4

African trypanosomes: the genome and adaptations for immune evasion

Gloria Rudenko[1]

Division of Cell and Molecular Biology, Sir Alexander Fleming Building, Imperial College London, Imperial College Road, South Kensington, London SW7 2AZ, U.K.

Abstract

The African trypanosome *Trypanosoma brucei* is a flagellated unicellular parasite transmitted by tsetse flies that causes African sleeping sickness in sub-Saharan Africa. Trypanosomes are highly adapted for life in the hostile environment of the mammalian bloodstream, and have various adaptations to their cell biology that facilitate immune evasion. These include a specialized morphology, with most nutrient uptake occurring in the privileged location of the flagellar pocket. In addition, trypanosomes show extremely high rates of recycling of a protective VSG (variant surface glycoprotein) coat, whereby host antibodies are stripped off of the VSG before it is re-used. VSG recycling therefore functions as a mechanism for cleaning the VSG coat, allowing trypanosomes to survive in low titres of anti-VSG antibodies. Lastly, *T. brucei* has developed an extremely sophisticated strategy of antigenic variation of its VSG coat allowing it to evade host antibodies. A single trypanosome has more than 1500 *VSG* genes, most of which are located in extensive silent arrays. Strikingly, most of these silent *VSG*s are pseudogenes, and we are still in the process of trying to understand how non-intact *VSG*s are recombined to produce genes encoding functional coats. Only one *VSG*

[1]email gloria.rudenko@imperial.ac.uk

is expressed at a time from one of approximately 15 telomeric *VSG* ES (expression site) transcription units. It is becoming increasingly clear that chromatin remodelling must play a critical role in ES control. Hopefully, a better understanding of these unique trypanosome adaptations will eventually allow us to disrupt their ability to multiply in the mammalian bloodstream.

Introduction

The African trypanosome *Trypanosoma brucei* is a unicellular flagellated protozoan parasite that is transmitted between mammalian hosts by tsetse flies, causing African sleeping sickness in sub-Saharan Africa. *T. brucei* has an unusually broad host range for a parasite and infects a varied range of African mammals in addition to humans and their livestock [1,2]. Very little is known about the specific adaptations of African trypanosomes that allow them to survive in such a diverse group of mammalian species. A striking feature of African trypanosomes is their ability to maintain chronic infections as extracellular parasites, rather than remaining hidden from the immune system within host cells as do the American trypanosome *Trypanosoma cruzi* (the causative agent of Chagas disease) or parasites such as *Leishmania* (causing leishmaniasis) or *Plasmodium* (the malaria parasite). While multiplying in the mammalian bloodstream, trypanosomes are exposed to continuous immune attack both by host antibodies and the complement system. As a consequence, *T. brucei* has developed a highly sophisticated strategy of antigenic variation of a protective VSG (variant surface glycoprotein) coat [3,4]. A host infected with a given VSG variant of trypanosomes eventually raises antibodies specific to that VSG, which leads to clearance of all cells that are recognized. However, as new VSG switch variants are continuously generated within a given trypanosome population, these (temporarily) escape recognition by antibodies and can prolong the infection. The trypanosome therefore continuously remains one step ahead of the host. In addition to this highly elaborate antigenic variation of VSG, *T. brucei* has developed numerous additional adaptations to its cell biology allowing it to flourish in the hostile niche of the mammalian bloodstream.

Trypanosome cell biology and immune evasion

VSG is very antigenic, and the trypanosome is only able to fully escape the consequences of host anti-VSG antibodies through switching between different VSG variants. However, in order to take up nutrients, the trypanosome needs to express invariant receptors and transporters on its cell surface. Various adaptations of trypanosome cell biology allow the cell to minimize exposure of these invariant surface proteins to host antibodies.

