA Glycosylation by a SWI/SNF2-like Protein. *Mol Cell* **17**: 441-451.

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

Dong,Y., Hakimi,M.A., Chen,X., Kumaraswamy,E., Cooch,N.S., Godwin,A.K., and Shiekhattar,R. (2003) Regulation of BRCC, a holoenzyme complex containing BRCA1

and BRCA2, by a signalosome-like subunit and its role in DNA repair. *Mol Cell* **12**: 1087-1099.

Donoho,G., Brenneman,M.A., Cui,T.X., Donoviel,D., Vogel,H., Goodwin,E.H. *et al.* (2003) Deletion of Brca2 exon 27 causes hypersensitivity to DNA crosslinks, chromosomal instability, and reduced life span in mice. *Genes Chromosomes Cancer* **36**: 317-331.

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.

Dougherty,W.G., Cary,S.M., and Parks,T.D. (1989) Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. *Virology* **171**: 356-364.

Drakas,R., Prisco,M., and Baserga,R. (2005) A modified tandem affinity purification tag technique for the purification of protein complexes in mammalian cells. *Proteomics* **5**: 132-137.

Dray,E., Siaud,N., Dubois,E., and Doutriaux,M.P. (2006) Interaction between Arabidopsis Brca2 and its partners Rad51, Dmc1, and Dss1. *Plant Physiol* **140**: 1059-1069.

Dreesen,O., and Cross,G.A. (2006) Consequences of telomere shortening at an active VSG expression site in telomerase-deficient Trypanosoma brucei. *Eukaryot Cell* **5**: 2114-2119.

Eisen,J.A., Sweder,K.S., and Hanawalt,P.C. (1995) Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res* **23**: 2715-2723.

El Sayed,N.M., Ghedin,E., Song,J., MacLeod,A., Bringaud,F., Larkin,C. *et al.* (2003) The sequence and analysis of Trypanosoma brucei chromosome II. *Nucleic Acids Res* **31**: 4856-4863.

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.

El Sayed,N.M., Myler,P.J., Bartholomeu,D.C., Nilsson,D., Aggarwal,G., Tran,A.N. *et al.* (2005a) The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease. *Science* **309**: 409-415.

El Sayed,N.M., Myler,P.J., Blandin,G., Berriman,M., Crabtree,J., Aggarwal,G. *et al.* (2005b) Comparative genomics of trypanosomatid parasitic protozoa. *Science* **309**: 404-409.

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.

Emery,H.S., Schild,D., Kellogg,D.E., and Mortimer,R.K. (1991) Sequence of RAD54, a Saccharomyces cerevisiae gene involved in recombination and repair. *Gene* **104**: 103-106.

Esashi,F., Christ,N., Gannon,J., Liu,Y., Hunt,T., Jasin,M., and West,S.C. (2005) CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. *Nature* **434**: 598-604.

Esashi,F., Galkin,V.E., Yu,X., Egelman,E.H., and West,S.C. (2007) Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. *Nat Struct Mol Biol*.

Esposito,M.S. (1978) Evidence that spontaneous mitotic recombination occurs at the two-strand stage. *Proc Natl Acad Sci U S A* **75**: 4436-4440.

Essers,J., Hendriks,R.W., Wesoly,J., Beerens,C.E., Smit,B., Hoeijmakers,J.H. *et al.* (2002a) Analysis of mouse Rad54 expression and its implications for homologous recombination. *DNA Repair (Amst)* **1**: 779-793.

Essers,J., Houtsmuller,A.B., van Veelen,L., Paulusma,C., Nigg,A.L., Pastink,A. *et al.* (2002b) Nuclear dynamics of RAD52 group homologous recombination proteins in response to DNA damage. *EMBO J* **21**: 2030-2037.

Featherstone,C., and Jackson,S.P. (1999) DNA double-strand break repair. *Curr Biol* **9**: R759-R761.

Feng,J.A., Johnson,R.C., and Dickerson,R.E. (1994) Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. *Science* **263**: 348-355.

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,M.A., Homans,S.W., Dwek,R.A., and Rademacher,T.W. (1988) The glycosylphosphatidylinositol membrane anchor of Trypanosoma brucei variant surface glycoprotein. *Biochem Soc Trans* **16**: 265-268.

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

French,C.A., Masson,J.Y., Griffin,C.S., O'Regan,P., West,S.C., and Thacker,J. (2002) Role of mammalian RAD51L2 (RAD51C) in recombination and genetic stability. *J Biol Chem* **277**: 19322-19330.

Fromont-Racine,M., Rain,J.C., and Legrain,P. (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet* **16**: 277-282.

Fuks,F., Milner,J., and Kouzarides,T. (1998) BRCA2 associates with acetyltransferase activity when bound to P/CAF. *Oncogene* **17**: 2531-2534.

Futamura,M., Arakawa,H., Matsuda,K., Katagiri,T., Saji,S., Miki,Y., and Nakamura,Y. (2000) Potential role of BRCA2 in a mitotic checkpoint after phosphorylation by hBUBR1. *Cancer Res* **60**: 1531-1535.

Futse,J.E., Brayton,K.A., Knowles,D.P., Jr., and Palmer,G.H. (2005) Structural basis for segmental gene conversion in generation of Anaplasma marginale outer membrane protein variants. *Mol Microbiol* **57**: 212-221.

Galio,L., Bouquet,C., and Brooks,P. (1999) ATP hydrolysis-dependent formation of a dynamic ternary nucleoprotein complex with MutS and MutL. *Nucleic Acids Res* **27**: 2325-2331.

Galkin,V.E., Esashi,F., Yu,X., Yang,S., West,S.C., and Egelman,E.H. (2005) BRCA2 BRC motifs bind RAD51-DNA filaments. *Proc Natl Acad Sci U S A* **102**: 8537-8542.

Garcia-Higuera,I., Taniguchi,T., Ganesan,S., Meyn,M.S., Timmers,C., Hejna,J. *et al.* (2001) Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell* **7**: 249-262.

Garcia-Salcedo,J.A., Gijon,P., Nolan,D.P., Tebabi,P., and Pays,E. (2003) A chromosomal SIR2 homologue with both histone NAD-dependent ADP-ribosyltransferase and deacetylase activities is involved in DNA repair in Trypanosoma brucei. *EMBO J* **22**: 5851-5862.

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.

Gayther,S.A., Mangion,J., Russell,P., Seal,S., Barfoot,R., Ponder,B.A. *et al.* (1997) Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. *Nat Genet* **15**: 103-105.

Gibson,W., Kanmogne,G., and Bailey,M. (1995) A successful backcross in Trypanosoma brucei. *Mol Biochem Parasitol* **69**: 101-110.

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.

Glover,L., Alsford,S., Beattie,C., and Horn,D. (2007) Deletion of a trypanosome telomere leads to loss of silencing and progressive loss of terminal DNA in the absence of cell cycle arrest. *Nucleic Acids Res* **35**: 872-880.

Glover,L., McCulloch,R., and Horn,D. (2008) Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Res*.

Godthelp,B.C., Artwert,F., Joenje,H., and Zdzienicka,M.Z. (2002) Impaired DNA damage-induced nuclear Rad51 foci formation uniquely characterizes Fanconi anemia group D1. *Oncogene* **21**: 5002-5005.

