Opinion TRENDS in Parasitology Vol.20 No.3 March 2004



# Antigenic variation in African trypanosomes: monitoring progress

**Richard McCulloch**

Wellcome Centre for Molecular Parasitology, University of Glasgow, Anderson College, 56 Dumbarton Road, Glasgow, G11 6NU, UK

**Antigenic variation is central to the success of African trypanosomes and other eukaryotic, bacterial and viral pathogens. Our understanding of the control and execution of this immune evasion strategy in trypanosomes is incomplete, despite the molecular basis of antigenic variation being first described over 20 years ago. Recent research progress in this field is highlighted here and some of the unresolved questions raised.**

African trypanosomes, such as *Trypanosoma brucei*, grow in the bloodstream and tissue fluids of mammals, and are subjected to specific and non-specific immune attack [1]. To hide invariant surface antigens from the immune response, trypanosomes shroud themselves in a variant surface glycoprotein (VSG) coat whose molecular identity they periodically change, allowing a part of the parasite population to avoid antibody-mediated killing throughout an infection. To ensure expression of a single *VSG* in one cell at one time, multiple telomeric sites of transcription, termed *VSG* expression sites (ES), have evolved and only one is actively transcribed at a time. In addition, large numbers of silent *VSG* genes have been dispersed around the genome. The number of ES and silent *VSG*s appears to amount to tens and hundreds of copies, respectively. The exact figure will soon be catalogued by the *T. brucei* genome sequencing initiative, which is conducted jointly between the Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/) and the Institute for Genome Research (http://www.tigr.org). However, even from the limited data published to date [2,3], we can say something about how this extensive *VSG* repertoire is accommodated and used.

The *T. brucei* genome comprises 11 diploid megabase-sized chromosomes, a set of intermediate-sized chromosomes and a large number of mini-chromosomes [4,5]. All classes of chromosome contain *VSG*s. Arrays of silent *VSG*s are present in the interior of the megabase chromosomes and towards at least one chromosome's telomere, where the majority of the *VSG*s are pseudogenes [3]. Whether or not all chromosome-internal *VSG*s are present within arrays, and the proportion of them that are intact, will explain how the gene repertoire expanded and diversified, and is activated during antigenic variation. It appears that most, but perhaps not all [3], *T. brucei* telomeres contain *VSG*s [6]. ESs appear to be found exclusively at the ends of the megabase and intermediate chromosomes, whereas stand-alone, silent *VSG*s are present at the telomeres of the minichromosomes. Within the mammal, *T. brucei* uses a class of ES called bloodstream ESs (BESs) which have complex structures (Figure 1), containing expression site-associated genes (ESAGs) that are co-transcribed with the *VSG*. Comparing six complete or partial BES [1,7] shows that their general structure is conserved, with the *VSG* proximal to the telomere and many kilobases from the promoter; the number of intervening ESAGs and their potential to encode functional gene products is, however, variable. *Trypanosoma brucei* is transmitted between mammals by tsetse. In the tsetse gut, the VSG coat is shed but, within the salivary glands, non-dividing, metacyclic form trypanosomes reacquire a VSG coat, selected from a repertoire of telomeric metacyclic ESs (MES). MES contain only a *VSG* coding sequence, but the presence of upstream ESAGs and pseudogenes suggests that the two types of ES are evolutionarily related [8,9].

Because of the division of the *VSG* repertoire between sites of transcription and silent non-transcribed loci, two

