

***Trypanosoma brucei* homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases**

Rebecca L. Barnes and Richard McCulloch*

The Wellcome Centre for Molecular Parasitology, University of Glasgow, Glasgow Biomedical Research Centre, 120 University Place, Glasgow, G12 8TA, UK

Received March 14, 2007; Revised April 4, 2007; Accepted April 5, 2007

ABSTRACT

Homologous recombination functions universally in the maintenance of genome stability through the repair of DNA breaks and in ensuring the completion of replication. In some organisms, homologous recombination can perform more specific functions. One example of this is in antigenic variation, a widely conserved mechanism for the evasion of host immunity. *Trypanosoma brucei*, the causative agent of sleeping sickness in Africa, undergoes antigenic variation by periodic changes in its variant surface glycoprotein (VSG) coat. VSG switches involve the activation of VSG genes, from an enormous silent archive, by recombination into specialized expression sites. These reactions involve homologous recombination, though they are characterized by an unusually high rate of switching and by atypical substrate requirements. Here, we have examined the substrate parameters of *T. brucei* homologous recombination. We show, first, that the reaction is strictly dependent on substrate length and that it is impeded by base mismatches, features shared by homologous recombination in all organisms characterized. Second, we identify a pathway of homologous recombination that acts preferentially on short substrates and is impeded to a lesser extent by base mismatches and the mismatch repair machinery. Finally, we show that mismatches during *T. brucei* recombination may be repaired by short-patch mismatch repair.

INTRODUCTION

Homologous recombination is a universal process in living organisms. The central enzyme of this reaction appears to have been conserved in all kingdoms of life (1) and in viruses: RecA in bacteria, Rad51 in eukaryotes, RadA in archaea and UvsX in phage T4 (2). In each case, the enzyme functions by catalysing the exchange of single-stranded DNA into intact DNA duplex, generating homologous pairing and promoting recombination (3). Furthermore, each orthologous protein binds single-stranded DNA in the form of a nucleoprotein filament, which has a highly similar structure in *Escherichia coli* (4), *Saccharomyces cerevisiae* (5) and *Archeoglobus fulgidus* (6). The purpose of homologous recombination is to repair DNA double-strand breaks (3) and to restart stalled or collapsed replication forks (7). In eukaryotes, it is also used in meiotic recombination (8) and, in some circumstances, telomere maintenance (9). Beyond these generalized functions, homologous recombination has been co-opted into specific functions in a diverse set of organisms. One example is mating type switching in yeast, where homologous recombination is induced by site-directed DNA lesions (10). In many pathogenic organisms, including bacteria, fungi and protists, homologous recombination can play a similarly specialized role in host immune evasion (11–13).

One way pathogens evade immunity is by antigenic variation, the periodic switching of surface antigens. In *Trypanosoma brucei*, a protistan parasite of mammals in Africa, antigenic variation involves switches in the variant surface glycoprotein (VSG) coat. The success of this strategy relies upon a *T. brucei* cell expressing a single VSG coat type at any one time, and the ability to switch to an antigenically distinct version, selected from an enormous archive of >1000 silent VSG genes (14); for recent reviews see (15–19). Singular VSG expression has involved the evolution of telomeric VSG transcription units, termed expression sites (ES), and transcriptional

*To whom correspondence should be addressed. Tel: 0044 141 330 5946; Fax: 0044 141 330 5422; Email: rmc9z@udcf.gla.ac.uk

control mechanisms that act upon them. *VSG* switching is dependent on recombination of the silent *VSGs* into the ES, and a number of such reactions have been described. The most commonly observed, at least early in *T. brucei* infections, are gene conversions that generate a copy of a silent *VSG* and transfer it to the ES. Such gene conversions encompass the *VSG* ORF and extend to regions of homology upstream and downstream. *VSG*-associated 70-bp repeats frequently demarcate the upstream conversion boundary (20,21), whilst the reaction can extend downstream to short blocks of homology in the *VSG* 3' ends (22) or to the telomeric repeats (23). Crossover exchanges between chromosome ends, termed reciprocal *VSG* switches, are also seen (24). Finally, segmental gene conversions are found where multiple *VSG* pseudogenes are recombined together to form novel, mosaic *VSGs* (25,26). These have been considered rare events, found late in infections. However, sequencing the *T. brucei* genome has revealed that *VSG* pseudogenes represent the substantial majority of the *VSG* archive (14), arguing that this is likely to be a significant mechanism of *VSG* switching (26,27).

Growing evidence suggests that *VSG* switching is closely linked to homologous recombination, since mutations in several key factors of homology-directed strand exchange, including RAD51 (28), a RAD51-related protein called RAD51-3 (29) and BRCA2 (C.Hartley and R.McCulloch, unpublished data), impair the immune evasion process. However, *T. brucei* *VSG* switching presents several unusual characteristics. First, the reaction can occur at very high rates (up to 1×10^{-2} switches/cell/generation)(30,31), significantly more frequent than the rates of general homologous recombination, which are more typical of random mutation (32). Second, recombination of *VSGs* frequently relies on flanking sequences, such as the 70 bp repeats, that are rather short and share limited homology (33,34), despite the fact that the *T. brucei* mismatch repair (MMR) machinery regulates homologous recombination to favour well-matched sequences, and estimates suggest that around 100 bp of homology are needed for efficient RAD51-mediated recombination (35). Finally, there appears to be a hierarchy in the substrates that are used by general homologous recombination in other organisms: sister chromatids appear to be the favoured substrate in both yeast (36) and mammals (32), while allelic sequences on chromosome homologues and homologous sequences at ectopic locations are progressively disfavoured (37,38). Recombinational activation of allelic *VSG* sequence on the sister chromatid would not result in a coat switch, compelling the reaction to search for silent *VSGs* throughout the genome. It has been argued that this is why the *VSG* archive is predominantly sub-telomeric, as these locations appear to escape such substrate constraints (39). Beyond that, it is unclear how the other characteristics of *VSG* switching are accommodated by homologous recombination. One study has suggested that the *T. brucei* homologous recombination is distinct from that of *S. cerevisiae* and its relatively close cousin *Leishmania*, in that the reaction can act on short stretches of homology and may have an elevated rate of strand exchange (40),

perhaps indicating modifications of the recombination machinery. Another study has suggested that it may be necessary for *T. brucei* to suppress MMR to allow homologous recombination to act during *VSG* switching (33). Here, we have sought to examine the parameters of homologous recombination in *T. brucei* in order to address these questions further. To do this we have used a transformation assay that allows us to measure the efficiency of *T. brucei* recombination, and to assess the pathways that operate.

In previous work, we characterized the *T. brucei* MMR machinery (41), which plays a critical role in maintaining genome stability and is conserved throughout evolution. The function of MMR is to recognize and excise base mismatches, which arise through replication errors, by chemical damage or during recombination between incompletely sequence-matched DNA molecules (42,43). Eukaryotic MMR is catalysed by homologues of bacterial MutS and MutL proteins, though the machinery has been elaborated, since most eukaryotes encode 3-7 MutS-related proteins (termed MSH 1-7) and 2-4 MutL-related proteins (MLH1-3 and PMS1-2). Mutation of either MSH2 or MLH1 in *T. brucei* demonstrates that MMR functions in correcting errors in the nuclear genome (41). Furthermore, MSH2, and by implication MMR, plays a role in constraining *T. brucei* homologous recombination to occur between well-matched sequences (35). A similar anti-recombination role for MMR has long been appreciated in other organisms, where it contributes to the suppression of excessive genome rearrangements and to speciation (44), though how it occurs remains to be determined fully. In this work, we extend the above analysis to show that *T. brucei* homologous recombination is strictly dependent on substrate length, in keeping with findings in other organisms. In addition, by comparing the recombination of diverged substrates of different lengths in wild-type cells and in MSH2 mutants, we have identified an MMR-independent homologous recombination pathway and find that a short-patch MMR pathway can correct base mismatches during *T. brucei* recombination.

MATERIALS AND METHODS

T. brucei strains, growth and transformation

Trypanosoma brucei bloodstream form cells were used and grown at 37°C in HMI-9 medium (45). The cells were of strain HTUB (35), which was derived from the MITat1.2 cell line by insertion of the hygromycin phosphotransferase (*HYG*) ORF into the tubulin array; *MSH2* heterozygous (+/-) and homozygous (-/-) mutants in this strain have been described before (35). Transformations to assay recombination efficiency were carried out by electroporation with 3 µg of DNA that had been PCR-amplified using Herculase (Stratagene) high-fidelity DNA polymerase (see below). Electroporation conditions were a single pulse at 1.4 kV, 25 µF using a Bio-Rad Gene Pulser II, and at least three transformations were performed for most constructs. For the transformations, *T. brucei* cells were grown maximally to 3×10^6 cells·ml⁻¹,

and minimally to 1.5×10^6 cells·ml⁻¹, harvested by centrifugation at 600g for 10 min at room temperature, and then re-suspended in Zimmerman post-fusion medium supplemented with 1% D-Glucose (46) to a concentration of 5×10^7 cells·ml⁻¹. A total of 2.5×10^7 cells were used per transformation and were allowed to recover following electroporation by growth in 10 ml of non-selective medium for 18 h before antibiotic selection. For this, the cells were harvested as before and then resuspended in HMI-9 containing 2.5 µg·ml⁻¹ phleomycin (Cayla) and spread in 1.0 ml aliquots over a 24-well dish. Between 2 and 10×10^6 cells were plated out in this way, depending on the construct being used (see Supplementary Figure 1). Transformation rates were measured by the number of wells containing phleomycin-resistant cells after 8–14 days growth. We have assumed that the cell population in each well is clonal, arising from a single transformant, though this underestimates the true number of transformants. However, calculating the likely correct number of transformants in each plate using the number of negative wells, and assuming a Poisson distribution of clones (40), does not alter the nature of the relationship between transformation rate and substrate length or homology (data not shown). In addition, such analysis does not alter

significantly the calculated minimum efficient processing segment, which was determined as described by Shen and Huang (1986).

