Microhomology-mediated deletion and gene conversion in African trypanosomes

Lucy Glover, Junho Jun and David Horn*

London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK

ABSTRACT

Antigenic variation in African trypanosomes is induced by DNA double-strand breaks (DSBs). In these protozoan parasites, DSB repair (DSBR) is dominated by homologous recombination (HR) and microhomology-mediated end joining (MMEJ), while non-homologous end joining (NHEJ) has not been reported. To facilitate the analysis of chromosomal end-joining, we established a system whereby inter-allelic repair by HR is lethal due to loss of an essential gene. Analysis of intrachromosomal end joining in individual DSBR survivors exclusively revealed MMEJ-based deletions but no NHEJ. A survey of microhomologies typically revealed sequences of between 5 and 20 bp in length with several mismatches tolerated in longer stretches. Mean deletions were of 54 bp on the side closest to the break and 284 bp in total. Break proximity, microhomology length and GC-content all favored repair and the pattern of MMEJ described above was similar at several different loci across the genome. We also identified interchromosomal gene conversion involving HR and MMEJ at different ends of a duplicated sequence. While MMEJ-based deletions were RAD51-independent, one-sided MMEJ was RAD51 dependent. Thus, we describe the features of MMEJ in *Trypanosoma brucei*, which is analogous to micro single-strand annealing; and RAD51 dependent, one-sided MMEJ. We discuss the contribution of MMEJ pathways to genome evolution, subtelomere recombination and antigenic variation.

INTRODUCTION

Homologous recombination (HR) and non-homologous end joining (NHEJ) make the major contribution to mitotic double-strand break-repair (DSBR) and the generation of genetic diversity in organisms ranging from fungi to mammals (1). Important roles for microhomology-mediated end joining (MMEJ) or other forms of 'alternative end-joining' have recently emerged in class switch recombination in B cells (2) and in cancer development (3,4). However, since MMEJ is only revealed when NHEJ is disrupted in these cells, the pathway appears to serve only a 'back-up' function (5). In contrast, MMEJ dominates end-joining reactions in trypanosomes (6), divergent protozoan parasites of humans and livestock that rely upon DSBR for effective antigenic variation and immune evasion (7). This suggests that MMEJ is a universally conserved pathway that is obscured or even suppressed in organisms with the capacity for NHEJ. The prominence of MMEJ in trypanosomes presents a unique opportunity to study the features of this pathway.

End-joining mechanisms are non-conservative, typically resulting in sequence loss. NHEJ may sometimes exploit just a few paired nucleotides and most commonly results in loss of <10 bp on either side of the break (1). MMEJ is distinct in that it allows imperfect, but directly repeated sequence of 5–20 nt flanking the break to recombine following annealing of the complementary strands from each repeated sequence (8). Single-strand annealing (SSA) is another non-conservative DSBR mechanism that allows more extensive directly repeated sequences to recombine (9). The products of MMEJ and SSA contain only one copy of the repeated sequence, with the deletion of sequences originally present between the two repeats. Thus, by analogy to SSA, MMEJ is also known as micro-SSA and this idea is supported by genetic analysis (4,8,10,11). However, alternative end-joining, typically defined genetically as KU70/80 or ligase 4 independent, can differ substantially in different organisms (12); in some cases, no or little microhomology is required, the size of the deletions can vary substantially and insertions are sometimes observed.

DSBR and DNA rearrangement are central to the process of antigenic variation in African trypanosomes (7), but the mechanisms underlying this process remain only partially characterized. Switching of the variant surface glycoprotein (VSG) coat depends upon

*To whom correspondence should be addressed. Tel: +44 20 7927 2352; Fax: +44 20 7636 8739; Email: david.horn@lshtm.ac.uk

monoallelic expression of a *VSG* gene at a telomere (13) and a large reservoir of subtelomeric *VSG* templates, that can be used to copy new *VSGs* into the active telomeric expression site (ES). While key factors required for NHEJ are absent or diverged in trypanosomatids (14), RAD51-independent MMEJ with exogenous templates has been described in *Trypanosoma brucei* (15) and these reactions have also been shown to be KU-independent in *in vitro* assays (14). However, little is known about chromosomal MMEJ or how this might contribute to *VSG* rearrangement. Here, we characterize the features of chromosomal MMEJ in *T. brucei*. We describe robust RAD51-independent, MMEJ-based deletions and also RAD51 dependent, one-sided, MMEJ-based gene conversion. Our findings suggest that these pathways have contributed to the evolution of a compact genome and to the switching of *VSG* gene expression, which typically involves recombination among short, repetitive flanking sequences.

MATERIALS AND METHODS

Trypanosoma brucei growth and manipulation

Lister 427, MITat1.2 (clone 221a), bloodstream form cells were grown in HMI-11 and transformed as described earlier (16). For limiting dilution cloning, cells were distributed over 96-well plates and were analyzed only if <40% of wells displayed cell growth after 5 days. Tetracycline was from Sigma and was used at $1\,\mu g\,ml^{-1}$.

Plasmid construction

Plasmid constructs for expression of the tetracycline repressor from the *TUB* locus (TetR-*BLE*), for tet-on expression of I-*Sce*I with an *N*-terminal SV40 nuclear ocalization signal from a *rRNA* spacer locus (I-*Sce*I-*HYG*) (17) and for integration of the R^{S}P cassette at the Tb11.02.2110 locus (6) were described previously.