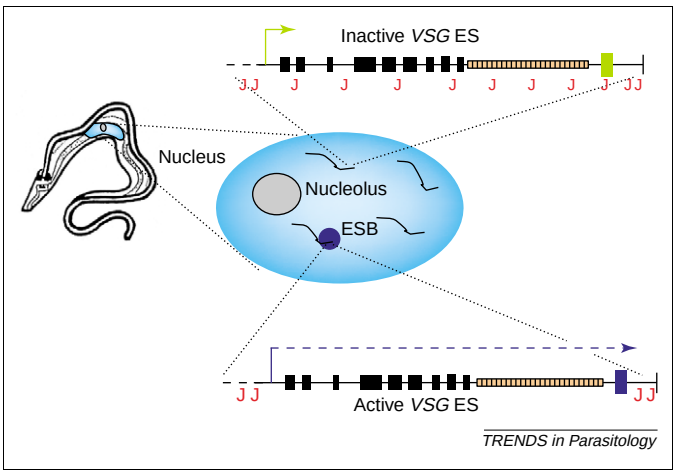


**Figure 1**. Variant surface glycoprotein transcriptional control and the expression site body. The long, slender replicative form of *Trypanosoma brucei*, found in the mammalian bloodstream, is depicted on the left. The insert represents the nucleus, containing the nucleolus and a distinct subnuclear structure termed the expression site body (ESB) in which variant surface glycoprotein (*VSG*) gene expression occurs. For simplicity, only four *T. brucei* chromosomes are depicted in the nucleus, and the inserts depict two telomeric sites of *VSG* transcription, termed expression sites (ES). The *VSG* (blue or green box) is always found adjacent to the telomere (vertical line), and is co-transcribed with several expression site-associated genes (black boxes) and flanked upstream by an array of degenerate 70 bp repeats (orange box). The inactive BES (top) contains the modified base J (indicated in red) throughout its coding sequence, as well as in sequences at the telomere and upstream of the BES promoter (green arrow). The active BES is present within the ESB, and transcription (blue broken line) occurs over all protein-coding sequences, where J is undetectable.

*Corresponding author:* Richard McCulloch (rmc9z@udcf.gla.ac.uk).

 1471-4922/$ - see front matter  doi:10.1016/j.pt.2003.12.004

processes can account for a switch in the expressed *VSG*. In one, active transcription of the BES encoding the VSG coat is suppressed, and a silent BES becomes actively transcribed. The second, more common, process involves recombination reactions that move silent *VSG*s into the BES. This is achieved often by gene conversion reactions where a copy is generated from a silent *VSG* and moved to the ES, deleting the resident *VSG* and leaving the silent *VSG* substrate genetically unaltered. This can occur directly only on the intact *VSG*s, not on the *VSG* pseudogenes. A substantial proportion of the *VSG* repertoire being pseudogenes would mean that a process called mosaic *VSG* formation, hitherto considered a rare reaction [10], would assume greater importance.

### *VSG* transcriptional control uses a subnuclear body

Many organisms contain gene families from which they express just one gene copy at a time or in a given cell. This has been termed allelic exclusion or monoallelic expression [11]. The discovery of a subnuclear body in *T. brucei* that contains the actively transcribed BES sheds light upon how this is controlled [12] (Figure 1). The *VSG* ESs are unusual in containing protein-coding genes, but are transcribed by RNA polymerase I (Pol I) [13], which normally transcribes an organism's ribosomal RNA genes. Navarro and Gull [12], using elegant *in situ* hybridisation analyses, demonstrated that the active BES is present within a RNA Pol I-containing body in a subnuclear location distinct from the nucleolus. Silent BESs are found outside this expression site body (ESB), which appears to be a discrete proteinaceous structure because it survives the removal of DNA by DNAase I.

The discovery of the ESB leads to several questions regarding its architecture and origin. The ESB and the nucleolus are both sites of Pol I transcription, but at least one nucleolar protein, fibrillarin, is absent from the ESB [12]. This could indicate that the two structures are rather distinct, which is an interesting possibility, or that the ESB arose as an evolutionary adaptation of the nucleolus. Navarro and Gull [12] suggest, on the basis of a lack of Pol I-extranucleolar staining, that the ESB is absent from *T. brucei* procyclic cells, a life cycle stage from the midgut of tsetse. Whether the ESB is reconstructed or reactivated as *T. brucei* infects a mammal or in later tsetse life cycle stages is not known. The broader evolutionary significance of the ESB is also worth considering: will related structures be the basis for monoallelic expression in other organisms? Antigenic variation driven by transcriptional switches also occurs in *Plasmodium* [14] and *Giardia* [15], although here Pol II probably drives the expression. Several examples of monoallelic expression from gene families occur in mammals [16] but, again, there could be a distinction with *T. brucei*, because when a gene is selected for expression or repression, this decision is permanent through further cell divisions.