Goggins,M., Schutte,M., Lu,J., Moskaluk,C.A., Weinstein,C.L., Petersen,G.M. *et al.* (1996) Germline BRCA2 gene mutations in patients with apparently sporadic pancreatic carcinomas. *Cancer Res* **56**: 5360-5364.

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,W.G., van Dijk,W.J., and Borst,P. (1993) The identification of hydroxymethyluracil in DNA of Trypanosoma brucei. *Nucleic Acids Res* **21**: 2039-2043.

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

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.

Grompe,M., and D'Andrea,A. (2001) Fanconi anemia and DNA repair. *Hum Mol Genet* **10**: 2253-2259.

Gruszynski,A.E., van Deursen,F.J., Albareda,M.C., Best,A., Chaudhary,K., Cliffe,L.J. *et al.* (2006) Regulation of surface coat exchange by differentiating African trypanosomes. *Mol Biochem Parasitol* **147**: 211-223.

Gudmundsdottir,K., Lord,C.J., Witt,E., Tutt,A.N., and Ashworth,A. (2004) DSS1 is required for RAD51 focus formation and genomic stability in mammalian cells. *EMBO Rep* **5**: 989-993.

Gull,K., Alsford,S., and Ersfeld,K. (1998) Segregation of minichromosomes in trypanosomes: implications for mitotic mechanisms. *Trends Microbiol* **6**: 319-323.

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

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.

Haber,J.E., and Hearn,M. (1985) Rad52-independent mitotic gene conversion in Saccharomyces cerevisiae frequently results in chromosomal loss. *Genetics* **111**: 7-22.

Haber,J.E., Ira,G., Malkova,A., and Sugawara,N. (2004) Repairing a double-strand chromosome break by homologous recombination: revisiting Robin Holliday's model. *Philos Trans R Soc Lond B Biol Sci* **359**: 79-86.

Hajduk,S.L., Hager,K., and Esko,J.D. (1992) High-density lipoprotein-mediated lysis of trypanosomes. *Parasitol Today* **8**: 95-98.

Hall,J.M., Lee,M.K., Newman,B., Morrow,J.E., Anderson,L.A., Huey,B., and King,M.C. (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* **250**: 1684-1689.

Hall,M.N., Hereford,L., and Herskowitz,I. (1984) Targeting of E. coli [beta]-galactosidase to the nucleus in yeast. *Cell* **36**: 1057-1065.

Hall,N., Berriman,M., Lennard,N.J., Harris,B.R., Hertz-Fowler,C., Bart-Delabesse,E.N. *et al.* (2003) The DNA sequence of chromosome I of an African trypanosome: gene content, chromosome organisation, recombination and polymorphism. *Nucleic Acids Res* **31**: 4864-4873.

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.

Hammarton,T.C., Clark,J., Douglas,F., Boshart,M., and Mottram,J.C. (2003) Stage-specific differences in cell cycle control in Trypanosoma brucei revealed by RNA interference of a mitotic cyclin. *J Biol Chem* **278**: 22877-22886.

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.

Hedges,S.B., Blair,J.E., Venturi,M.L., and Shoe,J.L. (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol Biol* **4**: 2.

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.

Hettmann,C., and Soldati,D. (1999) Cloning and analysis of a Toxoplasma gondii histone acetyltransferase: a novel chromatin remodelling factor in Apicomplexan parasites. *Nucl Acids Res* **27**: 4344-4352.

Hicks,G.R., and Raikhel,N.V. (1995) Protein import into the nucleus: an integrated view. *Annu Rev Cell Dev Biol* **11**: 155-188.

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.

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 Barry,J.D. (2005) The central roles of telomeres and subtelomeres in antigenic variation in African trypanosomes. *Chromosome Res* **13**: 525-533.

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

Howlett,N.G., Taniguchi,T., Olson,S., Cox,B., Waisfisz,Q., Die-Smulders,C. *et al.* (2002) Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* **297**: 606-609.

Hughes-Davies,L., Huntsman,D., Ruas,M., Fuks,F., Bye,J., Chin,S.F. *et al.* (2003) EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell* **115**: 523-535.

Hussain,S., Wilson,J.B., Medhurst,A.L., Hejna,J., Witt,E., Ananth,S. *et al.* (2004) Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways. *Hum Mol Genet* **13**: 1241-1248.

Hussain,S., Witt,E., Huber,P.A., Medhurst,A.L., Ashworth,A., and Mathew,C.G. (2003) Direct interaction of the Fanconi anaemia protein FANCG with BRCA2/FANCD1. *Hum Mol Genet* **12**: 2503-2510.

Ira,G., and Haber,J.E. (2002) Characterization of RAD51-independent break-induced replication that acts preferentially with short homologous sequences. *Mol Cell Biol* **22**: 6384-6392.

Ira,G., Satory,D., and Haber,J.E. (2006) Conservative inheritance of newly synthesized DNA in double-strand break-induced gene conversion. *Mol Cell Biol* **26**: 9424-9429.

Ivanov,E.L., Sugawara,N., Fishman-Lobell,J., and Haber,J.E. (1996) Genetic requirements for the single-strand annealing pathway of double- strand break repair in Saccharomyces cerevisiae. *Genetics* **142**: 693-704.

Jasin,M. (2002) Homologous repair of DNA damage and tumorigenesis: the BRCA connection. *Oncogene* **21**: 8981-8993.

Jeffreys,A.J., Wilson,V., and Thein,S.L. (1985) Hypervariable 'minisatellite' regions in human DNA. *Nature* **314**: 67-73.

Jeffreys,A.J., MacLeod,A., Tamaki,K., Neil,D.L., and Monckton,D.G. (1991) Minisatellite repeat coding as a digital approach to DNA typing. *Nature* **354**: 204-209.

Jeffreys,A.J., Neumann,R., and Wilson,V. (1990) Repeat unit sequence variation in minisatellites: A novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell* **60**: 473-485.

Jiricny,J. (2006) The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* **7**: 335-346.

Joenje,H., and Patel,K.J. (2001) The emerging genetic and molecular basis of Fanconi anaemia. *Nat Rev Genet* **2**: 446-457.

Johnson,P.J., Kooter,J.M., and Borst,P. (1987) Inactivation of transcription by UV irradiation of T. brucei provides evidence for a multicistronic transcription unit including a VSG gene. *Cell* **51**: 273-281.

Johnson,R.D., Liu,N., and Jasin,M. (1999) Mammalian XRCC2 promotes the repair of DNA double-strand breaks by homologous recombination. *Nature* **401**: 397-399.

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.

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.

Kalderon,D., Richardson,W.D., Markham,A.F., and Smith,A.E. (1984) Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* **311**: 33-38.

Keely,S.P., Renauld,H., Wakefield,A.E., Cushion,M.T., Smulian,A.G., Fosker,N. *et al.* (2005) Gene arrays at Pneumocystis carinii telomeres. *Genetics* **170**: 1589-1600.

Kennedy,P.G. (2004) Human African trypanosomiasis of the CNS: current issues and challenges. *J Clin Invest* **113**: 496-504.

Khanna,K.K., and Jackson,S.P. (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* **27**: 247-254.