To delete the 2110$_b$ allele an RsrII/XcmI fragment in pARD-NEO (16) was replaced with an RsrII/EcoRI fragment from pbRn5 (18) encompassing a portion of the *NPT* gene and an aldolase polyA signal. This pANΔHR construct was digested with SmaI/ApaI prior to transfection and correct integration was confirmed by PCR.

DNA analysis

For Southern blot analysis of DSBR, genomic DNA was digested with HindIII and Bsp120I and processed according to standard protocols. The *2110* probe was a 699 bp, *Sac*I fragment from pARD (16). A series of chromosome 11, *RFP*, *PAC* and *TUB*-specific primers were used to amplify and sequence repair junctions, typically using Taq polymerase in the presence of 1.5% DMSO. Direct sequencing of PCR products was carried out according to standard protocols.

RESULTS

MMEJ dominates repair in a chromosomal end-joining assay

We previously used the I-*Sce*I meganuclease to introduce a single DSB on *T. brucei* chromosome 11a (6); the R^{S}P$_a$ strain used contains a tetracycline-inducible I-*Sce*I gene and a single I-*Sce*I cleavage site adjacent to the Tb11.02.2110 gene (Figure 1A). Using this strain, >50% of cells survived I-*Sce*I induction and repaired the break and ~85% of these survivors underwent allelic HR using the 2110$_b$-allele as a repair template. In addition, two alternatives to allelic HR were revealed; among 26-independent repair events, three clones displayed ectopic HR and two displayed MMEJ (6), an insufficient number for any detailed analysis. To facilitate the isolation of survivors that display end joining and to characterize MMEJ in a chromosomal context, we devised an
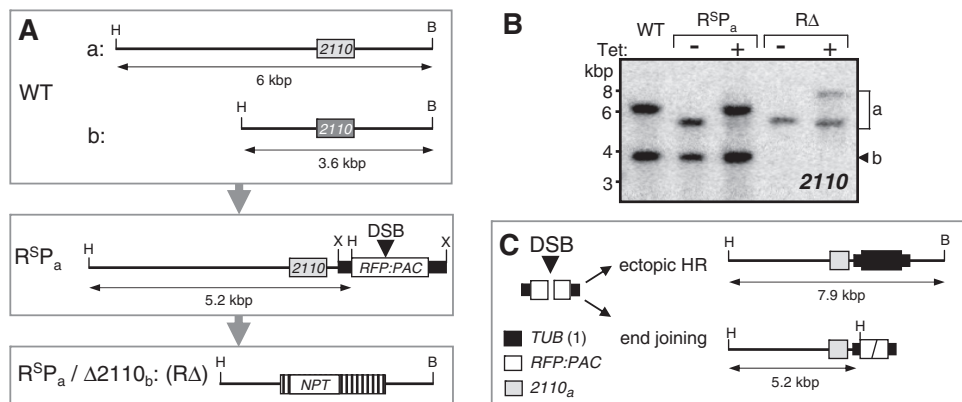
Figure 1. An experimental system to study end-joining. (**A**) The schematic maps illustrate the Tb11.02.2110 alleles in wild-type (WT), R^{S}P$_a$ and RΔ strains. The meganuclease cleavage site is embedded within a dsRed Fluorescent Protein (*RFP*)–Puromycin ACetyltransferase (*PAC*) fusion gene. B, Bsp120I; H, HindIII; X, XcmI. (**B**). The RΔ strain was validated by Southern blotting with WT and R^{S}P$_a$ controls. The RΔ and R^{S}P$_a$ strains were grown in the absence or presence of tetracycline (1 μg ml^{-1}) for 1 week. Genomic DNA was digested with Bsp120I and HindIII. Bands representing the 2110 alleles are indicated to the right. In the R^{S}P$_a$ strain, allelic HR regenerates the 6 kb allele while, in the RΔ strain, ectopic HR and end joining generate allele a fragments at 7.9 and 5.2 kbp, respectively; see fragment sizes in (A) and (C). (**C**) The schematic maps illustrate the result of ectopic HR and end joining expected to predominate in R^{S}P$_a$/Δ2110$_b$ survivors. The *TUB* sequences flanking the R^{S}P cassette promote R^{S}P pre-mRNA trans-splicing and polyadenylation and also allow ectopic HR which replaces R^{S}P with an α*TUB* gene copied from chromosome 1.
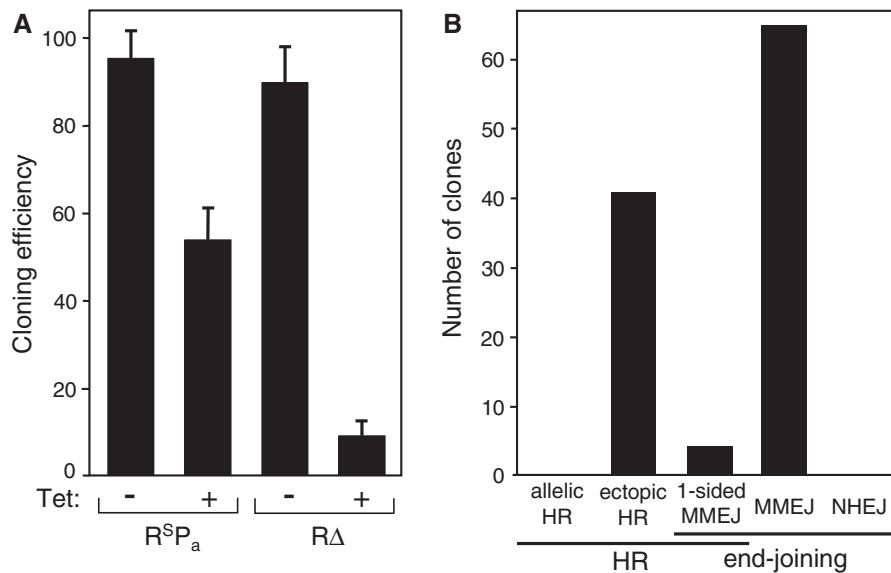
Figure 2. MMEJ is common in RΔ survivors. (**A**) As expected, the RΔ strain displays a reduced cloning efficiency after DSBR due to cell death after allelic HR. Data derived from dilution cloning in 96-well plates: −Tet, $n = 4$; +Tet, $n = 6$. (**B**) RΔ survivors display ectopic HR, one-sided MMEJ and MMEJ-based deletions as determined by DNA sequencing; $n = 110$.