The ESB imposes a new context within which we must consider how trypanosome VSG switching occurs. This body elegantly explains how a single BES is actively transcribed and the remainder are silenced: the factors for BES transcription are sequestered to the ESB and are absent from the nucleoplasm, and hence unavailable to the silent sites. Nevertheless, what are these factors and what parts of the BES are held within the body? BESs are large, around 40–60 kb, and we do not know if all or part of them is contained within the ESB. One might imagine that it is sufficient for the ESB to sequester just the BES promoter and the initiation factors. This seems unlikely, however, because at least some of the silent ES are partially transcribed [17]. In addition, ES promoters can direct gene expression when experimentally removed from their genomic context (e.g. onto extra-chromosomal plasmids [18]) and a ribosomal RNA promoter can functionally replace the ES promoter [19]. This argument would also suggest that there is no structural impediment to the (untested) possibility that the ESB also controls MES transcription in metacyclic form *T. brucei* (MES and BES promoters, despite both being transcribed by Pol I, are dissimilar in sequence [20,21]). Perhaps, this favors the idea suggested by Vanhamme *et al*. [22] that the active BES alone contains the elongation and processing factors that drive transcription over its entire length. Indeed, the ESB might be the site of the 'RNA factory' that these authors propose [22]. What is less clear is how a transcriptional switch from one BES to another occurs. To understand this, we must determine: (i) what factors mediate BES entry into, and hold it within, the ESB; (ii) what changes occur during a transcriptional switch; and (iii) when this occurs in the cell cycle. For example, is the ESB transiently unoccupied during a switch, but rapidly reacquires a BES? Alternatively, do two ESBs transiently arise, but are then reduced to one, or does the ESB transiently become able to contain two BES, but then resolves to a single BES-containing body? Chaves *et al*. [23] have come closest to answering this by identifying an unstable transcriptional switch intermediate where two BES, but not three [24], are close together in the nucleus and both express *VSG* mRNA and protein, implying that they are fully transcribed. Clearly, much remains to be learned, but the presence of a body for exclusive BES expression, and the coupling of transcription activation and inactivation seems to signal the death knell (at least for the BES) for stochastic models of transcriptional switching based on the proposed spread of heterochromatin from the *T. brucei* telomeres [25].

### Genetic controls of VSG switching by recombination

Homologous recombination is used in all organisms to repair damage to their genome, which can arise in many ways [26]. It might make sense to predict that trypanosomes have invested, during their evolution, in a specialized form of recombination to guide the movement of *VSG*s into the ES. This would be desirable to avoid potentially deleterious genetic changes that might arise accidentally during recombination of *VSG*s, which are dispersed throughout the genome. Surprisingly, the evidence that VSG switching simply exploits general pathways of homologous recombination repair now appears overwhelming. We have known for some time that the amount of *VSG* sequence copied during gene conversion reactions (Figure 2) extends for variable distances to conserved DNA sequences both upstream and downstream of the *VSG* ORF [27] (Figure 2). This bears the hallmarks of

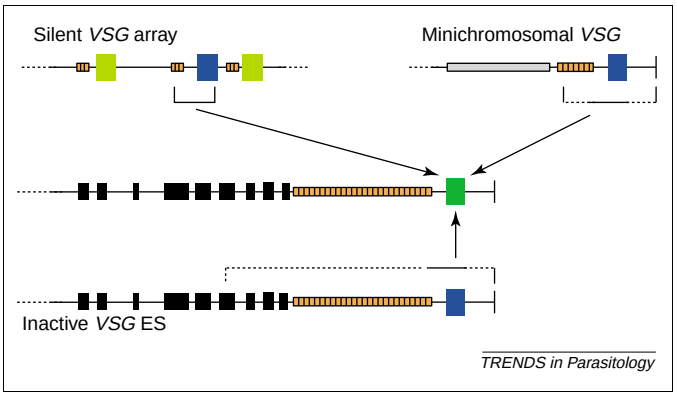


**Figure 2**. Variant surface glycoprotein switching by gene conversion. Switching of the expressed variant surface glycoprotein (*VSG*) by gene conversion involves copying a silent *VSG* (blue box) into the BES (see Figure 1), replacing the resident *VSG*. The silent *VSG* donor can be present in: (i) an array in the interior of a megabase chromosome (other *VSG*s are represented by pale green boxes); and (ii) at the telomere of a minichromosome (which contains a 177 bp repeated sequence; grey box); or (iii) in an inactive BES (only the bloodstream stage version is depicted). The amount of sequence copied during gene conversion is illustrated, and normally encompasses the *VSG* ORF and extends upstream to the 70 bp repeats (orange box), which appear to be uniquely associated with *VSG*s and play a role in recombination. The extent of converted sequence can vary (depicted by a broken line): the upstream conversion limit can extend to different parts of the 70 bp repeats, or into the conserved expression site-associated genes. *VSG* gene conversions encompassing the downstream repeats that make up the telomere have been also described.

homologous recombination, which is largely driven by the conserved DNA strand exchange enzyme RAD51 in eukaryotes [28]. Mutants of the *T. brucei* gene that encodes the best RAD51 candidate characterized to date are impaired in their ability to execute VSG switching [29], providing genetic evidence to support the molecular analyses of the reaction. In concordance with this, mutants of KU70/80, which are central to a competing recombination pathway, termed non-homologous recombination (NHEJ), that is known to mediate some transposition reactions [30], do not affect VSG switching [31].