Generation of *HYG* targeting constructs

For each transformation construct, 24 PCRs were performed, the products pooled and purified using a Qiagen PCR purification kit according to manufacturer's instructions; 1 QIAquick column was used per six PCRs, and the pooled PCR products resuspended in distilled H₂O. Each of the different sized PCR products were PCR-amplified either from p*HYG*wt::BLE, p*HYG*05::BLE or p*HYG*11::BLE (35), generating constructs with 0, 5 or 11% base mismatches relative to the HTUB (*HYG*wt) sequence. The oligonucleotide primers used to generate each construct were size-purified prior to the PCR and are named in Supplementary Figure 1 (sequences can be provided on request); the precise length of each targeting flank in the different PCR products, as well as the number of mismatched bases relative to *HYG*wt, is also detailed. For reasons that are unclear, a 125 bp construct with 11% sequence divergence could not be PCR-amplified and was omitted from the study. All the purified PCR products were examined by agarose gel

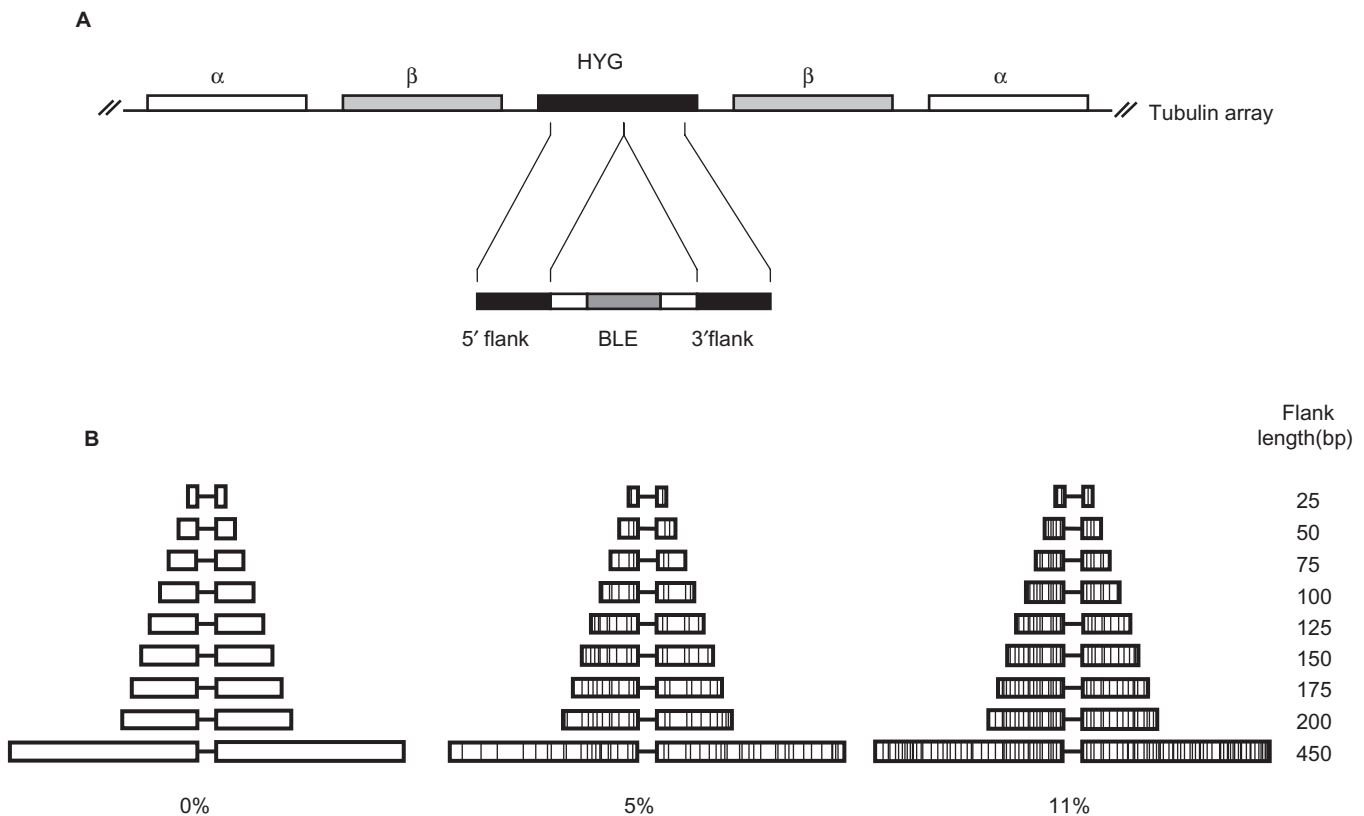


Figure 1. Assaying the length and sequence homology requirements of *T. brucei* recombination. (A) *HYG* (black box), encoding hygromycin resistance, was integrated into the tubulin array of bloodstream stage *T. brucei* cells, replacing an α -tubulin ORF (white box). *HYG*-transformed cells were then transformed with constructs containing a bleomycin resistance gene (*BLE*; dark grey box) and recombination flanks which target integration to *HYG*. (B) The *HYG* recombination flanks of the different constructs used in this study are diagrammed. In each, the 5' and 3' flanks are shown as boxes of decreasing size, depicting the different lengths (indicated) of the homology with *HYG*. The constructs were of three classes: the flanks had 100% sequence homology with *HYG* (indicated as 0% divergence), or had base mismatches (depicted by vertical lines) that reduced the homology to 95% (5% divergence) or 89% (11% divergence).

electrophoresis to ensure the lack of visible contaminating DNA molecules, and their concentration was determined spectrophotometrically.

Analysis of transformants

The hygromycin sensitivity or resistance of the *T. brucei* cells was determined by replica passaging 100 μ l of the phleomycin-resistant transformants into 1.5 ml of either non-selective HMI-9, or media containing 5 μ g.ml⁻¹ hygromycin (Roche). Growth was assessed microscopically 48 h later. For Southern analysis of the transformants, a 15 ml culture was grown to a density of $\sim 4 \times 10^6$ cells.ml⁻¹, harvested by centrifugation as before, resuspended in 500 μ l of 1 mM EDTA, 100 mM NaCl, 50 mM Tris-HCl pH 8.0 and lysed overnight at 37°C following the addition of SDS to 1% and proteinase K to 100 μ g.ml⁻¹. DNA was recovered by phenol/chloroform extraction and ethanol precipitation, and then resuspended in distilled H₂O. The genomic DNA samples were digested by restriction enzymes as described by the manufacturers and separated by electrophoresis, typically at ~ 30 V overnight, on 0.8% agarose gels (Seakem LE agarose, BioWhittaker Molecular Applications) made with 1 \times TAE buffer (40 mM Tris, 19 mM acetic acid, 1 mM EDTA) containing 0.2 μ g.ml⁻¹ ethidium bromide (Sigma). DNA was blotted by capillary transfer onto hybrid-XL membrane (Amersham), probed with α^{32} -P labelled DNA generated by random priming and washed to 0.2 \times SSC, 0.1% SDS at 65°C. Separation of intact *T. brucei* chromosomes was carried out on a Bio-Rad CHEF-DRIII apparatus. For this, each agarose plug contained $\sim 4 \times 10^7$ *T. brucei* cells, which had been grown in HMI-9 to a density of $\sim 2 \times 10^6$ cells.ml⁻¹, centrifuged as before, washed by resuspending the pelleted cells in 10 ml PSG (1 \times PBS, 1% w/v glucose), re-centrifuged and then resuspended in PSG at a concentration of 1.6×10^6 cells.ml⁻¹. The cells were then warmed at 37°C for 1 min and an equal volume of 1.4% Microsieve low-melt agarose (Flowgen) in H₂O added and mixed. Disposable plug moulds (BioRad) were filled with ~ 50 μ l agarose and placed at 4°C for ~ 4 h to set. The agarose plugs were then removed from the moulds, incubated in NDS buffer (0.5 M EDTA, 1 mM Tris base and 34.1 mM lauroyl sarcosine) pH 9.0 containing 1 mg.ml⁻¹ proteinase K at 50°C for ~ 24 h, transferred into NDS buffer pH 8.0 containing 1 mg.ml⁻¹ proteinase K at 50°C for ~ 24 h, and finally transferred into NDS buffer pH 8.0 for storage at 4°C. For electrophoresis, the plugs were washed four times at room temperature in 1 ml of 1 \times TB(0.1)E (0.089 M Tris-borate pH 8.0), 2 mM EDTA) for 1 h each, then placed in the wells of a 1.2% agarose (Seakem LE, BioWhittaker Molecular Applications) gel. The gel was electrophoresed at 15°C at 2.5 V.cm⁻¹ for 144 h with a 1400–700 s switch time and visualized by staining with 0.5 μ g.ml⁻¹ ethidium bromide and UV illumination. Sequences of the integrated DNA constructs were determined by performing PCR amplifications using oligonucleotide primers corresponding to the first and last 20 nt of the *HYG* ORF and *Taq* DNA polymerase (ABgene). The resulting PCR products were purified and

sequenced using oligonucleotide primers that read upstream and downstream from the bleomycin resistance cassette common to each construct.

RESULTS

Assaying *T. brucei* homologous recombination by targeted gene replacement

Previously, we described an assay to examine homologous recombination efficiency in *T. brucei* (35); this is summarized in Figure 1A. The assay relies upon a hygromycin phosphotransferase ORF (*HYG*) integrated into the tubulin array (47) on chromosome 1, providing a unique site for recombination. Recombination efficiency is determined by measuring the transformation rate of linear constructs containing a bleomycin resistance cassette (*BLE*) flanked by sequences derived from *HYG*. The advantage of this approach is that a single, defined site for recombination is analysed from which a number of parameters can be varied (see below). In addition, using a foreign sequence as a target reduces the potential for recombination into related sequences elsewhere in the genome, an issue that has influenced other studies where endogenous *T. brucei* sequences are used both as a genomic target and recombination substrate; e.g. (40,46,48). Although transformation of any organism is likely to be affected by a number of factors, including DNA concentration, transformation conditions and growth of the cells, this appears to be a reliable measure of recombination in *T. brucei*. The same assay has demonstrated the importance of *T. brucei* MMR in controlling homologous recombination between sequences containing base mismatches (35)(see below). Moreover, related transformation assays have quantified the role in recombination of a number of *T. brucei* genes, including *RAD51* (28,46), *MRE11* (48), *DMC1* (49) and two *RAD51*-related genes (*RAD51-3* and *RAD51-5*) (29), and have looked at the importance of target copy number (40). Most likely, this approach succeeds because virtually all stable transformants in *T. brucei* integrate linear DNA by homologous recombination, rather than by end-joining processes, and the formation of extra-chromosomal episomes following such transformation has not been described (50).

Using this assay we have shown previously that *T. brucei* homologous recombination, acting on substrates with targeting flanks of 450 bp, becomes increasingly less efficient as the level of sequence homology between the flanks and the genomic *HYG* target decreases (35). One percent of sequence divergence resulted in a 2.8-fold reduction in transformation efficiency, and 11% divergence caused a near 100-fold reduction. By mutating the *T. brucei* gene encoding MSH2, we showed that transformation efficiencies of substrates with 2–11% sequence divergence increased by around 9-fold, indicating that MMR is an important regulator of homologous recombination. Nevertheless, it is not the sole factor that determines the success or failure of recombination on such substrates, as the same decline in transformation efficiency with increasing divergence was seen in both

MSH2^{-/-} and wild-type cells. In this study, we have examined two further features of *T. brucei* homologous recombination. First, we tested the relationship between substrate length and recombination efficiency. Second, we asked if MMR has the same influence on homologous recombination between diverged substrates when the length of homology becomes short. To do this, a series of constructs with targeting flanks varying in size from 25 to 200 bp were generated (Figure 1B). These were derived by high-fidelity PCR from the previously described constructs with 450 bp flanks (35) that are either perfectly homologous to the genomic *HYG* sequence (0% sequence divergence), or that contain base mismatches resulting in either 5 or 11% divergence. Linear DNA was prepared by PCR amplification, rather than by restriction digestion, because non-homologous overhangs at the DNA ends, while insignificant for an integration flank of 450 bp, could have larger effects on recombination efficiencies mediated by the shorter substrates. The transformation efficiency of each construct was determined in *msh2*^{-/-}, *MSH2*^{+/-} and wild-type bloodstream stage *T. brucei* strains (35,41). The results of this analysis are graphed in Figures 2 and 3, and summarized in Tables 1 and 2.