Kieft,R., Brand,V., Ekanayake,D.K., Sweeney,K., Dipaolo,C., Reznikoff,W.S., and Sabatini,R. (2007) JBP2, a SWI2/SNF2-like protein, regulates de novo telomeric DNA glycosylation in bloodstream form Trypanosoma brucei. *Mol Biochem Parasitol* **156**: 24-31.

Kinebuchi,T., Kagawa,W., Enomoto,R., Tanaka,K., Miyagawa,K., Shibata,T. *et al.* (2004) Structural basis for octameric ring formation and DNA interaction of the human homologous-pairing protein Dmc1. *Mol Cell* **14**: 363-374.

Kojic,M., and Holloman,W.K. (2004) BRCA2-RAD51-DSS1 interplay examined from a microbial perspective. *Cell Cycle* **3**: 247-248.

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.

Kojic,M., Yang,H., Kostrub,C.F., Pavletich,N.P., and Holloman,W.K. (2003) The BRCA2-interacting protein DSS1 is vital for DNA repair, recombination, and genome stability in Ustilago maydis. *Mol Cell* **12**: 1043-1049.

Kojic,M., Zhou,Q., Lisby,M., and Holloman,W.K. (2006) Rec2 interplay with both Brh2 and Rad51 balances recombinational repair in Ustilago maydis. *Mol Cell Biol* **26**: 678-688.

Kojic,M., Zhou,Q., Lisby,M., and Holloman,W.K. (2005) Brh2-Dss1 interplay enables properly controlled recombination in Ustilago maydis. *Mol Cell Biol* **25**: 2547-2557.

Kooter,J.M., and Borst,P. (1984) Alpha-amanitin-insensitive transcription of variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. *Nucleic Acids Res* **12**: 9457-9472.

Kovalenko,O.V., Golub,E.I., Bray-Ward,P., Ward,D.C., and Radding,C.M. (1997) A novel nucleic acid-binding protein that interacts with human rad51 recombinase. *Nucleic Acids Res* **25**: 4946-4953.

Kowalczykowski,S.C. (2002) Molecular mimicry connects BRCA2 to Rad51 and recombinational DNA repair. *Nat Struct Biol* **9**: 897-899.

Kraakman-van der Zwet,M., Overkamp,W.J., van Lange,R.E., Essers,J., Duijn-Goedhart,A., Wiggers,I. *et al.* (2002) Brca2 (XRCC11) deficiency results in radioresistant DNA synthesis and a higher frequency of spontaneous deletions. *Mol Cell Biol* **22**: 669-679.

Kraemer,S.M., and Smith,J.D. (2006) A family affair: var genes, PfEMP1 binding, and malaria disease. *Curr Opin Microbiol* **9**: 374-380.

Krogan,N.J., Cagney,G., Yu,H., Zhong,G., Guo,X., Ignatchenko,A. *et al.* (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440**: 637-643.

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

Kulakova,L., Singer,S.M., Conrad,J., and Nash,T.E. (2006) Epigenetic mechanisms are involved in the control of Giardia lamblia antigenic variation. *Mol Microbiol* **61**: 1533-1542.

Kulasekara,H.D., and Blomfield,I.C. (1999) The molecular basis for the specificity of fimE in the phase variation of type 1 fimbriae of Escherichia coli K-12. *Mol Microbiol* **31**: 1171-1181.

Kunkel,T.A., and Erie,D.A. (2005) DNA mismatch repair. *Annu Rev Biochem* **74**: 681-710.

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., and Petes,T.D. (1988) Allelic and ectopic recombination between Ty elements in yeast. *Genetics* **119**: 549-559.

Kurumizaka,H., Aihara,H., Kagawa,W., Shibata,T., and Yokoyama,S. (1999) Human Rad51 amino acid residues required for Rad52 binding. *J Mol Biol* **291**: 537-548.

Kuzminov,A. (1995) A mechanism for induction of the SOS response in E. coli: insights into the regulation of reversible protein polymerization in vivo. *J Theor Biol* **177**: 29-43.

Kyes,S.A., Kraemer,S.M., and Smith,J.D. (2007) Antigenic variation in Plasmodium falciparum: gene organization and regulation of the var multigene family. *Eukaryot Cell* **6**: 1511-1520.

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.

Lanford,R.E., and Butel,J.S. (1984) Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* **37**: 801-813.

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

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

Lee,M., Daniels,M.J., and Venkitaraman,A.R. (2004) Phosphorylation of BRCA2 by the Polo-like kinase Plk1 is regulated by DNA damage and mitotic progression. *Oncogene* **23**: 865-872.

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.

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

Levinson,G., and Gutman,G.A. (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol* **4**: 203-221.

Levitus,M., Joenje,H., and de Winter,J.P. (2006) The Fanconi anemia pathway of genomic maintenance. *Cell Oncol* **28**: 3-29.

Li,W., Hesabi,B., Babbo,A., Pacione,C., Liu,J., Chen,D.J. *et al.* (2000) Regulation of double-strand break-induced mammalian homologous recombination by UBL1, a RAD51-interacting protein. *Nucleic Acids Res* **28**: 1145-1153.

Liang,F., Han,M., Romanienko,P.J., and Jasin,M. (1998) Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc Natl Acad Sci U S A* **95**: 5172-5177.

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.

Lin,H.R., Ting,N.S., Qin,J., and Lee,W.H. (2003) M phase-specific phosphorylation of BRCA2 by Polo-like kinase 1 correlates with the dissociation of the BRCA2-P/CAF complex. *J Biol Chem* **278**: 35979-35987.

Lindahl,T., and Wood,R.D. (1999) Quality control by DNA repair. *Science* **286**: 1897-1905.

Ling,C., Ishiai,M., Ali,A.M., Medhurst,A.L., Neveling,K., Kalb,R. *et al.* (2007) FAAP100 is essential for activation of the Fanconi anemia-associated DNA damage response pathway. *EMBO J* **26**: 2104-2114.

Liniger,M., Urwyler,S., Studer,E., Oberle,M., Renggli,C.K., and Roditi,I. (2004) Role of the N-terminal domains of EP and GPEET procyclins in membrane targeting and the establishment of midgut infections by Trypanosoma brucei. *Mol Biochem Parasitol* **137**: 247-251.

Linke,S.P., Sengupta,S., Khabie,N., Jeffries,B.A., Buchhop,S., Miska,S. *et al.* (2003) p53 interacts with hRAD51 and hRAD54, and directly modulates homologous recombination. *Cancer Res* **63**: 2596-2605.

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,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.

Liu,Y., Tarsounas,M., O'Regan,P., and West,S.C. (2007) Role of RAD51C and XRCC3 in genetic recombination and DNA repair. *J Biol Chem* **282**: 1973-1979.

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.

Lomonosov,M., Anand,S., Sangrithi,M., Davies,R., and Venkitaraman,A.R. (2003) Stabilization of stalled DNA replication forks by the BRCA2 breast cancer susceptibility protein. *Genes Dev* **17**: 3017-3022.

Lord,C.J., and Ashworth,A. (2007) RAD51, BRCA2 and DNA repair: a partial resolution. *Nat Struct Mol Biol* **14**: 461-462.