## Emerging complications

Examination of the role of MRE11 in trypanosome antigenic variation is not in keeping with that of RAD51. MRE11 is part of a trimeric enzyme complex with multiple proposed roles in both homologous recombination and NHEJ [32]. One role is in processing DNA double-strand breaks to make the ends recombinogenic, but MRE11 is also implicated in sensing and signalling DNA damage, bridging the ends of broken DNA molecules and telomere regulation. Homozygous mutants of *MRE11* demonstrate the importance of this gene in maintaining *T. brucei* genome integrity because they result in increased rates of chromosomal rearrangements [33]. These rearrangements are subtly different from those described previously in *Saccharomyces cerevisiae*: loss of chromosome-internal sequences was found in *T. brucei*, but not the inter-chromosome translocations and telomere deletions as seen in yeast. *MRE11* mutants display no detectable impairment in VSG switching, suggesting that the protein does not act with RAD51 in this process. Both genes do, however, influence homologous recombination because mutation of either leads to reduced rates of DNA integration into the *T. brucei* genome following transformation [33,34]. Together, these data indicate our incomplete understanding of trypanosome recombination and switching. Could it be that the initiating event in VSG switching, which remains unknown, somehow bypasses the roles of MRE11? Alternatively, might homologous recombination in *T. brucei* be slightly different from other organisms to accommodate the requirement for VSG switching? One hint of this can be seen when we look at DNA damage responses. *Trypanosoma brucei MRE11* mutants show little evidence for increased sensitivity to methyl methanesulfonate, as observed in other organisms: they do, however, display increased sensitivity to other DNA damaging agents.

*RAD51* remains, thus far, the only gene to be genetically implicated in regulating or catalyzing VSG switching. Despite this, *RAD51* mutants can still undergo antigenic variation by recombination [29]. Moreover, DNA transformation experiments show that at least some of the residual recombination activity in the absence of RAD51 looks like robust homologous recombination [35]. Other recombination reactions are revealed by this approach that fall between what we consider to be homologous recombination and NHEJ, in that they use short ($\sim$7–13 bp) regions of imperfect sequence homology. RAD51-independent recombination reactions, some using short regions of homology, have been described in bacteria and yeast [36,37] but, even in these organisms, the catalytic factors are mysterious, therefore, identifying them in trypanosomes poses a challenge. Moreover, we do not know what these further pathways might reveal about VSG switching. For instance, in what circumstances might the back-up recombination pathways contribute to VSG switching or *VSG* gene reassortment, and might the sequestration of the ES into the ESB impose constraints upon *T. brucei* recombination that these pathways help to overcome? In addition, all the above genetic analyses have been conducted in laboratory-adapted *T. brucei* strains that undergo antigenic variation at substantially lower rates than other, less-adapted strains [38]. The genetic basis of this difference remains to be determined.

## Does base J play a role in antigenic variation?

Trypanosomes contain a modified base called β-D-glucosyl(hydroxymethyl)uracil, or J, which replaces a fraction of thymine in the genome. J has been known for some time to have an interesting relationship with antigenic variation: it is detectable within the silent BES, but not within the active BES or in silent chromosome-internal VSGs [39]. J can be found in repeated sequences around the active BES (the telomeres and upstream 50 bp repeats), but not within the transcribed region (Figure 1). The impact of J on antigenic variation appeared to have been somewhat tarnished by the discovery that it is also present in other repeats around the *T. brucei* genome, and is found throughout the Kinetoplastida, as well as in *Diplonema* and *Euglena* [40], organisms that do not undergo antigenic variation. Definitive understanding of J's function has been hampered by difficulties in ablating it from the genome. Identification of the J biosynthetic enzymes has proved problematic, despite a good working model for a two-step pathway involving oxidation of thymine to

hydroxymethyluracil, followed by glucosylation [41,42]. van Leeuwen *et al*. [41] used chemical routes to reduce and increase the J content around tenfold, both of which caused increased gene expression from a silent ES, and an increase or decrease, respectively, in chromosome breaks around an ES. Identification of a J-specific binding protein [43] seemed likely to yield a clear indication of function, but mutants of the gene encoding this factor retain 5% of J and it appears that the protein promotes J maintenance or propagation [44].