T. brucei homologous recombination displays a strict dependence on substrate length

Comparing the transformation efficiencies of constructs with 100% sequence identity with the genomic *HYG* target indicates that *T. brucei* homologous recombination efficiency is dependent on substrate length. Over the range 200–50 bp, a linear relationship was found between transformation efficiency and substrate length (Figure 2). The average transformation rate for the 50 bp construct in the MMR-proficient cells (MMR+; either wild type or *MSH2*^{+/-}) was 0.80 ± 0.45 transformants $\times 10^{-6}$ cells, corresponding to a near 13-fold reduction in efficiency compared with the 200 bp construct ($10.16 \pm 0.38 \times 10^{-6}$). Beyond this range the linear relationship broke down. It appears that around 50 bp of sequence homology represents a lower threshold for *T. brucei* recombination, at which point the reaction becomes extremely inefficient: the 25-bp substrate displayed an average transformation rate in MMR+ cells ($0.02 \pm 0.06 \times 10^{-6}$) that was around 500-fold lower than the 200 bp substrate and 40-fold lower than the 50 bp substrate. It is notable, however, that transformants can be generated with 25 bp flanks, and that these can integrate by homologous recombination (Figure 4)(46). By comparing the transformation rate of the 200 bp construct with the restriction-digested plasmid containing 450 bp of homologous flank used previously (35), it appears that the rate of transformation reaches a plateau around 200 bp. This may indicate that *T. brucei* recombination becomes no more efficient with substrates longer than 200 bp, though it is also possible that this is not a reflection of recombination but represents the maximum transformation efficiency achievable in bloodstream stage cells under these *in vitro* conditions.

Mutation of *MSH2*, resulting in *T. brucei* cells with impaired MMR (MMR-), had no discernible

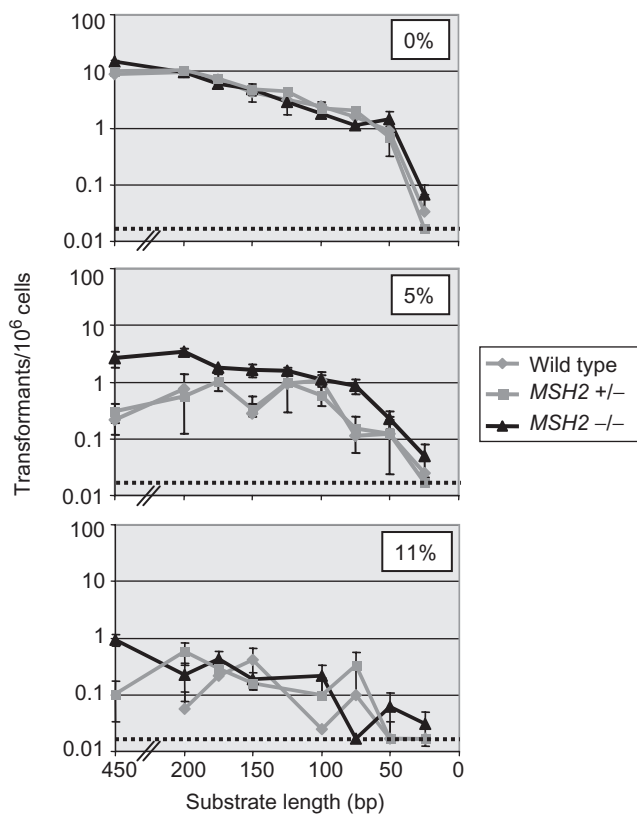


Figure 2. Transformation rate relative to substrate length and sequence homology. Plotted values represent the mean transformation efficiency for wild type, *MSH2* heterozygous (+/-) or *MSH2* homozygous (-/-) bloodstream stage *T. brucei* cells derived from at least three separate experiments; the error bars show standard error and the dotted lines depict the lowest detectable transformation rate in this assay. The upper graph depicts the relationship between transformation and substrate length with constructs that are 100% sequence matched (0% divergence), while the middle and lower graphs show the effect of increasing substrate divergence to 5% (95% sequence homology) and 11% (89% homology), respectively.

effect on recombination for any of 100% homologous constructs in the range 25–200 bp. This contrasts with the 450 bp parental construct, which showed a small (1.6-fold; Figure 3, Table 1) but significant increase in recombination in *msh2*^{-/-} cells relative to MMR+ cells (35). Elevation of recombination between similar-sized (350 bp), sequence-matched substrates is seen also in *S. cerevisiae* MMR mutants (51,52). The basis for this is unknown, but its absence on shorter substrates may indicate that MMR acts to suppress recombination on sequence-matched substrates only when they are of a significant length, perhaps because they are more prone to secondary structure during strand exchange.

Recombination of short substrates in *T. brucei* can occur by an MMR-independent pathway

On 450 bp substrates containing base mismatches relative to the genomic *HYG* sequence, increasing levels of sequence divergence had an exponentially deleterious effect on *T. brucei* homologous recombination (35).

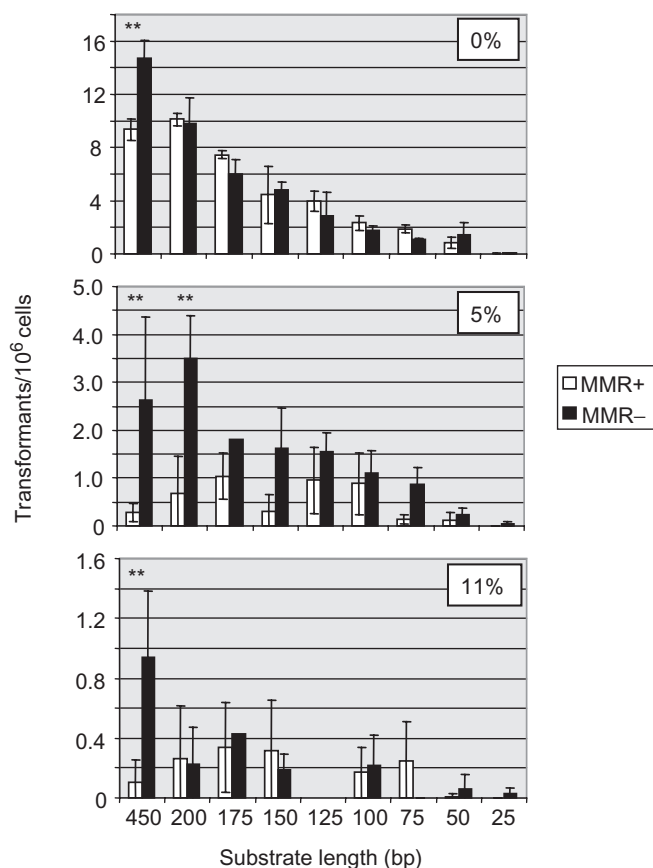


Figure 3. A comparison of transformation efficiency in mismatch repair-proficient and -deficient *T. brucei* cells. The values in the bar charts depict the average transformation rates of constructs with decreasing lengths of targeting flanks; error bars show standard error. In each case, white bars show the transformation rate of MMR-proficient cells (MMR+; wild type and *MSH2*+/-) and grey bars depict MMR-deficient cells (MMR-; *msh2*-/-). The top graph shows transformation rates of constructs with 100% sequence homology (0% sequence divergence) with the genomic *HYG* target, while the middle graph shows constructs that have 95% homology (5% divergence) and the lower graph shows constructs with 89% homology (11% divergence). Asterisks indicate a statistically significant difference between the transformation rate in the MMR+ and MMR- cells ($P < 0.05$; two sample *t*-test).

Furthermore, mutation of *MSH2* resulted in a mean 9-fold elevation in the rate of transformation when sequence divergence was $>2\%$ (Figure 3, Table 1). The transformation measurements in this study show that base mismatches and MMR regulation do not have a uniform effect on recombination over the range of substrate length analysed. This can be seen by examining the data in two ways. First, we examined the effect that mutating *MSH2* had on recombination of the 5 and 11% diverged substrates (Figures 2 and 3, Table 1). On the 5% diverged substrates, a statistically significant elevation in transformation efficiency in the MMR- cells relative to the MMR+ cells was observed only on the 450 and 200 bp substrates. In contrast, though the 5% diverged substrates smaller than 200 bp appeared to show a trend towards a slight increase in transformation in MMR- cells, this was not significant.

Table 1. The values shown are the fold 'increase' in the average transformation rate in *msh2*-/- cells relative to MMR-proficient cells (wild type and *MSH2*+/-) measured using constructs with different lengths of targeting flanks, and with 0, 5 or 11% sequence divergence (100, 95 and 89 homology, respectively) from the genomic *HYG* target; ND indicates not determined

Length (bp)	100%	95%	89%
450	1.57	9.29	9.4
200	0.97	5.3	4.4
175	0.81	1.73	1.68
150	1.06	5.22	0.64
125	0.78	1.63	ND
100	0.78	1.33	3.5
75	0.62	6.69	0.1
50	1.77	1.83	3
25	3.5	2.5	1.5

Table 2. The values show the fold 'decrease' in transformation rate of 5 and 11% diverged substrates relative to sequence-matched (0% divergence) substrates; data are shown from MMR-proficient (MMR+; wild type and *MSH2*+/-) and MMR-deficient (MMR-; *msh2*-/-) cells over a range of different substrate lengths

Length (bp)	MMR+		MMR-	
	5%	11%	5%	11%
450	33.4	93.4	5.6	15.6
200	15.4	203.2	2.8	66.8
175	7.1	29.6	3.3	14.2
150	14.6	16.2	3	26.8
125	3.9	ND	1.9	ND
100	2.8	38.9	1.6	8.6
75	13.8	8.5	1.3	55.5
50	6.7	40	6.5	23.7
25	1	1	1.4	2.3

Comparing the average transformation rates of these constructs confirms this (Table 1): the 450 and 200 bp substrates displayed 9.3- and 5.3-fold increases, respectively, in transformation rate in the MMR- cells relative to MMR+ cells, whereas averaging the data from all substrates below 200 bp revealed a mean 3.4-fold increase (range 1.3–6.7). On the 11% diverged substrates, impairment of MMR caused no significant elevation in transformation efficiency relative to the MMR+ cells on any substrate other than the longest (450 bp).

The second way we examined these data was to compare the transformation rates of the 5 and 11% diverged constructs, in both the MMR+ and MMR- cells, relative to the perfectly matched substrates. This is quantified in Table 2 by determining the extent of the reduction in average transformation rate of each 5 or 11% diverged substrate relative to the sequence-matched substrate of cognate length. In both the MMR+ and MMR- cells, at virtually all substrate lengths, increasing sequence divergence caused a progressively more severe impairment in transformation rate. However, in MMR+ cells this effect was more pronounced on 450 and 200 bp substrates than on any substrate shorter than 200 bp.