experimental system to eliminate survivors that use the major repair pathway of allelic HR. This was achieved by replacing the 2110_b-allele and other break-adjacent homology with a *NEO* selectable marker (Figure 1A). In the resulting R^{s}P$_a$/$\Delta 2110_b$ (RΔ) strains, allelic HR was expected to cause loss of heterozygocity, loss of the remaining 2110_a allele and, because the encoded *N*-terminal protein acetyltransferase is essential for growth (16), cell death. Disruption of a single *2110* allele had no detectable impact on the growth rate of the RΔ strains (data not shown).

To validate these strains, we devised a Southern blot analysis to distinguish between the *2110* alleles and the three modes of repair seen previously. Genomic DNA was extracted for Southern analysis prior to DSBR and 1 week after meganuclease induction and DSBR. Consistent with previous findings, the control R^{s}P$_a$ strain displayed a modified 2110_a allele prior to DSBR and, after DSBR, survivors displayed dominant allelic HR and reconstitution of the 'wild-type' allele a (Figure 1A and B). In contrast and as expected, the RΔ strain displayed a modified 2110_a allele and no 2110_b allele prior to DSBR and, after DSBR, survivors displayed DNA fragments consistent with ectopic HR and end joining (Figure 1B and C), but no allelic HR. Thus, allelic HR leads to cell death in the RΔ strain and a substantial proportion of survivors display repair via end-joining. A second independent RΔ strain displayed similar results (data not shown) so the first strain was selected for more detailed analysis.

We employed clonogenic assays to determine the proportion of RΔ cells that survive I-*Sce*I-induced lesions (Figure 2A). The R^{s}P$_a$ strain served as a control for this analysis and, consistent with previous findings, indicated >50% survival. In contrast, and consistent with cell death following allelic HR, the RΔ strain displayed <10%

survival (Figure 2A). To distinguish between different repair pathways and to quantify the relative contribution of each pathway in RΔ cells, we derived a panel of survivor clones. This approach was favored over analysis of mixed populations because DNA amplification is required to access the sequence of end-joining junctions (see below) and amplification of multiple, related DNA fragments is prone to 'template-switching' artifacts and amplification bias. In addition, the cloned-survivor approach improves the chance of revealing repair junctions in unanticipated locations since clones that initially fail to reveal a junction can be specifically targeted for further analysis.

We used limiting dilution under I-*Sce*I-inducing conditions to generate a panel of 107 survivor clones. In order to survey repair mechanisms, we prepared genomic DNA from the complete set of survivors. A series of primer pairs, either within or flanking the R^{s}P cassette, was used to amplify repair junctions by PCR that were then sequenced directly from amplified products. The process was iterative, starting with primers closer to the break and progressing to further distal sites until we amplified products from every survivor; three yielded a pair of 'repair fragments' either due to repair in two cells placed in the same well or in both replicated genomes from one cell. Sequencing the set of 110 DNA fragments revealed 65 MMEJ-based deletions (Table 1), 41 cases of ectopic HR, 4 cases of MMEJ-based gene conversion (see below), and no NHEJ (Figure 2B). Thus, end joining was exclusively microhomology mediated in our assay.