Lu,H., Guo,X., Meng,X., Liu,J., Allen,C., Wray,J. *et al.* (2005) The BRCA2-interacting protein BCCIP functions in RAD51 and BRCA2 focus formation and homologous recombinational repair. *Mol Cell Biol* **25**: 1949-1957.

Lundblad,V., and Blackburn,E.H. (1993) An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* **73**: 347-360.

Lustig,Y., Goldshmidt,H., Uliel,S., and Michaeli,S. (2005) The Trypanosoma brucei signal recognition particle lacks the Alu-domain-binding proteins: purification and functional analysis of its binding proteins by RNAi. *J Cell Sci* **118**: 4551-4562.

Lynch,H.T., Albano,W.A., Danes,B.S., Layton,M.A., Kimberling,W.J., Lynch,J.F. *et al.* (1984) Genetic predisposition to breast cancer. *Cancer* **53**: 612-622.

MacLeod,A., Turner,C.M., and Tait,A. (2001a) The detection of geographical substructuring of Trypanosoma brucei populations by the analysis of minisatellite polymorphisms. *Parasitology* **123**: 475-482.

MacLeod,A., Welburn,S., Maudlin,I., Turner,C.M., and Tait,A. (2001b) Evidence for multiple origins of human infectivity in Trypanosoma brucei revealed by minisatellite variant repeat mapping. *J Mol Evol* **52**: 290-301.

Majiwa,P.A., Matthyssens,G., Williams,R.O., and Hamers,R. (1985) Cloning and analysis of Trypanosoma (Nannomonas) congolense ILNat 2.1 VSG gene. *Mol Biochem Parasitol* **16**: 97-108.

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.

Marcello,L., and Barry,J.D. (2007a) From silent genes to noisy populations-dialogue between the genotype and phenotypes of antigenic variation. *J Eukaryot Microbiol* **54**: 14-17.

Marcello,L., and Barry,J.D. (2007b) Analysis of the VSG gene silent archive in Trypanosoma brucei reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Res* **17**: 1344-1352.

Marcello,L., Menon,S., Ward,P., Wilkes,J.M., Jones,N.G., Carrington,M., and Barry,J.D. (2007) VSGdb: a database for trypanosome variant surface glycoproteins, a large and diverse family of coiled coil proteins. *BMC Bioinformatics* **8**: 143.

Marmorstein,L.Y., Kinev,A.V., Chan,G.K., Bochar,D.A., Beniya,H., Epstein,J.A. *et al.* (2001) A human BRCA2 complex containing a structural DNA binding component influences cell cycle progression. *Cell* **104**: 247-257.

Marmorstein,L.Y., Ouchi,T., and Aaronson,S.A. (1998) The BRCA2 gene product functionally interacts with p53 and RAD51. *Proc Natl Acad Sci U S A* **95**: 13869-13874.

Marston,N.J., Richards,W.J., Hughes,D., Bertwistle,D., Marshall,C.J., and Ashworth,A. (1999) Interaction between the product of the breast cancer susceptibility gene BRCA2 and DSS1, a protein functionally conserved from yeast to mammals. *Mol Cell Biol* **19**: 4633-4642.

Martin,J.S., Winkelmann,N., Petalcorin,M.I., McIlwraith,M.J., and Boulton,S.J. (2005) RAD-51-Dependent and -Independent Roles of a Caenorhabditis elegans BRCA2-Related Protein during DNA Double-Strand Break Repair. *Mol Cell Biol* **25**: 3127-3139.

Maslov,D.A., Podlipaev,S.A., and Lukes,J. (2001) Phylogeny of the kinetoplastida: taxonomic problems and insights into the evolution of parasitism. *Mem Inst Oswaldo Cruz* **96**: 397-402.

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. *et al.* (2001b) Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev* **15**: 3296-3307.

Mathew,C.G. (2006) Fanconi anaemia genes and susceptibility to cancer. *Oncogene* **25**: 5875-5884.

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., and Gull,K. (1997) Commitment to differentiation and cell cycle re-entry are coincident but separable events in the transformation of African trypanosomes from their bloodstream to their insect form. *J Cell Sci* **110 (Pt 20)**: 2609-2618.

Matthews,K.R., and Gull,K. (1994) Evidence for an interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes. *J Cell Biol* **125**: 1147-1156.

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., Alexeev,A.A., and Kowalczykowski,S.C. (2003) A novel function of Rad54 protein. Stabilization of the Rad51 nucleoprotein filament. *J Biol Chem* **278**: 14029-14036.

McAllister,K.A., Bennett,L.M., Houle,C.D., Ward,T., Malphurs,J., Collins,N.K. *et al.* (2002) Cancer susceptibility of mice with a homozygous deletion in the COOH-terminal domain of the Brca2 gene. *Cancer Res* **62**: 990-994.

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.

McEachern,M.J., and Haber,J.E. (2006) Break-Induced Replication and Recombinational Telomere Elongation in Yeast. *Annu Rev Biochem*.

McIlwraith,M.J., Van Dyck,E., Masson,J.Y., Stasiak,A.Z., Stasiak,A., and West,S.C. (2000) Reconstitution of the strand invasion step of double-strand break repair using human Rad51 Rad52 and RPA proteins. *J Mol Biol* **304**: 151-164.

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.

Meetei,A.R., Medhurst,A.L., Ling,C., Xue,Y., Singh,T.R., Bier,P. *et al.* (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet* **37**: 958-963.

Mehr,I.J., and Seifert,H.S. (1998) Differential roles of homologous recombination pathways in Neisseria gonorrhoeae pilin antigenic variation, DNA transformation and DNA repair. *Mol Microbiol* **30**: 697-710.

Melen,K., and Julkunen,I. (1997) Nuclear Cotransport Mechanism of Cytoplasmic Human MxB Protein. *Journal of Biological Chemistry* **272**: 32353-32359.

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.

Michel,B. (2000) Replication fork arrest and DNA recombination. *Trends Biochem Sci* **25**: 173-178.

Michel,B., Flores,M.J., Viguera,E., Grompone,G., Seigneur,M., and Bidnenko,V. (2001) Rescue of arrested replication forks by homologous recombination. *Proc Natl Acad Sci U S A* **98**: 8181-8188.

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.

Michels,P.A., Liu,A.Y., Bernards,A., Sloof,P., Van der Bijl,M.M., Schinkel,A.H. *et al.* (1983) Activation of the genes for variant surface glycoproteins 117 and 118 in Trypanosoma brucei. *J Mol Biol* **166**: 537-556.

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

Mizuta,R., LaSalle,J.M., Cheng,H.L., Shinohara,A., Ogawa,H., Copeland,N. *et al.* (1997) RAB22 and RAB163/mouse BRCA2: proteins that specifically interact with the RAD51 protein. *Proc Natl Acad Sci U S A* **94**: 6927-6932.

Molyneux,D.H. (1983) Selective primary health care: strategies for control of disease in the developing world. VIII. African trypanosomiasis. *Rev Infect Dis* **5**: 945-956.

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.

Moreno-Herrero,F., de Jager,M., Dekker,N.H., Kanaar,R., Wyman,C., and Dekker,C. (2005) Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. *Nature* **437**: 440-443.

Morrison,L.J., Majiwa,P., Read,A.F., and Barry,J.D. (2005) Probabilistic order in antigenic variation of Trypanosoma brucei. *Int J Parasitol* **35**: 961-972.