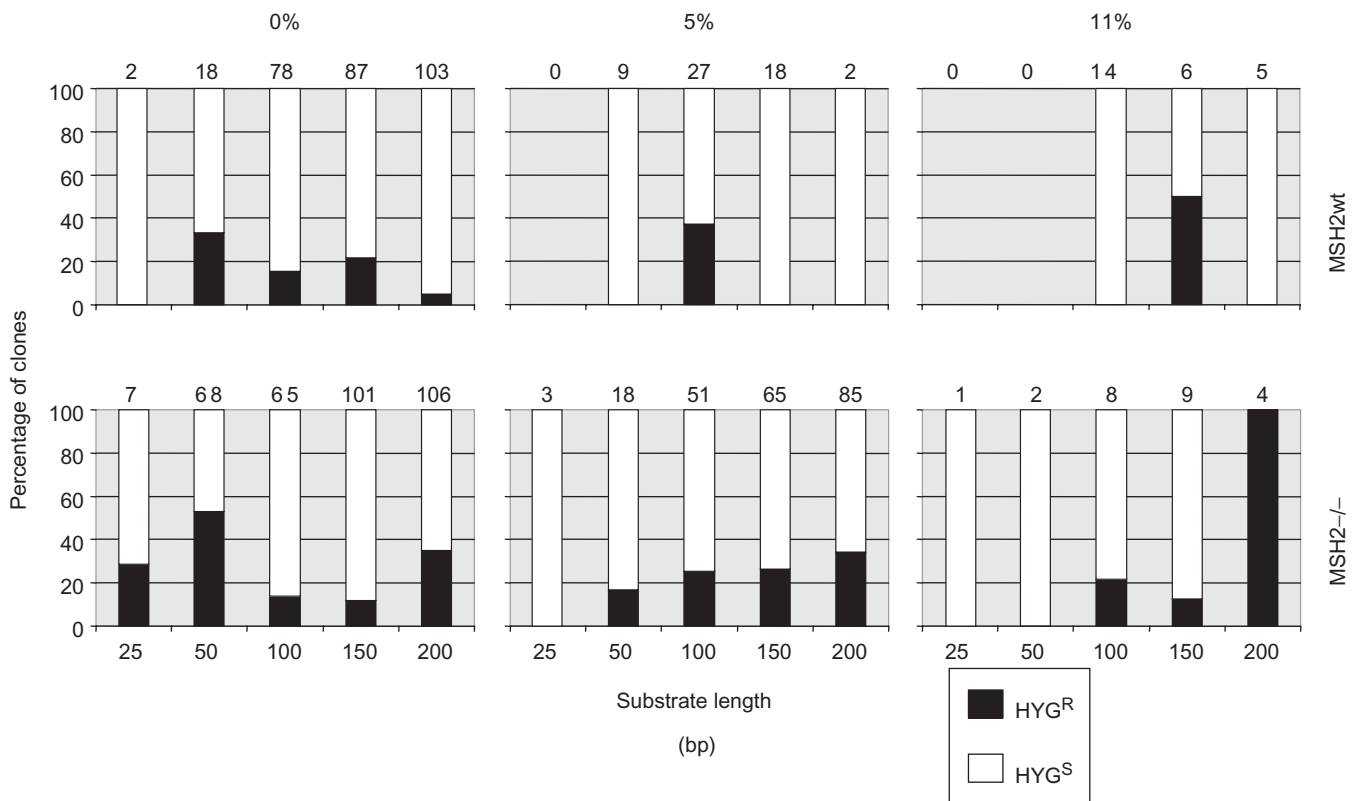


Figure 4. Antibiotic resistance patterns of transformants generated in the recombination assay. The bar charts show the proportion of hygromycin resistant (Hyg^R; black) and sensitive (Hyg^S; white) transformant clones recovered in wild type (wt) or *MSH2* homozygous (*-/-*) mutant *T. brucei* bloodstream stage cells following the transformation of constructs of varying substrate length and 0, 5 or 11% divergence from the genomic *HYG* sequence. Values above the charts indicate the number of transformants analysed.

In contrast, in MMR⁻ cells the extent to which sequence divergence impaired transformation was more uniform across all lengths. The average level of reduction in transformation of 5% diverged substrates smaller than 200 bp in MMR⁺ cells was 8.1-fold, compared with a 3-fold reduction of all substrates in MMR⁻ cells. For 11% diverged substrates, the reduction in transformation rate for substrates smaller than 200 bp was the same as all substrates in MMR⁻ cells (26.6- and 26.7-fold, respectively). Taken together, these data argue that *T. brucei* MMR has an important role in regulating recombination between diverged sequences when the substrates are long (around 200 bp or longer), but MMR has a less significant role on shorter substrates. Higher levels of sequence divergence presumably exacerbate this because the lengths of uninterrupted homology stretches in such substrates are reduced.

Most *T. brucei* integrative recombination occurs by homologous recombination

An important consideration in the above analysis, and in previous work using the same assay (35), is whether or not the experimental approach measures *T. brucei* homologous recombination on all substrates. For instance, it may be that MMR has less influence on recombination of short substrates because primarily non-homologous pathways of recombination act upon such sequences.

Indeed, we have shown that some DNA integration in *T. brucei* can occur by a pathway mediated by sequence microhomology (46). Homologous recombination of the constructs into the genomic *HYG* locus should lead to loss of hygromycin resistance (Hyg^S) due to disruption of the gene by *BLE*. We therefore tested the antibiotic resistances of a large number of transformants (Figure 4), revealing that a significant proportion of transformants in wild type, *MSH2*^{+/-} (data not shown) and *msh2*^{-/-} cells retained hygromycin resistance (Hyg^R). There was, however, no clear relationship between the retention of functional *HYG* and either substrate size or level of sequence homology. Although the numbers of Hyg^R transformants appeared to increase in wild-type cells as substrate length decreased from 200 to 50 bp at 0% divergence, the same trend was not apparent in *msh2*^{-/-} cells (Figure 4) or in *MSH2*^{+/-} cells (data not shown). Determining the average level at which Hyg^R cells arose, using the accumulated data for all substrate lengths, showed that this is not influenced by the extent of divergence: 15.1, 9.2 and 16.7 of transformants were Hyg^R for the 0, 5 or 11% diverged substrates in wild-type cells, respectively, compared with 28, 20.5 and 26.9% of transformants in *msh2*^{-/-} cells. These latter data do, however, indicate that Hyg^R transformants appeared to be slightly more prevalent (around 2-fold) when *MSH2* was mutated, perhaps indicating that

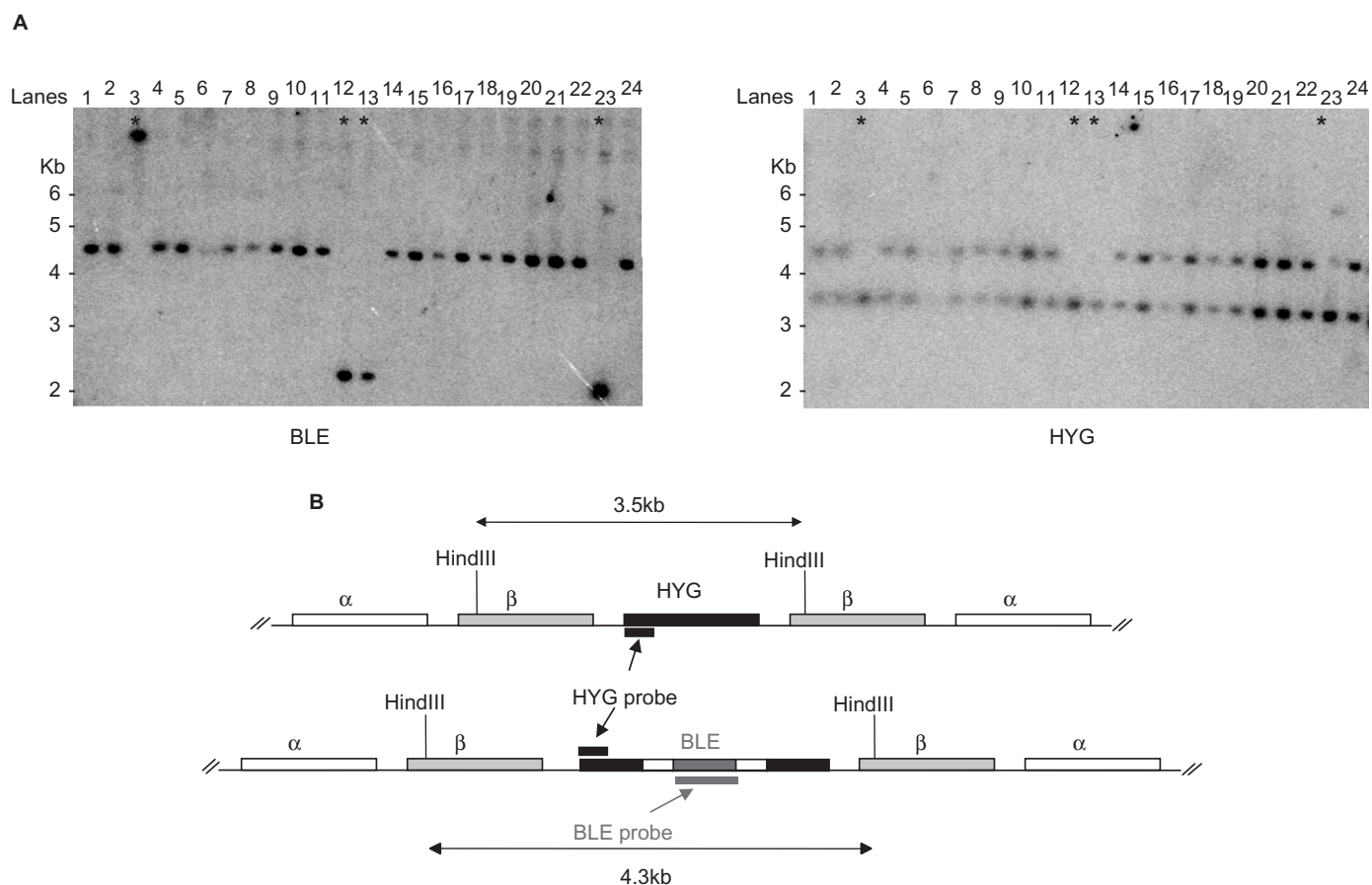


Figure 5. Genomic analysis of hygromycin-resistant transformants in *msh2*^{-/-} mutants. (A) A Southern blot of genomic DNA from *MSH2* homozygous mutant transformant clones digested with *Hind*III and probed with the bleomycin resistance gene ORF (*BLE*). The same blot was then stripped and re-probed with a portion of the hygromycin-resistance gene ORF (*HYG*). Asterisks denote transformants in which the constructs have integrated aberrantly. The constructs that gave rise to each transformant are in the following lanes: 1–3, 0%–50 bp; 4–5, 0%–100 bp; 6–7, 0%–150 bp; 8–9, 0%–200 bp; 10, 5%–150 bp; 11, 5%–200 bp; 12–13, 11%–100 bp; 14, 11%–150 bp; 15, 0%–100 bp; 16–17, 0%–150 bp; 18–19, 5%–100 bp; 20–22, 5%–200 bp; 23, 11%–100 bp; 24, 11%–150 bp. (B) A depiction of the expected restriction map of the *HYG* locus before (upper diagram) and after (lower diagram) homologous integration of *BLE*. The positions of probe fragments and the expected size of *Hind*III fragments are indicated.

MMR has an influence on the process that leads to their formation.

To examine the recombination pathway(s) that gives rise to *Hyg*^R transformants, a number of clones were examined by Southern analysis. The large majority of *Hyg*^R transformants had integrated the constructs by homologous recombination. In *msh2*^{-/-} cells, 20 of the 24 clones examined had maps indicative of *HYG* disruption by homologous integration of *BLE* (Figure 5). Re-probing the blot with a portion of the *HYG* ORF revealed that the antibiotic marker had been duplicated (based on the equivalent signal intensity of the two *HYG* bands) in these clones, explaining why they retained resistance to hygromycin. The four other *Hyg*^R *msh2*^{-/-} transformants had not integrated *BLE* into the tubulin *HYG* target, but instead into unmapped genomic locations, and the *HYG* marker was unaltered. A broadly similar pattern of *BLE* integration was seen for *Hyg*^R transformants in wild-type cells (data not shown). Here, 24 clones were examined, and 23 had disrupted *HYG* by homologous

recombination but retained a functional copy, whilst 1 had integrated *BLE* into an aberrant genomic location. The greater number of aberrant integrations in the *msh2*^{-/-} cells may indicate that this minor recombination pathway gains prominence in the absence of MMR. Furthermore, most aberrant integrations were seen using relatively short or diverged substrates: Three of the *msh2*^{-/-} clones had arisen from 11% diverged substrates of 100 bp (Figure 5, lanes 12, 13 and 23) and one had arisen from a 50 bp, 100% homologous substrate (Figure 5, lane 3); the wild type aberrant integrant arose from a 100 bp, 100% homologous substrate.