Microhomology pairing and junction formation

Our survey yielded 65 MMEJ junction sequences, a data set that provides sufficient information to describe the various characteristics of chromosomal MMEJ in *T. brucei*. Three sequences revealed MMEJ using a perfect 18-bp

Table 1. MMEJ junctions in RΔ survivors

Microhomology (MH) class	Junction type[a]	Survivor(s)	Sum Survivors	RFP Δ	PAC Δ	Total Δ[b]
1	X	1, 13, 30, 33, 49, 54, 60, 66, 76, 83, 86, 87, 88, 91, 93, 96, 98, 103, 104, 106	20	52	20	81
2	X	72	1	165	19	192
3	X/x	68	1	6	224	240
4	X	3, 14[c], 25, 28, 42, 81, 100, 107	8	231	47	291
	xa	10, 17, 29, 31, 63, 64, 92	7	227	57	291
	xb	43, 94, 97, 102	4	217	64	290
	(xa)	58	1	215	71	290
	(xb)	90	1	225	61	291
	Mixed	8[d] (4xa/xb), 56 (4X/xa), 41 [4X/(xb)]	3	217–231	47–64	290–291
5	(x)	34	1	15	288	312
6	(x)	5	1	6	330	341
7	X	27	1	129	327	465
8	X	50	1	430	59	498
9	X	12, 53	2	8	504	523
	(x)	45	1	7	511	522
10	x	20, 74, 78	3	58	454	522
11	X	7, 79	2	454	59	522
12	X	14[c]	1	235	355	598
13	X	48, 82	2	387	217	615
14	X	23	1	379	561	948
XcmI	X	40, 44, 85	3	953	961	1914

[a]Junction types are defined in the legend to Figure 4.
[b]The net loss of bp (Δ) includes 4 bp representing both single-stranded overhangs left after I-*Sce*I cleavage.
[c]Survivor 14 revealed two independent MMEJ events.
[d]The sequence from survivor 8 is not shown in Figure 4 because the 'mixed' junction generates frame-shifted, overlapping traces.

microhomology created during strain assembly. These sequences (gCCAgtccttgtgTGGgt) contain the XcmI sites (upper case characters) used in the assembly of the R^sP construct and are found 953- and 961 bp on either side of the break (Figure 1A). This is the longest and most break-distal microhomology we observed (Table 1 and see below), indicating that at least two factors co-operatively promote MMEJ-associated annealing, length of microhomology and proximity to the break.

For the remaining 62 sequences, we plotted the frequency of joining events that use microhomology blocks within 50 bp intervals along a physical map of the R^sP cassette (Figure 3A). We also illustrate the frequency of each microhomology pairing spanning the DSB (Figure 3B). These data show a strong bias for selection of microhomology closer to the DSB revealing mean deletions of 54 bp on the side closest to the break and 284 bp in total. Deletions ranged from 6- to 561 bp on one side and 81- to 948 bp in total (see Table 1 for full details).

Since template availability, chromatin structure and sequence context could impact the pattern of repair; we next asked whether MMEJ-based repair is similar at different loci across the *T. brucei* genome. A PCR assay using primers flanking the R^sP cassette was used to amplify repaired fragments from a large population of survivors (>50 000). Using this assay, we examined the pattern of repaired fragments in the current R^sP_a and RΔ strains and at three other loci: within the tubulin gene array (*TUB*), within an *rDNA* spacer (19) and within a subtelomeric *VSG* expression site (P_{ES}). A similar banding pattern in all cases (Figure 3C) suggested that MMEJ operates

similarly at multiple genomic loci, and the sequences of eleven MMEJ junctions from P_{ES} survivors were consistent with this interpretation (Supplementary Table S1). The contribution of MMEJ to survival, where known for the R^SP_a, RΔ and P_{ES} strains, was ~5 (6), 59 (Figure 2B) and 23%, respectively, indicating that the PCR assay is quantitative. The fainter banding pattern obtained with the *rDNA* and *TUB* samples therefore indicate that the relative contribution of MMEJ to repair is reduced at these tandem arrayed loci where multiple adjacent and allelic copies likely facilitate HR-based mechanisms.

Analysis of microhomology-pairing in our survey-set revealed 14 distinct classes (numbered MH1-14) with MH1 and MH4 accounting for 70% of all events (Table 1). Microhomologies are typically 5–20 bp in length with several mismatches tolerated in longer stretches (Figure 4A). Inspection of the junction sequences revealed multiple possible outcomes from a single microhomology pairing as exemplified by MH4 which yields five different junctions (Figure 4A). Four other classes revealed a mixture of two possible junctions emerging from a single microhomology pairing (3X/x, 4X/xa, 4X/[xb] in Figure 4A and 4xa/xb in Table 1) presumably due to differential processing on each strand.