Mowatt,M.R., Aggarwal,A., and Nash,T.E. (1991) Carboxy-terminal sequence conservation among variant-specific surface proteins of Giardia lamblia. *Mol Biochem Parasitol* **49**: 215-227.

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

Moynahan,M.E., Chiu,J.W., Koller,B.H., and Jasin,M. (1999) Brca1 controls homology-directed DNA repair. *Mol Cell* **4**: 511-518.

Moynahan,M.E., Pierce,A.J., and Jasin,M. (2001) BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol Cell* **7**: 263-272.

Nagoshi,Y.L., Alarcon,C.M., and Donelson,J.E. (1995) The putative promoter for a metacyclic VSG gene in African trypanosomes. *Mol Biochem Parasitol* **72**: 33-45.

Nakada,D., Shimomura,T., Matsumoto,K., and Sugimoto,K. (2003) The ATM-related Tel1 protein of Saccharomyces cerevisiae controls a checkpoint response following phleomycin treatment. *Nucleic Acids Res* **31**: 1715-1724.

Nakanishi,A., Han,X., Saito,H., Taguchi,K., Ohta,Y., Imajoh-Ohmi,S., and Miki,Y. (2007) Interference with BRCA2, which localizes to the centrosome during S and early M phase, leads to abnormal nuclear division. *Biochem Biophys Res Commun* **355**: 34-40.

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

Neto,J.L., Lira,C.B., Giardini,M.A., Khater,L., Perez,A.M., Peroni,L.A. *et al.* (2007) Leishmania replication protein A-1 binds in vivo single-stranded telomeric DNA. *Biochem Biophys Res Commun* **358**: 417-423.

Niedzwiedz,W., Mosedale,G., Johnson,M., Ong,C.Y., Pace,P., and Patel,K.J. (2004) The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Mol Cell* **15**: 607-620.

O'Brien,J., Wilson,I., Orton,T., and Pognan,F. (2000) Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. *Eur J Biochem* **267**: 5421-5426.

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.

Ochiai,K., Morimatsu,M., Tomizawa,N., and Syuto,B. (2001) Cloning and sequencing full length of canine Brca2 and Rad51 cDNA. *J Vet Med Sci* **63**: 1103-1108.

Oli,M.W., Cotlin,L.F., Shiflett,A.M., and Hajduk,S.L. (2006) Serum resistance-associated protein blocks lysosomal targeting of trypanosome lytic factor in Trypanosoma brucei. *Eukaryot Cell* **5**: 132-139.

Onyango,J.D., Burri,C., and Brun,R. (2000) An automated biological assay to determine levels of the trypanocidal drug melarsoprol in biological fluids. *Acta Trop* **74**: 95-100.

Oonuma,T., Morimatsu,M., Ochiai,K., and Syuto,B. (2003) Properties of the tumor suppressor gene brca2 in the cat. *J Vet Med Sci* **65**: 1123-1126.

Overath,P., Steverding,D., Chaudhri,M., Stierhof,Y.D., and Ziegelbauer,K. (1994) Structure and function of GPI-anchored surface proteins of Trypanosoma brucei. *Braz J Med Biol Res* **27**: 343-347.

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

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

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.

Pastink,A., Eeken,J.C.J., and Lohman,P.H.M. (2001) Genomic integrity and the repair of double-strand DNA breaks. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **480-481**: 37-50.

Patel,K.J., Yu,V.P., Lee,H., Corcoran,A., Thistlethwaite,F.C., Evans,M.J. *et al.* (1998) Involvement of Brca2 in DNA repair. *Mol Cell* **1**: 347-357.

Pays,E. (1985) Antigenic variability of African trypanosomes. *Ann Inst Pasteur Immunol* **136D**: 163-166.

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

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

Pays,E., Lips,S., Nolan,D., Vanhamme,L., and Perez-Morga,D. (2001) The VSG expression sites of Trypanosoma brucei: multipurpose tools for the adaptation of the parasite to mammalian hosts. *Mol Biochem Parasitol* **114**: 1-16.

Pays,E., Salmon,D., Morrison,L.J., Marcello,L., and Barry,J.D. (2007) Antigenic variation in *Trypanosoma brucei.* In Trypanosomes. After the genome. Barry,J.D., McCulloch,R., Mottram,J.C., and Acosta-Serrano,A. (eds). Horizon Bioscience, pp. 339-372.

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

Pays,E., Vanhamme,L., and 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.

Petalcorin,M.I., Galkin,V.E., Yu,X., Egelman,E.H., and Boulton,S.J. (2007) Stabilization of RAD-51-DNA filaments via an interaction domain in Caenorhabditis elegans BRCA2. *Proc Natl Acad Sci U S A* **104**: 8299-8304.

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

Pierce,A.J., Johnson,R.D., Thompson,L.H., and Jasin,M. (1999) XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev* **13**: 2633-2638.

Pittman,D.L., and Schimenti,J.C. (2000) Midgestation lethality in mice deficient for the RecA-related gene, Rad51d/Rad51l3. *Genesis* **26**: 167-173.

Powell,S.N., Willers,H., and Xia,F. (2002) BRCA2 keeps Rad51 in line. High-fidelity homologous recombination prevents breast and ovarian cancer? *Mol Cell* **10**: 1262-1263.

Proudfoot,C., and McCulloch,R. (2006) Trypanosoma brucei DMC1 does not act in DNA recombination, repair or antigenic variation in bloodstream stage cells. *Mol Biochem Parasitol* **145**: 245-253.

Proudfoot,C., and McCulloch,R. (2005) Distinct roles for two RAD51-related genes in Trypanosoma brucei antigenic variation. *Nucleic Acids Res* **33**: 6906-6919.

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.

Raz,B., Iten,M., Grether-Buhler,Y., Kaminsky,R., and Brun,R. (1997) The Alamar Blue assay to determine drug sensitivity of African trypanosomes (T.b. rhodesiense and T.b. gambiense) in vitro. *Acta Trop* **68**: 139-147.

Regis-da-Silva,C.G., Freitas,J.M., Passos-Silva,D.G., Furtado,C., ugusto-Pinto,L., Pereira,M.T. *et al.* (2006) Characterization of the Trypanosoma cruzi Rad51 gene and its role in recombination events associated with the parasite resistance to ionizing radiation. *Molecular and Biochemical Parasitology* **149**: 191-200.

Reid,S., Schindler,D., Hanenberg,H., Barker,K., Hanks,S., Kalb,R. *et al.* (2007) Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat Genet* **39**: 162-164.

Renzette,N., Gumlaw,N., Nordman,J.T., Krieger,M., Yeh,S.P., Long,E. *et al.* (2005) Localization of RecA in Escherichia coli K-12 using RecA-GFP. *Mol Microbiol* **57**: 1074-1085.

Ricchetti,M., Dujon,B., and Fairhead,C. (2003) Distance from the chromosome end determines the efficiency of double strand break repair in subtelomeres of haploid yeast. *J Mol Biol* **328**: 847-862.

Richards,F.F., Rosen,N.L., Onodera,M., Bogucki,M.S., Neve,R.L., Hotez,P. *et al.* (1981) Antigenic variation and the surface glycoproteins of Trypanosoma congolense. *Fed Proc* **40**: 1434-1439.