The above data suggest that two putatively distinct pathways operate in the context of this *T. brucei* recombination assay to generate hygromycin-resistant transformants. A minor pathway is integration events that do not target the tubulin locus, as directed by the terminal targeting flanks, but into other genomic loci. Although we have not mapped these integrations, the background rate at which they occur is very reminiscent of

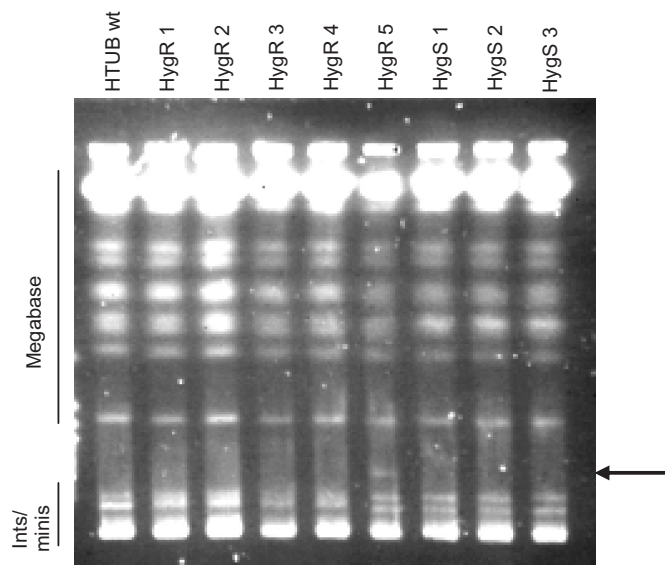


Figure 6. Pulse field gel electrophoretic analysis of transformants. Intact chromosomal DNA from untransformed cells (HTUBwt), and from 5 hygromycin resistant (Hyg^{R}) and 3 hygromycin sensitive (Hyg^{S}) transformants, is shown. Bands corresponding to the *T. brucei* megabase-sized chromosomes and to the intermediate-sized and mini-chromosomes (Ints/minis) are indicated. The arrow indicates the position of a novel, 370 kb molecule in the $\text{Hyg}^{\text{R}}5$ transformant.

the microhomology-mediated reactions we have described before (46). The prevalent pathway, in contrast, is a homologous recombination reaction associated with retention of a functional *HYG* gene. A number of processes could account for this. First, it is possible that duplication (or amplification) of *HYG* occurs at a relatively high rate in the growth of these clonal *T. brucei* lines, perhaps due to recombination within the multigenic tubulin locus. Second, larger scale changes in chromosome copy number (for instance, trisomy arising from replication errors) could be rather common in *T. brucei*. Finally, it is possible that a form of recombination, termed break-induced replication (53), could be commonly induced by such construct integrations, leading to the duplication of large stretches of chromosome 1, or even the complete chromosome. To examine this, we characterized a selection of Hyg^{S} and Hyg^{R} transformants from wild-type cells by pulsed field gel electrophoresis (Figure 6). One Hyg^{R} transformant ($\text{Hyg}^{\text{R}}5$), which arose as a result of an aberrant integration, contained a novel chromosome, 370 kb in size, that contained *BLE* and *HYG* sequence (data not shown). Such alterations in chromosome structure have been seen previously in microhomology-mediated reactions (46), providing further indirect evidence as to the recombination pathway involved in these integrations. In contrast, no differences in karyotype were seen in the other Hyg^{R} transformants (that had utilized homologous recombination; $\text{Hyg}^{\text{R}}1-4$) when compared with the Hyg^{S} cells or the parental strain, suggesting that large changes in chromosome structure are not associated with these events. Furthermore, probing of a Southern blot of these transformants (Supplementary Figure 3) indicated that there was no difference in the

relative amount of chromosomes 1 and 2. Since the tubulin array in which *HYG* is inserted is located on chromosome 1, this suggests that integration of *BLE* did not result in the generation of a new copy of the chromosome. Taken together, it seems likely that Hyg^{R} in some transformants is not a consequence of targeted integration by *BLE*, but results from events that amplify *HYG* in the tubulin locus and are a constant background process in the HTUB cell lines.

Construct integration occurs by independent invasion of the DNA ends, and mismatches are repaired by short-patch repair

To examine the mechanisms that contribute to the recombination and processing of the constructs during integration, we sequenced the DNA surrounding the homologously integrated *BLE* marker in a number of transformants clones generated in both wt and *msh2*^{-/-} cells using constructs with 50, 100, 150 or 200 bp flanks of 5 or 11% sequence divergence. We did not examine any clones that had integrated *BLE* aberrantly. The results in Figure 7 depict the pattern of residues in the transformant DNA that are mismatched between the constructs and wt *HYG* target. Most of the transformants had a *trans* pattern of sequences, in which the majority of mismatched residues corresponded to the sequence of the construct DNA on one side of *BLE* and the genomic *HYG* target on the other; no difference in this pattern was observed if the transformants were Hyg^{R} or Hyg^{S} (data not shown). Only clone 9 (from wt cells using a 100 bp, 5% diverged construct) was substantially different, with *HYG* sequence in both directions. Clone 23 (*msh2*^{-/-} cells with 200 bp, 5% construct) was somewhat different also, with predominantly *HYG*-derived residues. Such a *trans* pattern is consistent with a model for targeted gene replacement in yeast (54) and mammals (55) involving independent strand invasions by both arms of the construct (Figure 9). Such a model predicts that heteroduplex DNA can form by strand invasion of each construct end, which can result in 'sectored' or mixed sequence at the mismatched residue positions following replication of the heteroduplex (see Figure 9). However, no such mixed sequence was observed in this study: at each position where a base mismatch might form during strand invasion, the sequences were clearly either construct-derived or *HYG*-derived. This infers either that heteroduplex does not form or that repair of such mismatches is normally rapid and occurs before replication. The distribution of construct and *HYG*-derived sequences was very similar in the *msh2*^{-/-} and wt transformants, indicating that if mismatch repair occurs it can proceed independently of MSH2. In each transformant in which construct-derived residues were present in the transformed DNA, this was not continuous along the length of the flank, but was combined with *HYG*-derived residues. Most likely, *HYG*-derived residues that are distal to *BLE* in the flanks of transformants with an otherwise continuous tract of construct-derived sequence (clones 6, 22, 35, 30, 24, 37, 38) result from nucleolytic degradation of the ends of the linear molecules following transformation and before integration. This was most extensive in clone 30 (maximally 76 bp), whereas in a

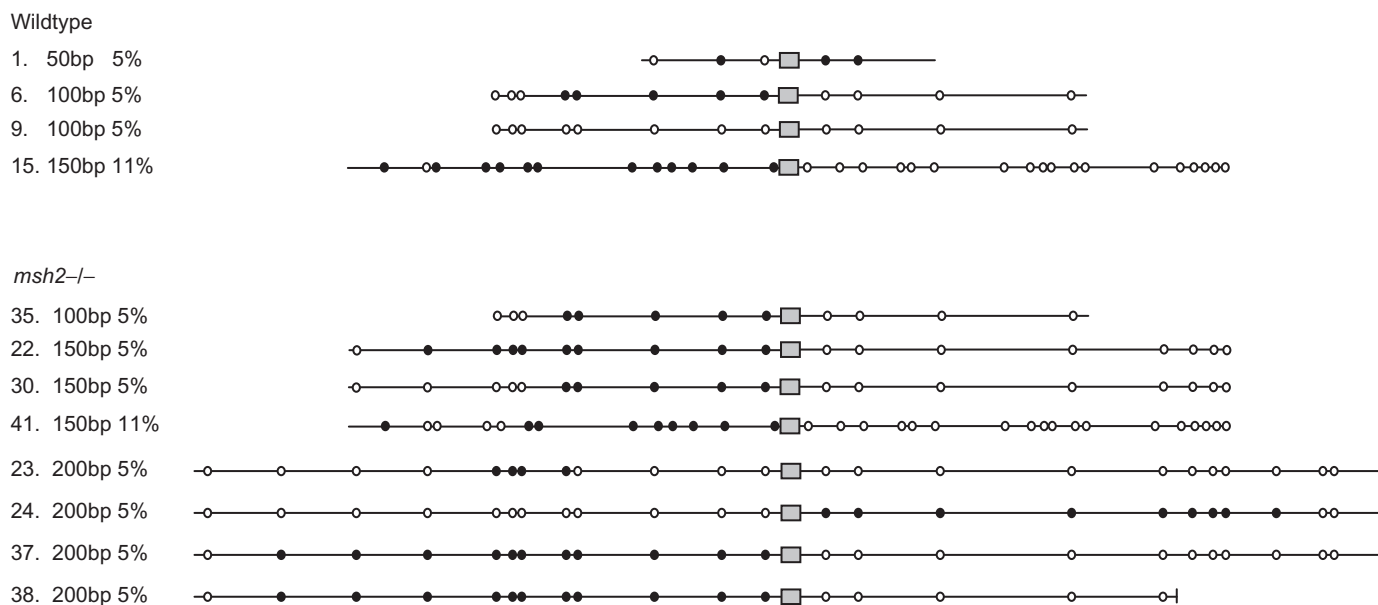


Figure 7. Mismatched base patterns in the DNA of integrated constructs. Circles denote the positions of bases that are mismatched between transformation constructs and the genomic *HYG* gene that is targeted by homologous recombination; filled circles denote that the sequence of a base matches the construct sequence, while open circles denote genomic *HYG* base sequence. The targeting flanks of the constructs are shown as lines, and the grey box indicates the central *BLE* resistance marker used for selection of transformants. Clone numbers for each DNA sequence are shown to the right, as well as the length of the flanks, the extent of sequence divergence relative to *HYG*, and whether the transformants were generated in wild type or *msh2*^{-/-} *T. brucei* cells.

number of other clones mismatched residues close to the ends of the construct molecules had been integrated (e.g. within 15 bp in clones 41 and 15, and 32 bp in clones 37 and 38), arguing that such degradation need not be substantial. In a number of other transformants, the flanks contained a discontinuous pattern of construct and *HYG*-derived sequence tracts (clones 1, 15, 41, 23), and some of the residues that were patterned in this way were positioned in close proximity (e.g. separated by only 2 and 8 bp in clones 15 and 23, respectively). This patterning is most simply explained by the formation of heteroduplex DNA during homologous integration and the repair of base mismatches prior to replication by short-patch mismatch repair.