Trypanosomes propel themselves through blood using a single flagellum [5] (Figure 1), and flagellar motility is essential for bloodstream-form *T. brucei* both *in vitro* and *in vivo,* but is not essential in insect-form trypanosomes
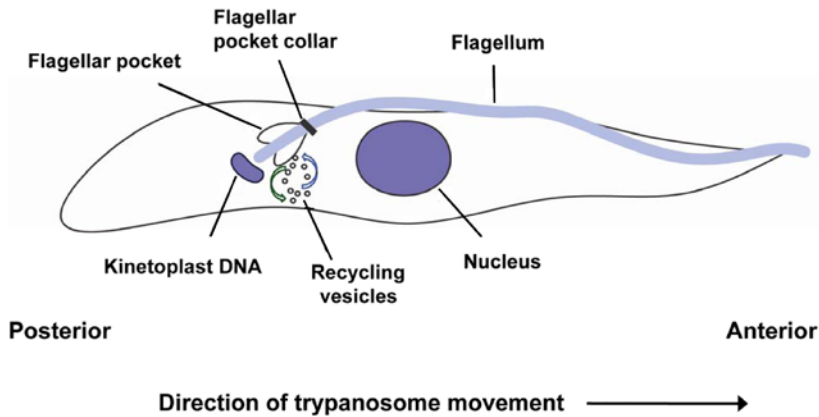
Figure 1. Cross-section of a trypanosome with key structures indicated
The nuclear and kinetoplast DNA are indicated as filled ovals. The flagellum extends from the flagellar pocket, which is where endocytosis, exocytosis and VSG recycling occur (indicated by recycling vesicles). The entire bloodstream-form trypanosome surface (including the flagellum) is coated with a dense homogeneous layer of VSG.

[6,7]. The trypanosome flagellum extends from a bulb-like invagination of the surface membrane known as a flagellar pocket [8,9]. This privileged location is the site for trypanosome nutrient uptake, and is where all endocytosis and exocytosis occurs [10]. The flagellar pocket forms a collar (which could be similar to a tight junction) around the exit point of the flagellum. This restriction presumably functions as a 'cuff', preventing easy access of host immune molecules such as antibodies and complement to the inside of the pocket [11]. Recently, a channel in the collar of the flagellar pocket has been described, which connects the inside of the flagellar pocket with the outside of the cell and is thought to facilitate the movement of fluid into the pocket [12]. Invariant trypanosome surface receptors within the flagellar pocket are therefore hidden from host antibodies. Relatively few invariant molecules appear to be present on the *T. brucei* outer cell surface, and these are possibly shielded from recognition by host antibodies by the rod-like VSG molecules [13].

Key for *T. brucei* survival in the bloodstream is this dense layer of 10 million VSG molecules, which are attached to the surface membrane through a GPI (glycosylphosphatidylinositol) anchor [14] and coat the entire surface of the cell, including the flagellum [13]. VSG is essential in bloodstream-form *T. brucei* even *in vitro*, and blocking its synthesis triggers a very rapid and precise precytokinesis cell-cycle arrest [15]. This argues for the presence of a cell-cycle checkpoint whereby VSG synthesis is 'sensed' during the trypanosome cell cycle, and progression is halted in the absence of adequate VSG to coat the daughter cells. In cells arrested after blocking VSG synthesis, a concurrent block in global protein synthesis is also triggered [16]. This translation arrest possibly functions to prevent the arrested trypanosome from growing in size, and thereby further diluting its protective VSG coat.

A striking feature of this protective VSG surface is that, despite the fact that VSG is packed at an extremely high (and possibly maximal) density on the trypanosome surface, the coat is highly fluid. VSG is recycled at an unusually high rate, and it has been estimated that the entire pool of VSG is turned over every 12 minutes [17]. This very high level of VSG recycling appears to serve as a coat 'self-cleaning' mechanism in which host molecules, such as antibodies, are stripped from internalized VSG before it is sent back to the cell surface. In addition to these very high rates of recycling, trypanosome motility also appears to play a role in facilitating the removal of VSG–antibody complexes. It has been proposed that hydrodynamic forces operate on the surface of a swimming trypanosome, which push VSG in a posterior direction [18] (Figure 2). These forces could result in VSG–antibody complexes within the VSG layer being preferentially swept back towards the flagellar pocket where they can be internalized, with the antibodies functioning analogously to 'sails'. In addition to the observed high rates of VSG recycling,
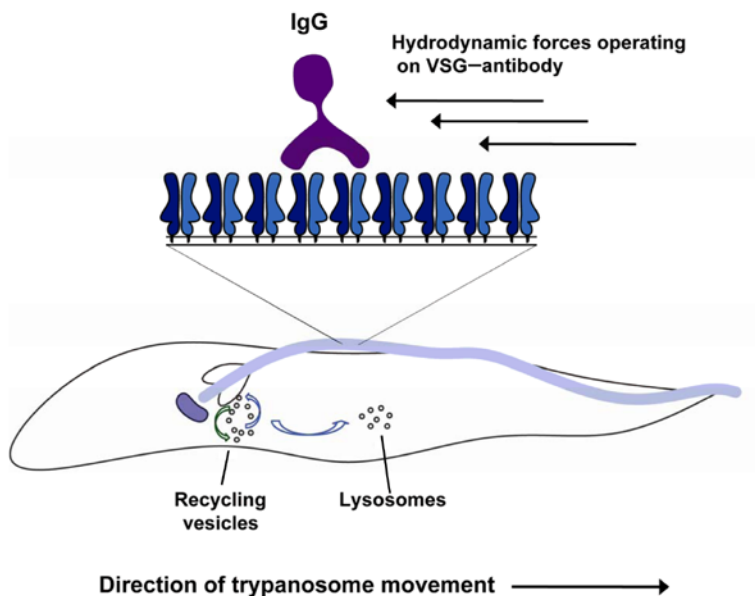