We next asked whether base composition influences MMEJ. The *RFP* and *PAC* sequences contain 63 and 73% GC base pairs, respectively and the sum of paired bases from all sequences depicted in Figure 4 is 724 (gray boxes) so we expected 492 GC-pairs (68%) if there is no bias. We observed 543 (75%) GC-pairs which indicates that GC base pairs significantly ($P < 0.0001$) favor productive annealing (Figure 4B); presumably due to
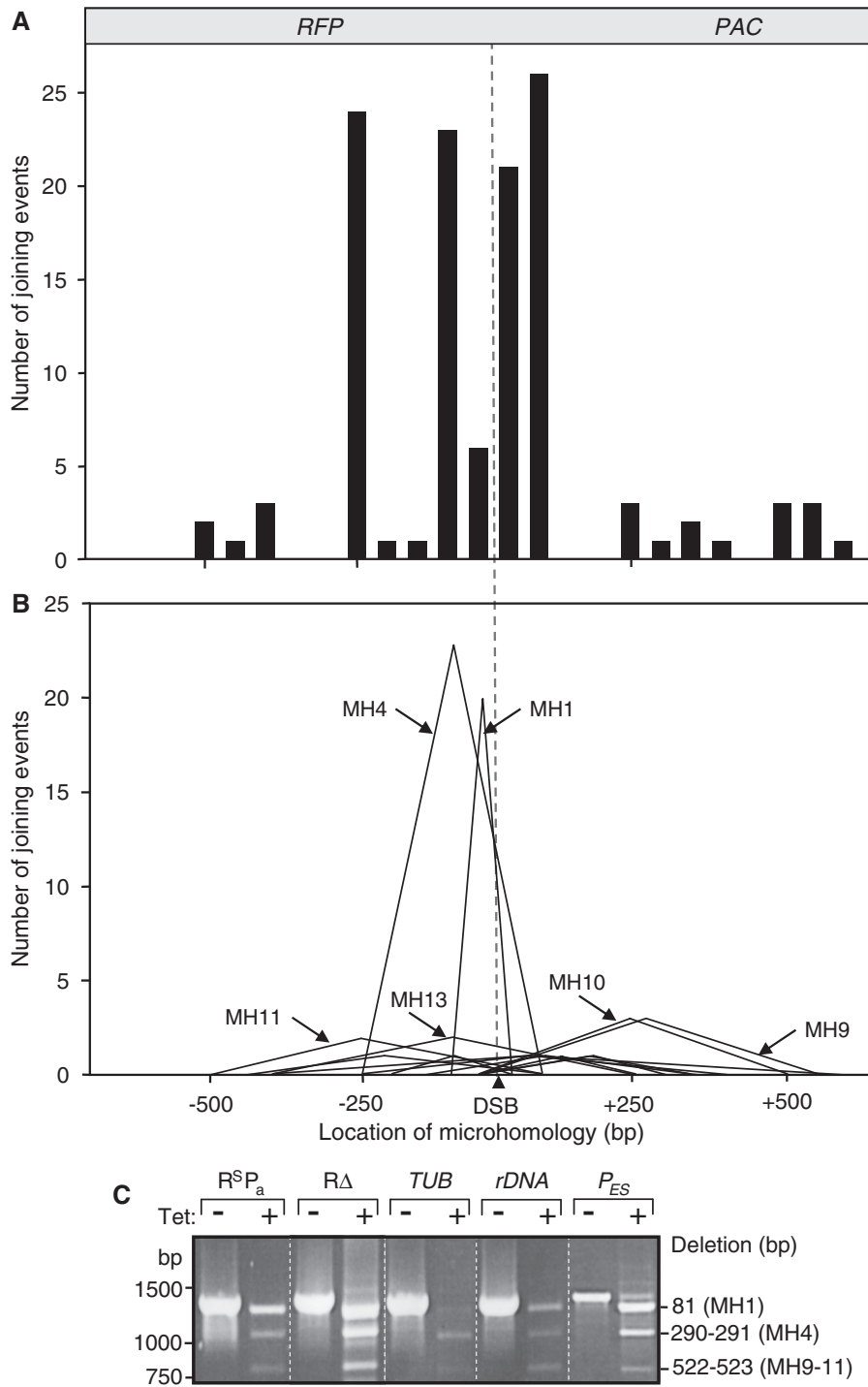
Figure 3. The distribution of microhomologies used for repair. (**A**) Frequency of microhomologies was mapped in 50-bp intervals on either side of the DSB. (**B**) Frequency of paired microhomologies was mapped as in A. All microhomology (MH) classes represented by >1 junction are indicated; see Table 1. (**C**) A PCR assay indicates a similar pattern of MMEJ at different loci across the genome. Products corresponding to frequent MH pairings and size of deletion are indicated; see Table 1.

increased stability. Furthermore, we observed 423/504 (84%) GC-pairs if MH1 repair, favored due to break proximity, was excluded from the analysis (Figure 4B). These results indicate that GC-rich microhomologies and long or break-proximal microhomologies act cooperatively to promote annealing and MMEJ.