DISCUSSION

The primary conclusion from this work is that *T. brucei* homologous recombination has substrate characteristics that are typical of those described in other organisms. The significance of this lies in our understanding of the relationship between homologous recombination and the critical immune evasion process of antigenic variation, which involves switches in *VSG* expression. Genetic evidence points to antigenic variation being closely linked to homologous recombination, despite the fact that *VSG* gene switching occurs at high rates (30,31), often acts upon rather short, diverged sequences (22,33,34) and frequently recombines *VSGs* on different chromosomes to the *VSG* ES (14,34), characteristics that are unusual for stochastic homologous recombination. We have shown previously that the efficiency of *T. brucei* homologous

recombination is dependent on substrate homology, which is at least partially controlled by the *T. brucei* MMR machinery (35). We now show here that recombination efficiency is also dependent on substrate length, at least over the range 25–200 bp. This indicates that specific features of *VSG* switching have not arisen through modifications of the *T. brucei* recombination machinery in these two key features. This conclusion is perhaps not surprising, as the primary function of homologous recombination is to repair DNA damage (7), and therefore fundamental reaction changes could have far-reaching effects on genome integrity. Any specificity of homologous recombination during *VSG* switching must therefore involve elements or factors that have not yet been uncovered and may be unique to *T. brucei*. Similar conclusions have been reached regarding antigenic variation in pathogenic *Neisseria* sp., where convergent evolution has produced a related immune evasion reaction that is also closely linked to homologous recombination (13).

The relationship we describe in *T. brucei* between substrate length and recombination rate (measured indirectly as transformation rate) appears to be conserved throughout evolution. In *E. coli*, three studies have recorded reductions in recombination efficiency over the range 74–20 bp (56), 405–27 bp (57) and 200–25 bp (58). In eukaryotes, *S. cerevisiae* recombination rate appears to decrease over 960–80 bp (59) or 2 kb–26 bp (60), and in mammalian cells a similar relationship has been recorded between an upper length of around 6.8 (61) to 10 kb (62) and a lower length between ~160 (63) and 330 bp (64). Variations in whether the relationship between

recombination rate and substrate length is linear or exponential and in the range of substrate sizes involved, both within and between organisms, presumably reflect differences in the assays used, meaning that it is difficult to compare the absolute sequence requirements of the recombination machineries between each organism. Nevertheless, a sharp drop in recombination efficiency below a lower threshold, referred to as the minimal efficient processing segment (MEPS)(57), has been described in *E. coli* (57), *S. cerevisiae* (59) and mammalian cells (63). We observed the same phenomenon in *T. brucei*, reinforcing our view that the recombination machinery in the parasite operates in the same manner. Despite this broad evolutionary conservation, previous reports have suggested that *T. brucei* recombination efficiency is not affected by substrate length over the range 50–400 bp (40,65), which is difficult to reconcile with the findings detailed here. One of these reports compared constructs that target distinct genomic locations (40), and it may be that differences in target accessibility (for instance, through chromatin structure or transcription level) affect recombination efficiency and confounded the analysis. Another explanation may be revealed by linear regression analysis (Figure 8), which suggests that the MEPS for *T. brucei* on sequence-matched substrates is around 31 bp, shorter than the estimates in either *S. cerevisiae* or mammals (~250 bp) (59), and closer to *E. coli* (around 25 bp). Intriguingly, transformation data from *Leishmania*, a related kinetoplastid parasite, failed to recover any integrants when DNA constructs with less than around 220 bp of targeting flank were used (66). Though this dichotomy with *T. brucei* may reflect the limitations of transformation in the two parasites rather than recombination, linear regression analysis of the (albeit limited) data set from Papadopoulou and Dumas (1997) is consistent with a MEPS of 190 bp in *Leishmania* (Supplementary Figure 2), closer to the size predicted in the two other eukaryotes. Perhaps, therefore, *T. brucei* has evolved to allow homologous recombination to operate, with reduced efficiency, on shorter substrates. More analysis will be needed to examine this, but it could be the result of differences in the activity of RAD51, differences in the factors that mediate RAD51 function, or a more active short-sequence recombination pathway.

The second conclusion from this work is that *T. brucei* contains an MMR-independent, or at least an MSH2-independent, pathway for homologous recombination that has not previously been described. On 5% diverged substrates in the size range 50–175 bp, and on 11% diverged substrates in the range 50–200 bp, this pathway appears to assume considerable significance, but is subordinate to MMR-dependent recombination on longer substrates (450 bp) containing these levels of sequence divergence. This argues that the reaction assumes a greater role during *T. brucei* homologous recombination as substrate length decreases. Characterization of transformants by antibiotic resistance and Southern analysis suggests that the considerable majority of *T. brucei* recombination, at all sequence lengths, occurs by homology, suggesting that the MMR-independent reaction is capable of precise

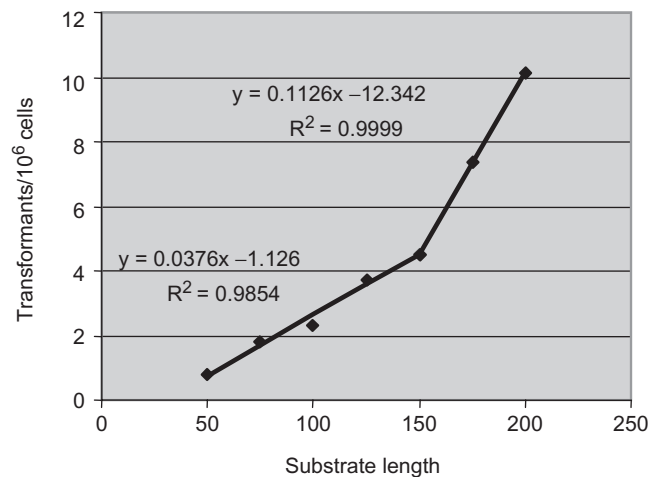


Figure 8. Linear regression analysis comparing the relationship between transformation rate and substrate length in *T. brucei*. Transformation rate was plotted for the combined wild type and *MSH2*^{+/-} (MMR proficient; MMR⁺) data for the sequence-matched substrates (0% divergence) in Table 1 relative to substrate length. The coefficient of determination (R^2) and the equations used to calculate the lines of best fit are indicated; only rates in the linear range (substrates from 200 to 50 bp) were included.

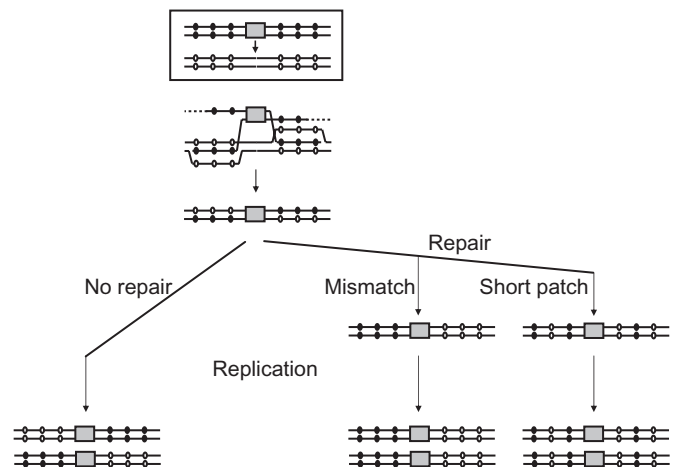


Figure 9. Gene targeting by two-end invasion. Grey boxes denote the *BLE* selectable marker. Lines indicate the flanks of the transformation construct or *HYG* target, within which black circles denote bases in the targeting construct that are mismatched with bases in *HYG*, which are shown in equivalent positions but as white circles. The boxed diagram denotes the route of targeted integration. Below is a model for integration showing how the 3' single-strand ends of the construct invade *HYG* (dotted lines denote 5'-3' nucleolytic degradation), and the result of resolution of the strand exchange intermediate. The pattern of mismatched bases in the integrated DNA is then shown following replication of the chromosome to yield two daughters, distinguishing the outcomes if no DNA repair of the mismatches occurs (left), or if mismatch repair precedes replication and occurs in favour of the lower DNA strand (middle), or if repair occurs by short-patch repair with an incomplete bias towards the lower strand (right).

integration. We have described a DNA repair process in *T. brucei* based on the joining of DNA molecules using very short stretches of homology, typically 5–15 bp in length (46), and similar reactions have been described in other eukaryotes (67,68) and potentially in bacteria (69–71).

We suspect that this is not the pathway responsible for most *T. brucei* MMR-independent recombination, for two reasons. First, examination of the transformation rates of the 5% diverged substrates 175 bp or smaller, and of the 11% diverged substrates 200 bp or smaller, suggests that transformation by the MMR-independent pathway can be as efficient as 1×10^{-6} transformants \cdot cells $^{-1}$, since these are conditions in which it is likely that most integrations occur by this route. This is considerably higher than the maximum transformation rate (0.1×10^{-6}) described for the microhomology reaction in *RAD51* wild-type *T. brucei* (46). Second, in most reactions where we have demonstrated that microhomology guides construct integration, this is associated with visible karyotype changes (46), suggesting that the reaction may be an end-joining process that exploits random DNA breaks, leading to genomic rearrangements. In contrast, such rearrangements are very rare in the extensive numbers of transformation events we have characterized here, including the conditions in which MMR-independent recombination would be expected to be prominent.

The presence of a putative MMR-independent recombination reaction in *T. brucei* is reminiscent of pathways described in *S. cerevisiae*, at least superficially. Although Rad51 is a central factor in DNA strand exchange during homologous recombination, some reactions in *S. cerevisiae* can occur in its absence; reviewed in (72). Rad52 appears to be essential for nearly all recombination in yeast, suggesting it contributes to both Rad51-dependent and -independent recombination. A relative of Rad52, termed Rad59, was identified in a search for factors required for recombination in *S. cerevisiae rad51* mutants (73), and the suggestion that both proteins contribute to Rad51-independent reactions is supported by findings that each can catalyse strand annealing (74). Rad59-dependent recombination requires less sequence homology than the Rad51-dependent reaction (53), and displays a lower degree of regulation by Msh2, and hence MMR, during recombination of diverged DNA sequences (75). A prediction, therefore, is that the MMR-independent recombination we see in *T. brucei* is RAD51-independent. We have not tested this directly, but some data appear consistent with this hypothesis. Linear regression analysis of the data presented here on substrate length and *T. brucei* transformation rate is best accounted for by two lines of best fit, consistent with two pathways operating (Figure 8). Above around \sim 150 bp a pathway predominates that has a MEPS of \sim 100 bp, whereas below \sim 150 bp a distinct pathway with a MEPS of \sim 31 bp. Interestingly, the longer MEPS is consistent with previous estimates of RAD51-dependent *T. brucei* recombination acting on long (450 bp) substrates (35). Despite the considerable inefficiency of homologous recombination on substrates shorter than the MEPS, the assay we have used here shows that such recombination can occur, as has been found in yeast (76), mammalian cells (64) and *E. coli* (where recombination of 25 bp substrates is predominantly RecA-independent)(58). Indeed, 80% of the transformants that arose from the constructs with only 25 bp of homology were Hyg^S (Figure 3), demonstrating that they had integrated by homologous recombination.

In *S. cerevisiae*, recombination of short substrates around 29–40 bp is not only Rad51-independent, but mutation of Rad51 increases the reaction efficiency (53). In *T. brucei*, transformation of constructs with very similar-sized regions of homology (24 bp) occurs at essentially the same frequency in *rad51* $^{-/-}$ and wild-type cells, suggesting the action of a RAD51-independent pathway (46). Despite these similarities, it is unclear what factors would catalyse RAD51-independent homologous recombination in *T. brucei*. Rad52 and Rad59 belong to a superfamily that is not conserved universally (77), and homologues of both proteins are either absent from the *T. brucei* genome (78), or are sufficiently diverged in sequence to have escaped detection. It is therefore possible that *T. brucei* contains recombination factors, thus far unidentified and distinct from RAD51, that can perform the functions equivalent to the pathway in which Rad59 and Rad52 act in *S. cerevisiae*.