Figure 2. Hydrodynamic forces operate on the surface coat of a swimming trypanosome

A cross-section of a trypanosome is indicated as in Figure 1. A cross-section of the VSG coat is indicated in the upper panel, with the VSG dimers attached to the lipid bilayer through GPI anchors. A schematic anti-VSG IgG antibody molecule is indicated at the top. As the trypanosome swims forward, hydrodynamic forces operating on the surface coat result in VSG–antibody complexes being preferentially swept backwards towards the flagellar pocket [18]. Anti-VSG IgG antibody is then stripped from the VSG in the lysosomes. Reprinted with modifications from *Cell*, vol. 131, Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N. and Overath, P., Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes, pp. 505–515, © 2007, with permission from Elsevier.

this preferential uptake of VSG–antibody complexes could be another adaptation of trypanosome cell biology that could allow a bloodstream-form trypanosome to survive for longer in a host, even in the presence of anti-VSG antibodies. Despite these delaying tactics, eventually the titre of anti-VSG antibodies rises to the point where the trypanosomes are invariably overwhelmed, and only those that have switched to another VSG can survive.

Changing the active VSG

A single trypanosome contains an extensive repertoire of more than 1500 *VSG* genes and pseudogenes encoding different VSG variants [19]. Only one of these many *VSG*s is expressed at a time, and is transcribed from one of approximately 15 telomeric *VSG* ES (expression site) transcription units [20]. *T. brucei* can switch its active VSG in different ways (Figure 3). The simplest mechanism is transcriptional activation of a new ES and silencing of the old one (Figure 3D). However, this VSG switching event is only effective early in an infection, as this mechanism accesses the limited pool of the approximately 15 ES-located *VSG*s [20]. Alternatively, the trypanosome can
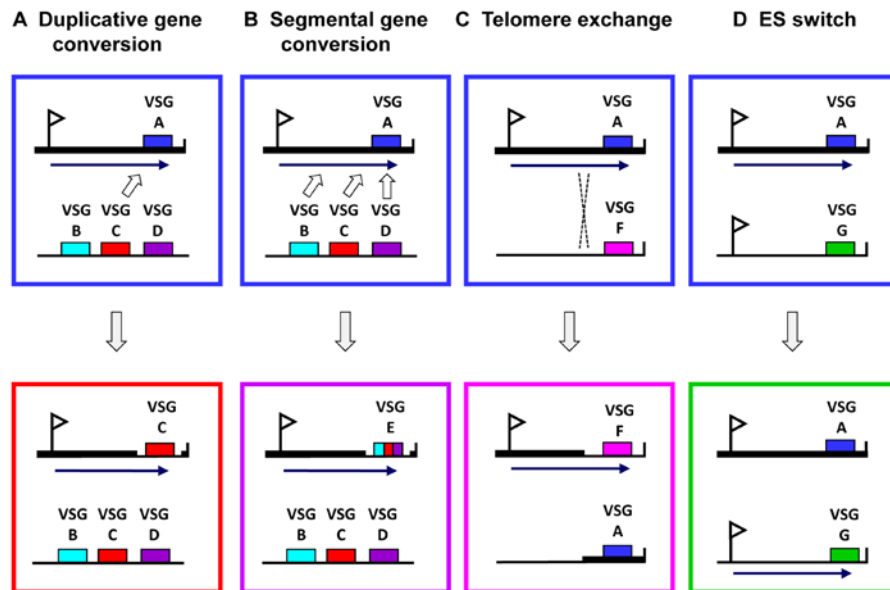