One-sided MMEJ-based gene conversion

The junction sequences from four survivor clones displayed a gene conversion-based interchromosomal repair mechanism. This resulted from allelic HR on one side of the R^sP cassette and MMEJ on the other, replacing R^sP with a segment from the *NPT* cassette on chromosome
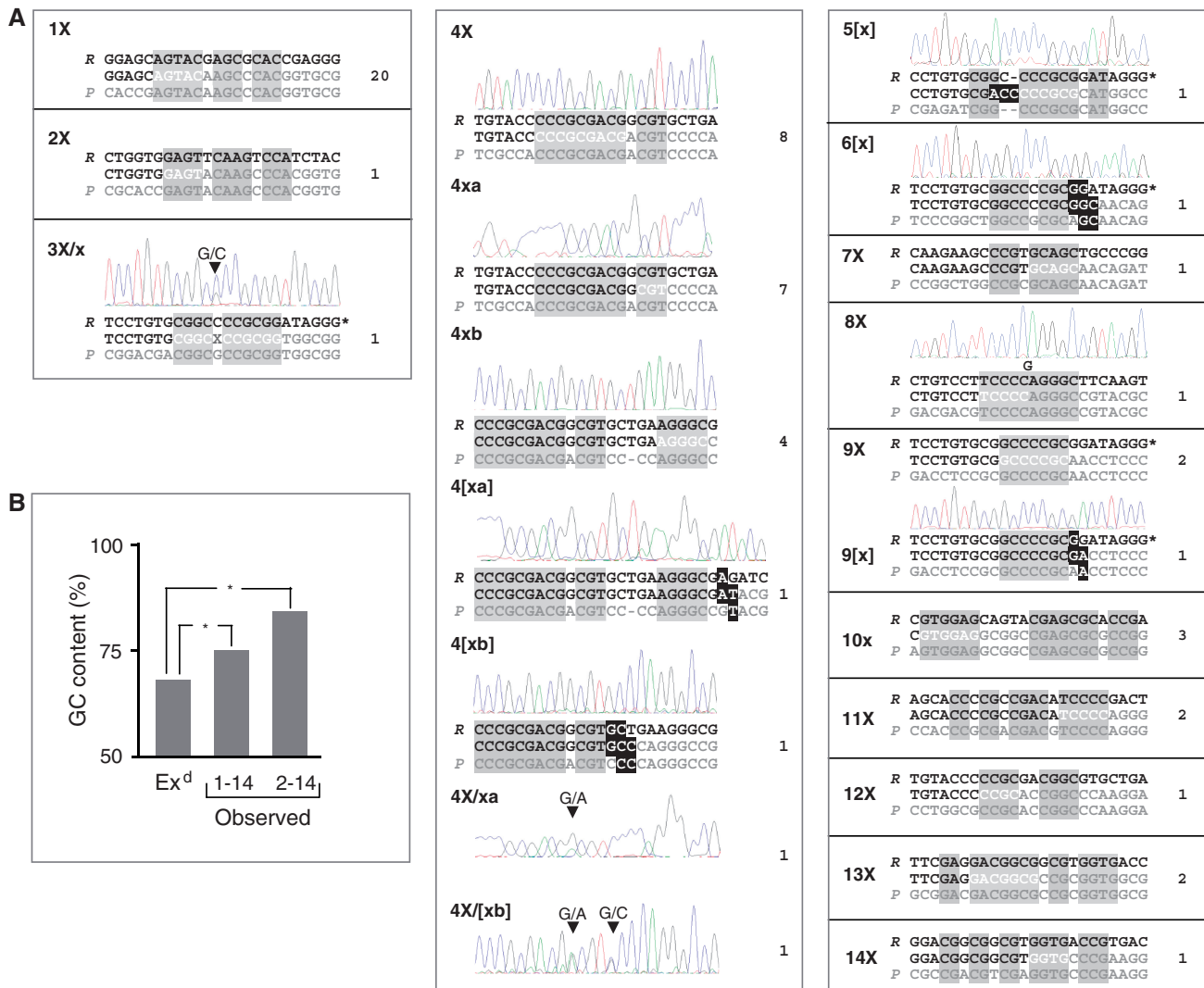
Figure 4. Microhomology classes; see Table 1 for more details. (**A**) MMEJ junctions (sense strand only) are shown for all 14 MH classes identified. The parental *RFP* (*R*) and *PAC* (*P*) sequences are shown above and below, respectively. The template switch site is indicated with white lettering. The junction types are as follows: X, processed within the largest homology patch; x, processed within a 'minor' homology patch; [x] processed outside the homology patches. Microhomologies are highlighted (≥ 3 bp patches plus 2 bp patches if within 1 bp of a larger patch). Sequence traces are shown for [x] junctions, 3X/x, 8X and all MH4 sub-classes (see the text), and the numbers of each sub-class recorded is indicated to the right. Asterisk at MH3, 5, 6 and 9 indicates the I-*Sce*I cleaved terminus. (**B**) GC-content is over-represented within microhomology patches. *$P < 0.0001$ as determined using a χ^2 test.

11b (Figure 5A). A second PCR amplification of fragments spanning each gene conversion tract confirmed the expected size in each case (Figure 5B) and sequencing indicated that clones 4 and 71 used a related microhomology, while clones 24 and 73 used the same microhomology but generated different junctions (Figure 5C).

MMEJ is typically RAD51-independent. To assess the role of RAD51 in chromosomal MMEJ-based deletion and gene conversion, we applied a PCR assay to populations of survivors from $R^s P_a$ strains with wild-type *RAD51* expression or with rad51 disrupted. An assay for MMEJ-based deletions indicated robust activity in the absence of RAD51 (Figure 6A, upper panel), and sequencing of nine rad51 survivors revealed exclusively MMEJ-based deletions (data not shown). In contrast, one-sided MMEJ-based gene conversion was specifically ablated in the rad51 null strain (Figure 6A, lower panel). Sequencing confirmed that the major products detected using this assay in wild-type RAD51 cells both represented one-sided MMEJ (Figure 6B and C). Thus, one-sided MMEJ-based gene conversion can use allelic or ectopic homology on chromosome 1. The RAD51-requirement suggests that gene conversion is initiated by HR within tubulin sequence and resolved by MMEJ. We also analyzed an $R^s P_a$ survivor for which we previously failed to identify a repair mechanism [see Figure 5B, lane 4 in (6)] and this survivor was also found to have arisen through one-sided ectopic MMEJ (Figure 6B and C, iii). Segments copied from chromosome 1 in these RAD51-dependent, one-sided MMEJ reactions ranged from 28 to 1084 bp.