Rigaut,G., Shevchenko,A., Rutz,B., Wilm,M., Mann,M., and Seraphin,B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**: 1030-1032.

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.

Rivas,S., Romeis,T., and Jones,J.D. (2002) The Cf-9 disease resistance protein is present in an approximately 420-kilodalton heteromultimeric membrane-associated complex at one molecule per complex. *Plant Cell* **14**: 689-702.

Robbins,J., Dilworth,S.M., Laskey,R.A., and Dingwall,C. (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**: 615-623.

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

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

Roth,C., Jacquemot,C., Giroud,C., Bringaud,F., Eisen,H., and Baltz,T. (1991) Antigenic variation in Trypanosoma equiperdum. *Res Microbiol* **142**: 725-730.

Rouse,J., and Jackson,S.P. (2002) Interfaces between the detection, signaling, and repair of DNA damage. *Science* **297**: 547-551.

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., 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.

Rutter W Jr, Valenzuela P, Bell G E, Holland M, Hager G L, Degennaro L J, and Bishop R J (1976) The role of DNA-dependent RNA polymerase in transcriptive specificity. In The Organization and Expression of the Eukaryotic Genome. Bradbury,E.M., and Javeherian K (eds). New York: Academic Press, pp. 279-293.

Saeki,H., Siaud,N., Christ,N., Wiegant,W.W., van Buul,P.P., Han,M. *et al.* (2006) Suppression of the DNA repair defects of BRCA2-deficient cells with heterologous protein fusions. *Proc Natl Acad Sci U S A* **103**: 8768-8773.

Salmon,D., Paturiaux-Hanocq,F., Poelvoorde,P., Vanhamme,L., and Pays,E. (2005) Trypanosoma brucei: growth differences in different mammalian sera are not due to the species-specificity of transferrin. *Exp Parasitol* **109**: 188-194.

Sambrook,J., Fritsch,E.F., and Maniatis,T. (1989) Molecular cloning: a laboratory manual N.Y.: Cold Spring Harbor.

Sauvageau,S., Ploquin,M., and Masson,J.Y. (2004) Exploring the multiple facets of the meiotic recombinase Dmc1. *Bioessays* **26**: 1151-1155.

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

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

Schild,D., Lio,Y.C., Collins,D.W., Tsomondo,T., and Chen,D.J. (2000) Evidence for simultaneous protein interactions between human Rad51 paralogs. *J Biol Chem* **275**: 16443-16449.

Schimanski,B., Nguyen,T.N., and Gunzl,A. (2005) Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot Cell* **4**: 1942-1950.

Schweizer,J., Tait,A., and Jenni,L. (1988) The timing and frequency of hybrid formation in African trypanosomes during cyclical transmission. *Parasitol Res* **75**: 98-101.

Scully,R., Chen,J., Plug,A., Xiao,Y., Weaver,D., Feunteun,J. *et al.* (1997) Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88**: 265-275.

Sechman,E.V., Rohrer,M.S., and Seifert,H.S. (2005) A genetic screen identifies genes and sites involved in pilin antigenic variation in Neisseria gonorrhoeae. *Mol Microbiol* **57**: 468-483.

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

Sehorn,M.G., Sigurdsson,S., Bussen,W., Unger,V.M., and Sung,P. (2004) Human meiotic recombinase Dmc1 promotes ATP-dependent homologous DNA strand exchange. *Nature* **429**: 433-437.

Seigneur,M., Bidnenko,V., Ehrlich,S.D., and Michel,B. (1998) RuvAB acts at arrested replication forks. *Cell* **95**: 419-430.

Seitz,E.M., Brockman,J.P., Sandler,S.J., Clark,A.J., and Kowalczykowski,S.C. (1998) RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev* **12**: 1248-1253.

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

Shapiro,S.Z., Naessens,J., Liesegang,B., Moloo,S.K., and Magondu,J. (1984) Analysis by flow cytometry of DNA synthesis during the life cycle of African trypanosomes. *Acta Trop* **41**: 313-323.

Sharan,S.K., and Bradley,A. (1997) Murine Brca2: sequence, map position, and expression pattern. *Genomics* **40**: 234-241.

Sharan,S.K., Morimatsu,M., Albrecht,U., Lim,D.S., Regel,E., Dinh,C. *et al.* (1997) Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature* **386**: 804-810.

Shea,C., Glass,D.J., Parangi,S., and Van der Ploeg,L.H. (1986) Variant surface glycoprotein gene expression site switches in Trypanosoma brucei. *J Biol Chem* **261**: 6056-6063.

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.

Sheader,K., te,V.D., and Rudenko,G. (2004) Bloodstream form-specific up-regulation of silent vsg expression sites and procyclin in Trypanosoma brucei after inhibition of DNA synthesis or DNA damage. *J Biol Chem* **279**: 13363-13374.

Shen,Z., Pardington-Purtymun,P.E., Comeaux,J.C., Moyzis,R.K., and Chen,D.J. (1996) Associations of UBE2I with RAD52, UBL1, p53, and RAD51 proteins in a yeast two-hybrid system. *Genomics* **37**: 183-186.

Shevchenko,A., Jensen,O.N., Podtelejnikov,A.V., Sagliocco,F., Wilm,M., Vorm,O. *et al.* (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc Natl Acad Sci U S A* **93**: 14440-14445.

Shin,D.S., Pellegrini,L., Daniels,D.S., Yelent,B., Craig,L., Bates,D. *et al.* (2003) Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *EMBO J* **22**: 4566-4576.

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.

Shivji,M.K., Davies,O.R., Savill,J.M., Bates,D.L., Pellegrini,L., and Venkitaraman,A.R. (2006) A region of human BRCA2 containing multiple BRC repeats promotes RAD51-mediated strand exchange. *Nucleic Acids Res* **34**: 4000-4011.

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.

Shu,Z., Smith,S., Wang,L., Rice,M.C., and Kmiec,E.B. (1999) Disruption of muREC2/RAD51L1 in mice results in early embryonic lethality which can Be partially rescued in a p53(-/-) background. *Mol Cell Biol* **19**: 8686-8693.

Siaud,N., Dray,E., Gy,I., Gerard,E., Takvorian,N., and Doutriaux,M.P. (2004) Brca2 is involved in meiosis in Arabidopsis thaliana as suggested by its interaction with Dmc1. *EMBO J* **23**: 1392-1401.

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

Smith,A.B., Esko,J.D., and Hajduk,S.L. (1995) Killing of trypanosomes by the human haptoglobin-related protein. *Science* **268**: 284-286.

Smith,A.B., and Hajduk,S.L. (1995) Identification of haptoglobin as a natural inhibitor of trypanocidal activity in human serum. *Proc Natl Acad Sci U S A* **92**: 10262-10266.

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

Smogorzewska,A., Matsuoka,S., Vinciguerra,P., McDonald,E.R., III, Hurov,K.E., Luo,J. *et al.* (2007) Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell* **129**: 289-301.

Sogin,M.L., Elwood,H.J., and Gunderson,J.H. (1986) Evolutionary diversity of eukaryotic small-subunit rRNA genes. *Proc Natl Acad Sci U S A* **83**: 1383-1387.