Should we conclude that J plays no role in trypanosome antigenic variation? The answer must be no because its function remains elusive. The fact that J is present in other organisms does not preclude a purpose that serves VSG switching, and allowed trypanosomes to co-opt it into that role. Most work has suggested that J is probably bound by proteins that lead to altered chromatin states, thus changing transcription or recombination rates [44]. This is possible, but we might also consider that J serves to repel the interaction of some undefined factors. Could J mark the silent BESs and prevent their uptake into the ESB, or might the ESB reverse or block J biosynthesis on a BES to allow its transcription? Both J and the ESB are present in bloodstream-stage *T. brucei*, but both become undetectable in procyclic cells [12,45]. J might also directly repel recombination factors, which could stabilize repeat tracts [41]. In terms of antigenic variation, the presence of J in silent BESs could direct recombination towards the transcribed BES, where VSG switching would have an immediate effect. In this regard, it is noteworthy that no comparison of recombination rates at genomic sites with and without J has been reported.

**Acknowledgements**

The number of references cited in this article is limited, and I apologize to those people whose important contributions have not been identified. I am grateful to Dave Barry, Chris Proudfoot and Rebecca Barnes for critical reading of the manuscript. Work in my laboratory is supported by the Royal Society and Wellcome Trust.

**References**

1 Donelson, J.E. (2003) Antigenic variation and the African trypanosome genome. *Acta Trop.* 85, 391–404
2 Hall, N. *et al.* (2003) The DNA sequence of chromosome I of an African trypanosome: gene content, chromosome organisation, recombination and polymorphism. *Nucleic Acids Res.* 31, 4864–4873
3 El Sayed, N.M. *et al.* (2003) The sequence and analysis of *Trypanosoma brucei* chromosome II. *Nucleic Acids Res.* 31, 4856–4863
4 Melville, S.E. *et al.* (1998) The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* and the assignment of chromosome markers. *Mol. Biochem. Parasitol.* 94, 155–173
5 Ersfeld, K. *et al.* (1999) Nuclear and genome organization of *Trypanosoma brucei*. *Parasitol. Today* 15, 58–63
6 Barry, J.D. *et al.* (2003) Why are parasite contingency genes often associated with telomeres? *Int. J. Parasitol.* 33, 29–45
7 Berriman, M. *et al.* (2002) The architecture of variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 122, 131–140
8 Graham, S.V. *et al.* (1999) A structural and transcription pattern for variant surface glycoprotein gene expression sites used in metacyclic stage *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 103, 141–154
9 LaCount, D.J. *et al.* (2001) Analysis of a donor gene region for a variant surface glycoprotein and its expression site in African trypanosomes. *Nucleic Acids Res.* 29, 2012–2019
10 Thon, G. *et al.* (1990) Trypanosome variable surface glycoproteins: composite genes and order of expression. *Genes Dev.* 4, 1374–1383
11 Borst, P. (2002) Antigenic variation and allelic exclusion. *Cell* 109, 5–8
12 Navarro, M. and Gull, K. (2001) A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. *Nature* 414, 759–763
13 Gunzl, A. *et al.* (2003) RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryot. Cell* 2, 542–551
14 Duffy, M.F. *et al.* (2003) Regulation of antigenic variation in *Plasmodium falciparum*: censoring freedom of expression? *Trends Parasitol.* 19, 121–124
15 Nash, T.E. (2002) Surface antigenic variation in *Giardia lamblia*. *Mol. Microbiol.* 45, 585–590
16 Ohlsson, R. *et al.* (1998) Monoallelic expression: 'there can only be one'. *Trends Genet.* 14, 435–438
17 Vanhamme, L. *et al.* (2000) Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. *Mol. Microbiol.* 36, 328–340
18 Vanhamme, L. *et al.* (2001) Control and function of the bloodstream variant surface glycoprotein expression sites in *Trypanosoma brucei*. *Int. J. Parasitol.* 31, 523–531
19 Rudenko, G. *et al.* (1995) A ribosomal DNA promoter replacing the promoter of a telomeric VSG gene expression site can be efficiently switched on and off in *T. brucei*. *Cell* 83, 547–553
20 Kim, K.S. and Donelson, J.E. (1997) Co-duplication of a variant surface glycoprotein gene and its promoter to an expression site in African trypanosomes. *J. Biol. Chem.* 272, 24637–24645
21 Ginger, M.L. *et al.* (2002) *Ex vivo* and *in vitro* identification of a consensus promoter for VSG genes expressed by metacyclic-stage trypanosomes in the tsetse fly. *Eukaryot. Cell* 1, 1000–1009
22 Vanhamme, L. *et al.* (2001) An update on antigenic variation in African trypanosomes. *Trends Parasitol.* 17, 338–343
23 Chaves, I. *et al.* (1999) Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *EMBO J.* 18, 4846–4855
24 Ulbert, S. *et al.* (2002) Expression site activation in *Trypanosoma brucei* with three marked variant surface glycoprotein gene expression sites. *Mol. Biochem. Parasitol.* 120, 225–235
25 Borst, P. and Ulbert, S. (2001) Control of VSG gene expression sites. *Mol. Biochem. Parasitol.* 114, 17–27
26 van Gent, D.C. *et al.* (2001) Chromosomal stability and the DNA double-stranded break connection. *Nat. Rev. Genet.* 2, 196–206
27 Borst, P. and Cross, G.A. (1982) Molecular basis for trypanosome antigenic variation. *Cell* 29, 291–303
28 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
29 McCulloch, R. and Barry, J.D. (1999) A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes Dev.* 13, 2875–2888
30 Di Fagagna, F.F. *et al.* (2003) The Gam protein of bacteriophage Mu is an orthologue of eukaryotic Ku. *EMBO Rep.* 4, 47–52
31 Conway, C. *et al.* (2002) Ku is important for telomere maintenance, but not for differential expression of telomeric VSG genes, in African trypanosomes. *J. Biol. Chem.* 277, 21269–21277
32 D'Amours, D. and Jackson, S.P. (2002) The Mre11 complex: at the crossroads of DNA repair and checkpoint signalling. *Nat. Rev. Mol. Cell Biol.* 3, 317–327
33 Robinson, N.P. *et al.* (2002) Inactivation of Mre11 does not affect VSG gene duplication mediated by homologous recombination in *Trypanosoma brucei*. *J. Biol. Chem.* 277, 26185–26193
34 Tan, K.S. *et al.* (2002) *Trypanosoma brucei* MRE11 is non-essential but influences growth, homologous recombination and DNA double-strand break repair. *Mol. Biochem. Parasitol.* 125, 11–21
35 Conway, C. *et al.* (2002) Two pathways of homologous recombination in *Trypanosoma brucei*. *Mol. Microbiol.* 45, 1687–1700
36 Paques, F. and Haber, J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol. Mol. Biol. Rev.* 63, 349–404
37 Lovett, S.T. *et al.* (2002) Crossing over between regions of limited homology in *Escherichia coli*. RecA-dependent and RecA-independent pathways. *Genetics* 160, 851–859