Figure 5. One-sided MMEJ-based gene conversion. (**A**) The schematic map illustrates four allelic one-sided MMEJ events; the gray box indicates HR and the lines indicate the locations of microhomologies, in the aldolase (*ALD*) processing sequence (cross-hatched box) linked to the *NPT* gene in this case. In clones 24 and 73, the terminal 21 codons of the *2110* gene are replaced with 11 new codons and a stop codon; presumably allowing the expression of a functional protein. (**B**) PCR amplification of fragments spanning the gene conversion tract reveals a product of the expected size in each case; wild-type (wt) cells serve as a control. The locations of the primers are indicated in A (small arrows). (**C**) The four one-sided allelic MMEJ junction sequences are shown; other details as in Figure 4A.

DISCUSSION

In SSA, repeated sequences flanking a DSB are able to anneal once single-stranded regions are exposed. Resolution requires processing by digestion of the single-stranded tails and gap filling (9). MMEJ is also known as micro-SSA and our results provide insight into the equivalent processing steps in *T. brucei* MMEJ. For example, the MH4-type events we report indicate removal of single-stranded tails at five alternative locations. In this case, cleavage or digestion of the single-stranded tail is most common within a 9 or 3 bp microhomology patch. Comparison with MH1-type events, where processing is exclusively within the longer microhomology patch, suggests that a nearby, out-of-register 5 bp patch of microhomology (Figure 4A, 4xb) distorts the paired sequence and promotes processing at different locations. In addition, four junction sequences revealed a mixture of two possible outcomes (3X/x, 4X/xa, 4X/[xb] and 4xa/xb), which suggests staggered processing

of the single-stranded tails and the maintenance of mismatched base pairs until the next DNA replication cycle.

A DSB initiates a homology search which, if successful, results in HR. This search for homology is typically successful at diploid loci in *T. brucei* while non-conservative MMEJ generates a deletion. Using our experimental system, ∼5% of cells survive DSBR via MMEJ at a single copy diploid locus and this is increased to ∼60% in the RΔ strains where allelic HR is lethal. A bias towards deletions associated with DSBR by end joining in trypanosomatids may have made a major contribution to genome evolution. Specifically, MMEJ-mediated deletions, with a mean size of 284 bp in our assays, may be responsible for the remarkably compact organization of trypanosomatid genomes, which are comprised of ∼50% protein-coding sequence. MMEJ may also be responsible for many of the synteny gaps found in trypanosomatid chromosomes which often reflect loss of redundant genes in a lineage (20).

Interestingly, at a subtelomeric *VSG* expression site, ∼25% of cells survive DSBR using MMEJ-based repair. This increased contribution to DSBR when compared to other loci could reflect reduced competition with allelic HR or could equally be explained by another feature of subtelomeric chromatin. To further explore the contribution of microhomology to DNA rearrangement at subtelomeric loci, we examined the available *VSG* expression site sequences which are thought to reflect host–parasite interactions (21). Remarkably, we identified three examples of putative microhomology-based deletion (Supplementary Figure S1). In two cases, putative parental sequences were found at multiple subtelomeres and, consistent with frequent recombination, the putative derived sequences were also found at multiple subtelomeres. Retention of the junction sequences suggests that these deletions occurred relatively recently. At only 12–18 bp though, the deleted segments are short relative to those described above. We also noted that larger deletions involving the loss of ESAGs could have resulted from SSA mediated by flanking homology; ESAG3 (pseudo)genes in the case of ESAG4/8 deletion in BES12 and BES14 for example.

Approximately 30 copies of the tubulin sequence are present in tandem on chromosome 1 and the presence of tubulin gene segments flanking our reporter cassette allowed for ectopic HR between chromosomes 1 and 11. This revealed 41 (37%) such repair events when allelic HR was eliminated and clearly demonstrates the capacity to search genome wide for suitable homologous templates for DSBR in *T. brucei*. In addition, a subset of these interactions initiated one-sided MMEJ. These robust ectopic recombination pathways may reflect mechanisms that are important for subtelomeric interactions that underpin *VSG* recombination and antigenic variation.

So what contribution does MMEJ make to antigenic variation? A switch in *VSG* expression involves replacement of the previously active *VSG* at the single active telomeric ES (13) and this proceeds via RAD51 dependent and independent mechanisms (22) that are not fully understood. This is typically an interchromosomal gene conversion process that relies upon imperfect homology within

Figure 6. One-sided MMEJ-based gene conversion is RAD51 dependent. (**A**) PCR assays indicate a similar pattern of MMEJ-based deletion in RAD51 and rad51 null strains (upper panel) while one-sided MMEJ-based gene conversion is ablated in rad51 null strains (lower panel). The locations of the primers are indicated in B. (**B**) The schematic map illustrates three ectopic one-sided MMEJ events. (i) and (ii) are from Figure 6A and (iii) is from (6); other details as in Figure 5A. (**C**) The three one-sided ectopic MMEJ junction sequences are shown; other details as in Figure 4A.