Identification of an MMR-independent pathway of homologous recombination could be important in understanding antigenic variation and genetic variability in *T. brucei*. Although antigenic variation is impaired in RAD51 mutants, *VSG* gene conversion reactions can still be catalysed (28). It is plausible that the MMR-independent pathway, if it is RAD51-independent, could explain these residual *VSG* switching events. There is also no evidence that *VSG* switching is influenced by MMR (35), which appears to be at odds with the requirement for RAD51, since this pathway should be suppressed by MMR. However, although assays for recombination in yeast *rad51* mutants reveal functions for Rad59 in defined recombination pathways, it is very likely that in wild-type cells Rad51, Rad59 and Rad52 actually act together (53,75). *Trypanosoma brucei* *VSG* switching may therefore occur by a specific recombination pathway that requires RAD51, but is directed towards an MMR-independent route by unidentified factors that act like Rad59. Indeed, this could explain the ability of *VSG* switching to use rather short and dissimilar DNA sequences as substrates, such as the 70 bp repeats upstream of *VSG* genes. Furthermore, although nothing is known about the genetic requirements of segmental gene conversion involving the *VSG* pseudogenes, it is clear that *VSG* genes share very little primary sequence homology (<25% identity between encoded amino acids in >95% of the *VSG* repertoire; L.Marcello and J.D.Barry, personal communication). It is therefore possible that an MMR-independent reaction, active on short stretches of homology, would be involved. It is important to note that to date we have only assayed *T. brucei* homologous recombination at one interstitial site. Although some work has suggested that *T. brucei* recombination occurs at equivalent efficiencies in different genomic locations (40), it is important to examine this systematically and to assess the pathways used. For instance, it is clear in other organisms that other factors, such as transcription (79), can influence recombination, and some work has suggested that different pathways of recombination are active in interstitial relative to subtelomeric environments (80). Moreover, recombination can contribute to the maintenance of

telomeres (9), and the pathway(s) that acts in this regard in *T. brucei* is unknown (81).

In any model in which MMR triggers the rejection of recombination between mismatched DNA substrates, heteroduplex DNA must form during the strand exchange step. In MMR+ cells in which recombinants escape rejection, such heteroduplex is likely to be repaired, whereas it should be visible in MMR- cells. An MMR-independent pathway may or may not generate heteroduplex, since it could avoid MMR surveillance by using short, perfectly matched sequences or because the strand exchange mechanism between mismatched molecules avoids triggering MMR. We cannot readily distinguish these possibilities through sequencing the integrated DNA in these experiments. In no transformant, even in MMR- cells, did we find evidence for the DNA having a mixture of construct and genomic *HYG* sequence at the mismatched residues, which would arise if heteroduplex formed and was not repaired prior to replication, as illustrated in Figure 9. Instead, we saw a predominant *trans* pattern of construct sequence to one side of the integrated marker and *HYG* sequence on the other. This is most readily explained by independent strand invasion and annealing between a 3' single strand on each side of the *BLE* marker and the genomic *HYG* target, suggesting that the mechanism of homologous recombination during targeted gene replacement in *T. brucei* is equivalent to that in yeast and mammals (54,55). The one exception we found, where all the sequence is *HYG*-derived, occurred in wild-type cells (clone 9, Figure 7). This could result from MMR of the heteroduplex in favour of the recipient, *HYG* DNA. Alternatively, it could indicate that construct integration occasionally occurs by annealing of a single strand encompassing both arms of the construct, followed by mismatch repair; such a mechanism has been seen in yeast (82), though appears to be rare (54). One distinction between targeted gene replacement in *T. brucei* and *S. cerevisiae* is revealed here: mutation of *MSH2* in *T. brucei* either increased the frequency of transformation or had no effect, depending on substrate length, whereas the same mutation in *S. cerevisiae* results in reduced integration rates (83). This may indicate that the parasite *MSH2* protein does not promote recombination in the ways it has been found to do elsewhere (84).

The lack of clear evidence for heteroduplex DNA is perplexing. In wild-type cells, the lack of mixed sequence at mismatched residues could be explained by MMR that acts following construct integration and is directed towards one or other DNA strand by, for instance, the direction of replication (Figure 9). However, the same pattern was seen in the *msh2*-/- transformants. In addition, we did not see any striking difference in this pattern as construct length changes, which could arise through a shift in the substrate requirements of the recombination pathways being used. It is conceivable that in all these conditions most strand exchange is limited to perfectly matched sequences between the construct and *HYG* target, and crossovers then occur on these intermediates to incorporate both strands of the construct DNA. However, a number of findings argue against this. First, we did not observe construct sequence on both sides

of *BLE*, and it is not clear how crossover integration could be directional in this way. Second, it is difficult to explain by this model the rather common appearance of transformants in which tracts of construct sequence were interrupted by patches of *HYG* sequence in the 'left' arms (clones 1, 15, 41, 23; Figure 7). Finally, the length of homology that mediates strand pairing would have to have been very short in some cases (e.g. 15 bp in clone 41, and 12 bp in clones 6, 35 and 30; Figure 7). For these reasons, we suggest that heteroduplex does form during homologous integration and that it is repaired by short-patch MMR. This would explain the predominant pattern of sequence we see in the integrated DNA (Figure 9). Furthermore, it most readily explains transformants in which construct-derived and *HYG*-derived residues were found in close proximity. Such a situation is consistent with repair of two putative mismatches in different directions (in favour of the invading and recipient strands), which is distinct from long-patch MMR, where a mismatch triggers excision of extensive regions of DNA to allow repair (43). However, given that the *trans* pattern of sequence in the transformants was predominantly with construct sequence to the 'left' of *BLE* and *HYG* sequence to the 'right', and that many tracts were continuous stretches of either construct or *HYG* residues, it seems likely that some feature(s) biases the direction of short-patch repair. The nicks that are present during strand invasion cannot explain this, as repair would then be in the same direction for each end of the construct. Replication or transcription seems like a reasonable alternative, though this cannot be absolute, since we see discontinuous tracts in a number of transformants, and one clone has reversed the predominant pattern (clone 23; Figure 7). Short-patch MMR has not been described in *T. brucei*, but has been found in other organisms (including during recombination), though the molecular machinery is still being described (85,86). A consequence of the lack of identifiable heteroduplex DNA in this study is that we cannot determine how extensive the lengths of strand exchange intermediates are in these experiments, nor whether this differs in wild-type and *msh2*-/- cells, as has been reported in yeast (87,88). In addition, this approach does not allow us to test our hypothesis that the putative MMR-independent pathway we propose utilizes shorter substrates. However, given the existence of mixed tracts of construct and *HYG* sequence at all substrate lengths, it is likely that heteroduplex is formed in all conditions.

A footnote in this study, which was not appreciated in our previous analyses (35), is that we find significant levels of duplication of the *HYG* locus in this assay. The available evidence suggests that this is due to the generation of additional copies of *HYG* in the tubulin array during growth of the clonal HTUB *T. brucei* cell lines, rather than being a consequence of the targeted gene replacement. This 'spread' of a resistance marker in tubulin has been described previously (89), and shown to be due to unequal sister chromatid exchange. At least in our experiments this appears to be rather frequent. Though this may be a result of the antibiotic exerting a selective pressure for increased gene product, it is notable that high frequencies of allelic gene conversion have been

described in *T. brucei* (90) and that considerable variation in repetitive sequences has been described between *T. brucei* strains (91). Our data also hint that impairment in MMR may enhance *HYG* amplification, perhaps because MMR-mediated suppression of recombination between divergent sequences is alleviated, leading to greater rearrangements in the *T. brucei* genome.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGMENTS

RMcC is a Royal Society University Research Fellow and thanks the Medical Research Council and Wellcome Trust for grant support. Thanks to all our colleagues in WCMP for comments and discussions, in particular Lucio Marcello and Dave Barry for sharing data on *VSG* sequence homology, Craig Lapsley for sterling DNA help and Rachel Dobson and Claire Hartley for critical reading of the manuscript. Funding to pay the Open Access publication charges for this article was provided by the Wellcome Trust.

REFERENCES

- Aravind,L., Walker,D.R. and Koonin,E.V. (1999) Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res.*, **27**, 1223–1242.
- Fujisawa,H., Yonesaki,T. and Minagawa,T. (1985) Sequence of the T4 recombination gene, *uvrX*, and its comparison with that of the *recA* gene of *Escherichia coli*. *Nucleic Acids Res.*, **13**, 7473–7481.
- West,S.C. (2003) Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell Biol.*, **4**, 435–445.
- 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.
- 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.
- McIlwraith,M.J., Hall,D.R., Stasiak,A.Z., Stasiak,A., Wigley,D.B. and West,S.C. (2001) RadA protein from *Archaeoglobus fulgidus* forms rings, nucleoprotein filaments and catalyses homologous recombination. *Nucleic Acids Res.*, **29**, 4509–4517.
- Michel,B., Grompone,G., Flores,M.J. and Bidnenko,V. (2004) Multiple pathways process stalled replication forks. *Proc. Natl Acad. Sci. USA*, **101**, 12783–12788.
- Neale,M.J. and Keeney,S. (2006) Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature*, **442**, 153–158.
- McEachern,M.J. and Haber,J.E. (2006) Break-induced replication and recombinational telomere elongation in yeast. *Annu. Rev. Biochem.*, **75**, 111–136.
- Vengrova,S. and Dalgaard,J.Z. (2006) The wild-type *Schizosaccharomyces pombe* mat1 imprint consists of two ribonucleotides. *EMBO Rep.*, **7**, 59–65.
- Deutsch,K.W., Moxon,E.R. and Welles,T.E. (1997) Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol. Mol. Biol. Rev.*, **61**, 281–293.
- Barry,J.D. and McCulloch,R. (2001) Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. *Adv. Parasitol.*, **49**, 1–70.
- Kline,K.A., Sechman,E.V., Skaar,E.P. and Seifert,H.S. (2003) Recombination, repair and replication in the pathogenic Neisseriae: the 3 R's of molecular genetics of two human-specific bacterial pathogens. *Mol. Microbiol.*, **50**, 3–13.
- Berriman,M., Ghedin,E., Hertz-Fowler,C., Blandin,G., Renauld,H., Bartholomeu,D.C., Lennard,N.J., Caler,E., Hamlin,N.E. *et al.* (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science*, **309**, 416–422.
- Taylor,J.E. and Rudenko,G. (2006) Switching trypanosome coats: what's in the wardrobe? *Trends Genet.*, **22**, 614–620.
- McCulloch,R. (2004) Antigenic variation in African trypanosomes: monitoring progress. *Trends Parasitol.*, **20**, 117–121.
- 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.
- Borst,P. (2002) Antigenic variation and allelic exclusion. *Cell*, **109**, 5–8.
- Pays,E., Vanhamme,L. and Perez-Morga,D. (2004) Antigenic variation in *Trypanosoma brucei*: facts, challenges and mysteries. *Curr. Opin. Microbiol.*, **7**, 369–374.
- 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.
- Aline,R.Jr, MacDonald,G., Brown,E., Allison,J., Myler,P., Rothwell,V. and Stuart,K. (1985) (TAA)_n within sequences flanking several intrachromosomal variant surface glycoprotein genes in *Trypanosoma brucei*. *Nucleic Acids Res.*, **13**, 3161–3177.
- Borst,P. and Cross,G.A. (1982) Molecular basis for trypanosome antigenic variation. *Cell*, **29**, 291–303.
- de Lange,T., Kooter,J.M., Michels,P.A. and Borst,P. (1983) Telomere conversion in trypanosomes. *Nucleic Acids Res.*, **11**, 8149–8165.
- 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.
- Roth,C.W., Longacre,S., Raibaud,A., Baltz,T. and Eisen,H. (1986) The use of incomplete genes for the construction of a *Trypanosoma equiperdum* variant surface glycoprotein gene. *EMBO J.*, **5**, 1065–1070.
- Kamper,S.M. and Barbet,A.F. (1992) Surface epitope variation via mosaic gene formation is potential key to long-term survival of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **53**, 33–44.
- Barry,J.D., Marcello,L., Morrison,L.J., Read,A.F., Lythgoe,K., Jones,N., Carrington,M., Blandin,G., Bohme,U. *et al.* (2005) What the genome sequence is revealing about trypanosome antigenic variation. *Biochem. Soc. Trans.*, **33**, 986–989.
- McCulloch,R. and Barry,J.D. (1999) A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes Dev.*, **13**, 2875–2888.
- Proudfoot,C. and McCulloch,R. (2005) Distinct roles for two RAD51-related genes in *Trypanosoma brucei* antigenic variation. *Nucleic Acids Res.*, **33**, 6906–6919.
- Turner,C.M. and Barry,J.D. (1989) High frequency of antigenic variation in *Trypanosoma brucei* rhodesiense infections. *Parasitology*, **99**(Pt 1), 67–75.
- 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.
- Johnson,R.D. and Jasin,M. (2000) Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *EMBO J.*, **19**, 3398–3407.
- Blundell,P.A., Rudenko,G. and Borst,P. (1996) Targeting of exogenous DNA into *Trypanosoma brucei* requires a high degree of homology between donor and target DNA. *Mol. Biochem. Parasitol.*, **76**, 215–229.
- Borst,P., Rudenko,G., Taylor,M.C., Blundell,P.A., van Leeuwen,F., Bitter,W., Cross,M. and McCulloch,R. (1996) Antigenic variation in trypanosomes. *Arch. Med. Res.*, **27**, 379–388.
- 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.
- Kadyk,L.C. and Hartwell,L.H. (1992) Sister chromatids are preferred over homologs as substrates for recombinational repair in *Saccharomyces cerevisiae*. *Genetics*, **132**, 387–402.