38 Barry, J.D. and McCulloch, R. (2001) Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. *Adv. Parasitol.* 49, 1–70
39 van Leeuwen, F. *et al.* (1997) Localization of the modified base J in telomeric VSG gene expression sites of *Trypanosoma brucei*. *Genes Dev.* 11, 3232–3241
40 Dooijes, D. *et al.* (2000) Base J originally found in kinetoplastida is also a minor constituent of nuclear DNA of *Euglena gracilis*. *Nucleic Acids Res.* 28, 3017–3021
41 van Leeuwen, F. *et al.* (1998) Biosynthesis and function of the modified DNA base beta-D-glucosyl- hydroxymethyluracil in *Trypanosoma brucei*. *Mol. Cell. Biol.* 18, 5643–5651
42 Ulbert, S. *et al.* (2002) Expression of the human DNA glycosylase hSMUG1 in *Trypanosoma brucei* causes DNA damage and interferes with J biosynthesis. *Nucleic Acids Res.* 30, 3919–3926
43 Cross, M. *et al.* (1999) The modified base J is the target for a novel DNA-binding protein in kinetoplastid protozoans. *EMBO J.* 18, 6573–6581
44 Cross, M. *et al.* (2002) J-binding protein increases the level and retention of the unusual base J in trypanosome DNA. *Mol. Microbiol.* 46, 37–47
45 Gommers-Ampt, J. *et al.* (1991) A novel DNA nucleotide in *Trypanosoma brucei* only present in the mammalian phase of the life-cycle. *Nucleic Acids Res.* 19, 1745–1751