repetitive sequences upstream of *VSG* genes (7); '70 bp' repeats that are widely distributed among subtelomeres. Break-induced replication (BIR), typically a RAD51 dependent process in *Saccharomyces cerevisiae* (23), has been proposed as a mechanism of telomere conversion based *VSG* switching (7). Recent work in *T. brucei* indicates both RAD51-dependent and -independent BIR with suppression of the RAD51-dependent pathway by TOPO3α (24). Since we demonstrate RAD51 dependent and independent pathways, MMEJ is an excellent candidate for mediating 70 bp repeat recombination. Gene conversion tracts involving BIR or two-sided recombination could be initiated, terminated or both by MMEJ and this could explain why antigenic variation is relatively insensitive to regulation by mismatch repair (25,26). It is also interesting to note in this respect that translocations involved in class switch recombination in B cells (2), as well as replication fork breakage-induced rearrangements in human cells (27), may use a micro-BIR pathway. MMEJ-based equivalents of the three major HR mechanisms, gene conversion, BIR and SSA, and one-sided gene conversion are all possible and we describe the latter two, micro-SSA and one-sided gene conversion, in a chromosomal context in *T. brucei*. Thus, the gene conversions we describe could reflect important pathways of *VSG* rearrangement and antigenic variation. However, it will be challenging to distinguish between the use of microhomology or longer tracts of homology following recombination among highly repetitive *T. brucei* sequences.

MMEJ is considered a backup end-joining mechanism in cells where NHEJ operates. Unusually in *T. brucei*, MMEJ dominates end joining (6), but little is known

about this repair mechanism in trypanosomes. We developed strains to facilitate studies on chromosomal MMEJ in *T. brucei* and show that proximity to the break, number and proportion of matched bases and GC-content all promote pairing and MMEJ-based deletions. In addition, we show that one-sided MMEJ can mediate interchromosomal gene conversion.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

FUNDING

REFERENCES

1. Lieber,M.R. (2010) The mechanism of double-strand DNA break repair by the nonclassical DNA end-joining pathway. *Annu. Rev. Biochem.*, **79**, 181–211.
2. Yan,C.T., Boboila,C., Souza,E.K., Franco,S., Hickernell,T.R., Murphy,M., Gumaste,S., Geyer,M., Zarrin,A.A., Manis,J.P. *et al.* (2007) IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature*, **449**, 478–482.
3. Zhu,C., Mills,K.D., Ferguson,D.O., Lee,C., Manis,J., Fleming,J., Gao,Y., Morton,C.C. and Alt,F.W. (2002) Unrepaired DNA

breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell*, **109**, 811–821.

4. McVey,M. and Lee,S.E. (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.*, **24**, 529–538.

5. Nussenzweig,A. and Nussenzweig,M.C. (2007) A backup DNA repair pathway moves to the forefront. *Cell*, **131**, 223–225.

6. Glover,L., McCulloch,R. and Horn,D. (2008) Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Res.*, **36**, 2608–2618.

7. Boothroyd,C.E., Dreesen,O., Leonova,T., Ly,K.I., Figueiredo,L.M., Cross,G.A. and Papavasiliou,F.N. (2009) A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. *Nature*, **459**, 278–281.

8. Decottignies,A. (2007) Microhomology-mediated end joining in fission yeast is repressed by pku70 and relies on genes involved in homologous recombination. *Genetics*, **176**, 1403–1415.

9. Haber,J.E. (2006) Transpositions and translocations induced by site-specific double-strand breaks in budding yeast. *DNA Repair*, **5**, 998–1009.

10. Lee,K. and Lee,S.E. (2007) *Saccharomyces cerevisiae* Sae2- and Tel1-dependent single-strand DNA formation at DNA break promotes microhomology-mediated end joining. *Genetics*, **176**, 2003–2014.

11. Liang,L., Deng,L., Nguyen,S.C., Zhao,X., Maulion,C.D., Shao,C. and Tischfield,J.A. (2008) Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Res.*, **36**, 3297–3310.

12. Yu,A.M. and McVey,M. (2010) Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.*, **38**, 5706–5717.

13. 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.

14. 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.

15. 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.

16. Ingram,A.K., Cross,G.A. and Horn,D. (2000) Genetic manipulation indicates that *ARD1* is an essential N^a-acetyltransferase in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **111**, 309–317.

17. 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.

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

19. Glover,L. and Horn,D. (2009) Site-specific DNA double-strand breaks greatly increase stable transformation efficiency in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **166**, 194–197.

20. El-Sayed,N.M., Myler,P.J., Blandin,G., Berriman,M., Crabtree,J., Aggarwal,G., Caler,E., Renauld,H., Worthey,E.A., Hertz-Fowler,C. *et al.* (2005) Comparative genomics of trypanosomatid parasitic protozoa. *Science*, **309**, 404–409.

21. Hertz-Fowler,C., Figueiredo,L.M., Quail,M.A., Becker,M., Jackson,A., Bason,N., Brooks,K., Churcher,C., Fahkro,S., Goodhead,I. *et al.* (2008) Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS ONE*, **3**, e3527.

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

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

24. Kim,H.S. and Cross,G.A. (2010) TOPO3a influences antigenic variation by monitoring expression-site-associated *VSG* switching in *Trypanosoma brucei*. *PLoS Pathog.*, **6**, e1000992.

25. Barnes,R.L. and McCulloch,R. (2007) *Trypanosoma brucei* homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases. *Nucleic Acids Res.*, **35**, 3478–3493.

26. 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.

27. Hastings,P.J., Ira,G. and Lupski,J.R. (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.*, **5**, e1000327.