Sogin,M.L., Gunderson,J.H., Elwood,H.J., Alonso,R.A., and Peattie,D.A. (1989) Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from Giardia lamblia. *Science* **243**: 75-77.

Sonoda,E., Sasaki,M.S., Buerstedde,J.M., Bezzubova,O., Shinohara,A., Ogawa,H. *et al.* (1998) Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death. *EMBO J* **17**: 598-608.

Sonoda,E., Takata,M., Yamashita,Y.M., Morrison,C., and Takeda,S. (2001) Homologous DNA recombination in vertebrate cells. *Proc Natl Acad Sci U S A* **98**: 8388-8394.

Stark,J.M., Pierce,A.J., Oh,J., Pastink,A., and Jasin,M. (2004) Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol Cell Biol* **24**: 9305-9316.

Sternberg,J.M. (2004) Human African trypanosomiasis: clinical presentation and immune response. *Parasite Immunol* **26**: 469-476.

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

Stracker,T.H., Theunissen,J.W., Morales,M., and Petrini,J.H. (2004) The Mre11 complex and the metabolism of chromosome breaks: the importance of communicating and holding things together. *DNA Repair (Amst)* **3**: 845-854.

Sturzbecher,H.W., Donzelmann,B., Henning,W., Knippschild,U., and Buchhop,S. (1996) p53 is linked directly to homologous recombination processes via RAD51/RecA protein interaction. *EMBO J* **15**: 1992-2002.

Sugawara,N., Ira,G., and Haber,J.E. (2000) DNA length dependence of the single-strand annealing pathway and the role of Saccharomyces cerevisiae RAD59 in double-strand break repair. *Mol Cell Biol* **20**: 5300-5309.

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.

Sun,H., Treco,D., and Szostak,J.W. (1991) Extensive 3'-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the ARG4 recombination initiation site. *Cell* **64**: 1155-1161.

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 Klein,H. (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* **7**: 739-750.

Sung,P., Krejci,L., Van Komen,S., and Sehorn,M.G. (2003) Rad51 recombinase and recombination mediators. *J Biol Chem* **278**: 42729-42732.

Swaffield,J.C., Melcher,K., and Johnston,S.A. (1995) A highly conserved ATPase protein as a mediator between acidic activation domains and the TATA-binding protein. *Nature* **374**: 88-91.

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.

Szabo,C., Masiello,A., Ryan,J.F., and Brody,L.C. (2000) The breast cancer information core: database design, structure, and scope. *Hum Mutat* **16**: 123-131.

Tait,A., and Turner,C.M. (1990) Genetic exchange in Trypanosoma brucei. *Parasitol Today* **6**: 70-75.

Takanami,T., Mori,A., Takahashi,H., and Higashitani,A. (2000) Hyper-resistance of meiotic cells to radiation due to a strong expression of a single recA-like gene in Caenorhabditis elegans. *Nucleic Acids Res* **28**: 4232-4236.

Takata,M., Sasaki,M.S., Sonoda,E., Fukushima,T., Morrison,C., Albala,J.S. *et al.* (2000) The Rad51 paralog Rad51B promotes homologous recombinational repair. *Mol Cell Biol* **20**: 6476-6482.

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

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

Takata,M., Tachiiri,S., Fujimori,A., Thompson,L.H., Miki,Y., Hiraoka,M. *et al.* (2002) Conserved domains in the chicken homologue of BRCA2. *Oncogene* **21**: 1130-1134.

Takata,M., Yamamoto,K., Matsushita,N., Kitao,H., Hirano,S., and Ishiai,M. (2006) The Fanconi anemia pathway promotes homologous recombination repair in DT40 cell line. *Subcell Biochem* **40**: 295-311.

Tamura,T., Konishi,Y., Makino,Y., and Mikoshiba,K. (1996) Mechanisms of transcriptional regulation and neural gene expression. *Neurochem Int* **29**: 573-581.

Tan,T.L., Essers,J., Citterio,E., Swagemakers,S.M., de Wit,J., Benson,F.E. *et al.* (1999) Mouse Rad54 affects DNA conformation and DNA-damage-induced Rad51 foci formation. *Curr Biol* **9**: 325-328.

Taniguchi,T., and D'Andrea,A.D. (2006) Molecular pathogenesis of Fanconi anemia: recent progress. *Blood* **107**: 4223-4233.

Tarsounas,M., Davies,A.A., and West,S.C. (2004) 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.

Tashiro,S., Kotomura,N., Shinohara,A., Tanaka,K., Ueda,K., and Kamada,N. (1996) S phase specific formation of the human Rad51 protein nuclear foci in lymphocytes. *Oncogene* **12**: 2165-2170.

Tavtigian,S.V., Simard,J., Rommens,J., Couch,F., Shattuck-Eidens,D., Neuhausen,S. *et al.* (1996) The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat Genet* **12**: 333-337.

Tetley,L., Turner,C.M., Barry,J.D., Crowe,J.S., and Vickerman,K. (1987) Onset of expression of the variant surface glycoproteins of Trypanosoma brucei in the tsetse fly studied using immunoelectron microscopy. *J Cell Sci* **87 (Pt 2)**: 363-372.

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

Thon,G., Baltz,T., Giroud,C., and Eisen,H. (1990) Trypanosome variable surface glycoproteins: composite genes and order of expression. *Genes Dev* **4**: 1374-1383.

Thorslund,T., Esashi,F., and West,S.C. (2007) Interactions between human BRCA2 protein and the meiosis-specific recombinase DMC1. *EMBO J.*

Thorslund,T., and West,S.C. (2007) BRCA2: a universal recombinase regulator. *Oncogene* **26**: 7720-7730.

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., Yuan,S.S., Lee,E.Y., and Sung,P. (1998) Nuclease activities in a complex of human recombination and DNA repair factors Rad50, Mre11, and p95. *J Biol Chem* **273**: 21447-21450.

Turner,C.M. (1997) The rate of antigenic variation in fly-transmitted and syringe-passaged infections of Trypanosoma brucei. *FEMS Microbiol Lett* **153**: 227-231.

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.

Turner,C.M., McLellan,S., Lindergard,L.A., Bisoni,L., Tait,A., and MacLeod,A. (2004) Human infectivity trait in Trypanosoma brucei: stability, heritability and relationship to sra expression. *Parasitology* **129**: 445-454.

Tutt,A., Bertwistle,D., Valentine,J., Gabriel,A., Swift,S., Ross,G. *et al.* (2001) Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. *EMBO J* **20**: 4704-4716.

Ui,A., Seki,M., Ogiwara,H., Onodera,R., Fukushige,S.i., Onoda,F., and Enomoto,T. (2005) The ability of Sgs1 to interact with DNA topoisomerase III is essential for damage-induced recombination. *DNA Repair* **4**: 191-201.

Ullu,E., Matthews,K.R., and Tschudi,C. (1993) Temporal order of RNA-processing reactions in trypanosomes: rapid trans splicing precedes polyadenylation of newly synthesized tubulin transcripts. *Mol Cell Biol* **13**: 720-725.

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 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 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.

Vandenberg,C.J., Gergely,F., Ong,C.Y., Pace,P., Mallery,D.L., Hiom,K., and Patel,K.J. (2003) BRCA1-independent ubiquitination of FANCD2. *Mol Cell* **12**: 247-254.