37. Richardson, C., Moynahan, M.E. and Jasin, M. (1998) Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. *Genes Dev.*, **12**, 3831–3842.
38. Burgess, S.M. and Kleckner, N. (1999) Collisions between yeast chromosomal loci in vivo are governed by three layers of organization. *Genes Dev.*, **13**, 1871–1883.
39. 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.
40. Wickstead, B., Ersfeld, K. and Gull, K. (2003) The frequency of gene targeting in *Trypanosoma brucei* is independent of target site copy number. *Nucleic Acids Res.*, **31**, 3993.
41. 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.
42. Jiricny, J. (2006) The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol.*, **7**, 335–346.
43. Modrich, P. (2006) Mechanisms in eukaryotic mismatch repair. *J. Biol. Chem.*, **281**, 30305–30309.
44. Harfe, B.D. and Jinks-Robertson, S. (2000) DNA mismatch repair and genetic instability. *Annu. Rev. Genet.*, **34**, 359–399.
45. 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.
46. Conway, C., Proudfoot, C., Burton, P., Barry, J.D. and McCulloch, R. (2002) Two pathways of homologous recombination in *Trypanosoma brucei*. *Mol. Microbiol.*, **45**, 1687–1700.
47. Seebeck, T., Whittaker, P.A., Imboden, M.A., Hardman, N. and Braun, R. (1983) Tubulin genes of *Trypanosoma brucei*: a tightly clustered family of alternating genes. *Proc. Natl Acad. Sci. USA*, **80**, 4634–4638.
48. 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.
49. 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.
50. Alsford, N.S., Navarro, M., Jamnadass, H.R., Dunbar, H., Ackroyd, M., Murphy, N.B., Gull, K. and Ersfeld, K. (2003) The identification of circular extrachromosomal DNA in the nuclear genome of *Trypanosoma brucei*. *Mol. Microbiol.*, **47**, 277–289.
51. Datta, A., Hendrix, M., Lipsitch, M. and Jinks-Robertson, S. (1997) Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. *Proc. Natl Acad. Sci. USA*, **94**, 9757–9762.
52. Nicholson, A., Hendrix, M., Jinks-Robertson, S. and Crouse, G.F. (2000) Regulation of mitotic homeologous recombination in yeast. Functions of mismatch repair and nucleotide excision repair genes. *Genetics*, **154**, 133–146.
53. 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.
54. Langston, L.D. and Symington, L.S. (2004) Gene targeting in yeast is initiated by two independent strand invasions. *Proc. Natl Acad. Sci. USA*, **101**, 15392–15397.
55. Li, J., Read, L.R. and Baker, M.D. (2001) The mechanism of mammalian gene replacement is consistent with the formation of long regions of heteroduplex DNA associated with two crossing-over events. *Mol. Cell Biol.*, **21**, 501–510.
56. Watt, V.M., Ingles, C.J., Urdea, M.S. and Rutter, W.J. (1985) Homology requirements for recombination in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **82**, 4768–4772.
57. Shen, P. and Huang, H.V. (1986) Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics*, **112**, 441–457.
58. Lovett, S.T., Hurley, R.L., Sutura, V.A.Jr, Aubuchon, R.H. and Lebedeva, M.A. (2002) Crossing over between regions of limited homology in *Escherichia coli*. RecA-dependent and RecA-independent pathways. *Genetics*, **160**, 851–859.
59. Jinks-Robertson, S., Michelitch, M. and Ramcharan, S. (1993) Substrate length requirements for efficient mitotic recombination in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **13**, 3937–3950.
60. Ahn, B.Y., Dornfeld, K.J., Fagrelus, T.J. and Livingston, D.M. (1988) Effect of limited homology on gene conversion in a *Saccharomyces cerevisiae* plasmid recombination system. *Mol. Cell Biol.*, **8**, 2442–2448.
61. Hasty, P., Rivera-Perez, J. and Bradley, A. (1991) The length of homology required for gene targeting in embryonic stem cells. *Mol. Cell Biol.*, **11**, 5586–5591.
62. Deng, C. and Capecchi, M.R. (1992) Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell Biol.*, **12**, 3365–3371.
63. Rubnitz, J. and Subramani, S. (1984) The minimum amount of homology required for homologous recombination in mammalian cells. *Mol. Cell Biol.*, **4**, 2253–2258.
64. Ayares, D., Chekuri, L., Song, K.Y. and Kucherlapati, R. (1986) Sequence homology requirements for intermolecular recombination in mammalian cells. *Proc. Natl Acad. Sci. USA*, **83**, 5199–5203.
65. Shen, S., Arhin, G.K., Ullu, E. and Tschudi, C. (2001) In vivo epitope tagging of *Trypanosoma brucei* genes using a one step PCR-based strategy. *Mol. Biochem. Parasitol.*, **113**, 171–173.
66. Papadopoulou, B. and Dumas, C. (1997) Parameters controlling the rate of gene targeting frequency in the protozoan parasite *Leishmania*. *Nucleic Acids Res.*, **25**, 4278–4286.
67. Boulton, S.J. and Jackson, S.P. (1996) *Saccharomyces cerevisiae* Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J.*, **15**, 5093–5103.
68. Guirouilh-Barbat, J., Huck, S., Bertrand, P., Pirzio, L., Desmaze, C., Sabatier, L. and Lopez, B.S. (2004) Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Mol. Cell*, **14**, 611–623.
69. de Vries, J. and Wackernagel, W. (2002) Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *Proc. Natl Acad. Sci. USA*, **99**, 2094–2099.
70. Prudhomme, M., Libante, V. and Claverys, J.P. (2002) Homologous recombination at the border: insertion-deletions and the trapping of foreign DNA in *Streptococcus pneumoniae*. *Proc. Natl Acad. Sci. USA*, **99**, 2100–2105.
71. Shiraishi, K., Hanada, K., Iwakura, Y. and Ikeda, H. (2002) Roles of RecJ, RecO, and RecR in RecET-mediated illegitimate recombination in *Escherichia coli*. *J. Bacteriol.*, **184**, 4715–4721.
72. Symington, L.S. (2002) Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol. Mol. Biol. Rev.*, **66**, 630–670, table.
73. Bai, Y. and Symington, L.S. (1996) A Rad52 homolog is required for RAD51-independent mitotic recombination in *Saccharomyces cerevisiae*. *Genes Dev.*, **10**, 2025–2037.
74. Wu, Y., Sugiyama, T. and Kowalczykowski, S.C. (2006) DNA annealing mediated by rad52 and rad59 proteins. *J. Biol. Chem.*, **281**, 15441–15449.
75. Spell, R.M. and Jinks-Robertson, S. (2003) Role of mismatch repair in the fidelity of *R. Genetics*, **165**, 1733–1744.
76. Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. and Cullin, C. (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **21**, 3329–3330.
77. Iyer, L.M., Koonin, E.V. and Aravind, L. (2002) Classification and evolutionary history of the single-strand annealing proteins, RecT, Redbeta, ERF and RAD52. *BMC Genomics*, **3**, 8.
78. El Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., Aggarwal, G., Tran, A.N., Ghedin, E., Worthey, E.A., Delcher, A.L. et al. (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*, **309**, 409–415.
79. Aguilera, A. (2002) The connection between transcription and genomic instability. *EMBO J.*, **21**, 195–201.
80. 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.
81. Dreesen, O. and Cross, G.A. (2006) Telomerase-independent stabilization of short telomeres in *Trypanosoma brucei*. *Mol. Cell Biol.*, **26**, 4911–4919.
 82. Leung, W., Malkova, A. and Haber, J.E. (1997) Gene targeting by linear duplex DNA frequently occurs by assimilation of a single strand that is subject to preferential mismatch correction. *Proc. Natl Acad. Sci. USA*, **94**, 6851–6856.
 83. Langston, L.D. and Symington, L.S. (2005) Opposing roles for DNA structure-specific proteins Rad1, Msh2, Msh3, and Sgs1 in yeast gene targeting. *EMBO J.*, **24**, 2214–2223.
 84. Saparbaev, M., Prakash, L. and Prakash, S. (1996) Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics*, **142**, 727–736.
 85. Coic, E., Gluck, L. and Fabre, F. (2000) Evidence for short-patch mismatch repair in *Saccharomyces cerevisiae*. *EMBO J.*, **19**, 3408–3417.
 86. Muheim-Lenz, R., Buterin, T., Marra, G. and Naegeli, H. (2004) Short-patch correction of C/C mismatches in human cells. *Nucleic Acids Res.*, **32**, 6696–6705.
 87. Alani, E., Reenan, R.A. and Kolodner, R.D. (1994) Interaction between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. *Genetics*, **137**, 19–39.
 88. Chen, W. and Jinks-Robertson, S. (1999) The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. *Genetics*, **151**, 1299–1313.
 89. Ten Asbroek, A.L., Mol, C.A., Kieft, R. and Borst, P. (1993) Stable transformation of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **59**, 133–142.
 90. Leal, S., Acosta-Serrano, A., Morris, J. and Cross, G.A. (2004) Transposon mutagenesis of *Trypanosoma brucei* identifies glycosylation mutants resistant to concanavalin A. *J. Biol. Chem.*, **279**, 28979–28988.
 91. 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.