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

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 (Pt 21)**: 2661-2671.

Venkitaraman,A.R. (2002) Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* **108**: 171-182.

Vickerman,K. (1985) Developmental cycles and biology of pathogenic trypanosomes. *Br Med Bull* **41**: 105-114.

Vickerman,K. (1978) Antigenic variation in trypanosomes. *Nature* **273**: 613-617.

Voloshin,O.N., Wang,L., and Camerini-Otero,R.D. (1996) Homologous DNA pairing promoted by a 20-amino acid peptide derived from RecA. *Science* **272**: 868-872.

Walgraffe,D., Devaux,S., Lecordier,L., Dierick,J.F., Dieu,M., Van den,A.J. *et al.* (2005) Characterization of subunits of the RNA polymerase I complex in Trypanosoma brucei. *Mol Biochem Parasitol* **139**: 249-260.

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,X., Andreassen,P.R., and D'Andrea,A.D. (2004) Functional Interaction of Monoubiquitinated FANCD2 and BRCA2/FANCD1 in Chromatin. *Mol Cell Biol* **24**: 5850-5862.

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

West,S.C. (2003) Molecular views of recombination proteins and their control. *Nat Rev Mol Cell Biol* **4**: 435-445.

West,S.C., Cassuto,E., and Howard-Flanders,P. (1981) Homologous pairing can occur before DNA strand separation in general genetic recombination. *Nature* **290**: 29-33.

White,C.I., and Haber,J.E. (1990) Intermediates of recombination during mating type switching in Saccharomyces cerevisiae. *EMBO J* **9**: 663-673.

Wickstead,B., Ersfeld,K., and Gull,K. (2003a) Repetitive elements in genomes of parasitic protozoa. *Microbiol Mol Biol Rev* **67**: 360-75, table.

Wickstead,B., Ersfeld,K., and Gull,K. (2004) The small chromosomes of Trypanosoma brucei involved in antigenic variation are constructed around repetitive palindromes. *Genome Res* **14**: 1014-1024.

Wickstead,B., Ersfeld,K., and Gull,K. (2003b) The frequency of gene targeting in Trypanosoma brucei is independent of target site copy number. *Nucl Acids Res* **31**: 3993.

Wijers,D.J., and Willet,K.C. (1960) Factors that may influence the infection rate of Glossina palpalis with Trypanosoma gambiense. II. The number and morphology of the trypano-somes present in the blood of the host at the time of the infected feed. *Ann Trop Med Parasitol* **54**: 341-350.

Wilson,J.B., Yamamoto,K., Marriott,A.S., Hussain,S., Sung,P., Hoatlin,M.E. *et al.* (2008) FANCG promotes formation of a newly identified protein complex containing BRCA2, FANCD2 and XRCC3. *Oncogene*.

Wilson,J.H., and Elledge,S.J. (2002) Cancer. BRCA2 enters the fray. *Science* **297**: 1822-1823.

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.

Wong,A.K., Pero,R., Ormonde,P.A., Tavtigian,S.V., and Bartel,P.L. (1997) RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2. *J Biol Chem* **272**: 31941-31944.

Wong,J.M., Ionescu,D., and Ingles,C.J. (2003) Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2. *Oncogene* **22**: 28-33.

Wong,Z., Wilson,V., Patel,I., Povey,S., and Jeffreys,A.J. (1987) Characterization of a panel of highly variable minisatellites cloned from human DNA. *Ann Hum Genet* **51**: 269-288.

Wooster,R., Bignell,G., Lancaster,J., Swift,S., Seal,S., Mangion,J. *et al.* (1995) Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**: 789-792.

Wooster,R., Neuhausen,S.L., Mangion,J., Quirk,Y., Ford,D., Collins,N. *et al.* (1994) Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science* **265**: 2088-2090.

Xia,B., Dorsman,J.C., Ameziane,N., de Vries,Y., Rooimans,M.A., Sheng,Q. *et al.* (2007) Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat Genet* **39**: 159-161.

Xia,B., Sheng,Q., Nakanishi,K., Ohashi,A., Wu,J., Christ,N. *et al.* (2006a) Control of BRCA2 Cellular and Clinical Functions by a Nuclear Partner, PALB2. *Mol Cell* **22**: 719-729.

Xia,B., Sheng,Q., Nakanishi,K., Ohashi,A., Wu,J., Christ,N. *et al.* (2006b) Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell* **22**: 719-729.

Xia,F., Taghian,D.G., DeFrank,J.S., Zeng,Z.C., Willers,H., Iliakis,G., and Powell,S.N. (2001) Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. *Proc Natl Acad Sci U S A* **98**: 8644-8649.

Xong,H.V., Vanhamme,L., Chamekh,M., Chimfwembe,C.E., Van den,A.J., Pays,A. *et al.* (1998) A VSG expression site-associated gene confers resistance to human serum in Trypanosoma rhodesiense. *Cell* **95**: 839-846.

Xu,Z., Fulop,Z., Zhong,Y., Evinger,A.J., III, Zan,H., and Casali,P. (2005) DNA lesions and repair in immunoglobulin class switch recombination and somatic hypermutation. *Ann N Y Acad Sci* **1050**: 146-162.

Yang,H., Jeffrey,P.D., Miller,J., Kinnucan,E., Sun,Y., Thoma,N.H. *et al.* (2002) BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science* **297**: 1837-1848.

Yang,W., and Steitz,T.A. (1995) Crystal structure of the site-specific recombinase gamma delta resolvase complexed with a 34 bp cleavage site. *Cell* **82**: 193-207.

Yang,Y.G., Herceg,Z., Nakanishi,K., Demuth,I., Piccoli,C., Michelon,J. *et al.* (2005) The Fanconi anemia group A protein modulates homologous repair of DNA double-strand breaks in mammalian cells. *Carcinogenesis* **26**: 1731-1740.

Yu,D.S., Sonoda,E., Takeda,S., Huang,C.L., Pellegrini,L., Blundell,T.L., and Venkitaraman,A.R. (2003) Dynamic control of Rad51 recombinase by self-association and interaction with BRCA2. *Mol Cell* **12**: 1029-1041.

Yu,V.P., Koehler,M., Steinlein,C., Schmid,M., Hanakahi,L.A., van Gool,A.J. *et al.* (2000) Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation. *Genes Dev* **14**: 1400-1406.

Yuan,S.S., Lee,S.Y., Chen,G., Song,M., Tomlinson,G.E., and Lee,E.Y. (1999) BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo. *Cancer Res* **59**: 3547-3551.

Zhang,J.R., Hardham,J.M., Barbour,A.G., and Norris,S.J. (1997) Antigenic variation in Lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. *Cell* **89**: 275-285.

Zhang,Q.Y., DeRyckere,D., Lauer,P., and Koomey,M. (1992) Gene conversion in Neisseria gonorrhoeae: evidence for its role in pilus antigenic variation. *Proc Natl Acad Sci U S A* **89**: 5366-5370.

Zhou,Q., Kojic,M., Cao,Z., Lisby,M., Mazloum,N.A., and Holloman,W.K. (2007) Dss1 interaction with Brh2 as a regulatory mechanism for recombinational repair. *Mol Cell Biol* **27**: 2512-2526.

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

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