Figure 3. Main mechanisms of VSG switching in *T. brucei*
The large open boxes indicate trypanosomes, and the small filled boxes indicate *VSG*s. The active *VSG* is transcribed from a telomeric *VSG* ES, with the ES promoter indicated by a flag, and transcription by an arrow. (**A**) Most important during a chronic infection is duplicative gene conversion, whereby the trypanosome can copy virtually any *VSG* into the active ES. (**B**) Segmental gene conversion is the same process, except that small regions of *VSG* are gene-converted into the active ES sometimes from multiple *VSG*s. This process allows the trypanosome to use *VSG* pseudogenes to produce a new chimaeric *VSG*. (**C**) A VSG switch mediated by a telomere exchange involves DNA recombination on two telomeres (indicated by a cross). (**D**) An ES switch involves the transcriptional activation of a new ES and the silencing of the old one.

use DNA rearrangements including gene conversion and telomere exchange to exchange the active *VSG* [3] (Figures 3A–3C). During a switch mediated by gene conversion, a silent *VSG* is copied into the active ES and replaces the old *VSG* (Figure 3A). After a telomere exchange, there is no loss of DNA, and a cross-over occurs on two telomere ends, resulting in a previously inactive telomeric *VSG* to be inserted into the active ES (Figure 3C). Gene conversion is by far the most important mechanism for *VSG* switching during a chronic infection, as it enables potentially any *VSG* in the genome to be copied into the active ES, replacing the old *VSG*.

Relatively little is known about the DNA recombination machinery mediating *VSG* switching. The variable *VSG* cassettes are invariably flanked upstream by characteristic 70 bp simple sequence repeats. Silent *VSG*s located in the *VSG* basic copy arrays or silent mini-chromosomes normally have several 70 bp repeats located upstream of each gene. These 70 bp repeats also extend for typically tens of kilobases upstream of the telomeric *VSG*s located in ESs [21]. *VSG*s also invariably contain conserved 3′ regions located within the *VSG*s themselves. These two regions of homology around the *VSG* (70 bp repeats upstream and the 3′ conserved region downstream) are key for facilitating gene conversion between dissimilar *VSG*s, and faciliate *VSG* duplication through gene conversion into the active ES.

Three *T. brucei* proteins involved in homologous recombination have been shown to play a role in VSG switching: RAD51, BRCA2 and RAD51-3 [22–24]. However, knockouts of these proteins do not completely inhibit *VSG* switching, indicating a certain degree of redundancy. Proteins involved in homologous recombination that are exclusively involved in the DNA rearrangement events necessary for antigenic variation still need to be identified. It had been considered likely that DNA double-strand break repair is involved in *VSG* switching, and there is now experimental evidence supporting this: this elegant approach used a *T. brucei* strain where a heterologous recognition sequence for the yeast endonuclease I-SceI was introduced next to the 70 bp repeats of the active ES [25]. Expression of the exogenous yeast endonuclease I-SceI within these trypanosomes was subsequently induced, and *VSG* switching was monitored. Introduction of a double-strand DNA break adjacent to the 70 bp repeats of the active ES dramatically increased rates of *VSG* switching approximately 250-fold. The authors were also able to detect naturally occurring double-strand DNA breaks within the 70 bp repeats of the actively transcribed ES. This indicates that repair of these DNA breaks may be a natural step during *VSG* switching. The 70 bp sequence itself appeared to facilitate *VSG* switching in these experiments, and could possibly be the target of a sequence-specific endonuclease in *T. brucei*, making it more prone to double-strand DNA breaks than other regions of the genome. Unlike in other organisms, double-strand DNA break repair in *T. brucei* appears to primarily involve microhomology-mediated DNA joining rather than non-homologous end joining [26]. This specific repair pathway might facilitate the gene conversion events involved in *VSG* switching.

A detailed discussion of the mechanisms of DNA recombination that facilitate *VSG* switching is outside the scope of this particular chapter, but this subject has been reviewed recently in [4].

The VSG repertoire

The huge *VSG* repertoire is silent with the exception of the *VSG* located in the active ES. While some *VSGs* are found in silent ESs, silent or basic copy *VSGs* are also present in extensive subtelomeric arrays as aneuploid single-copy regions attached to diploid chromosome cores [27] (Figure 4). Although the *VSG* repertoires of different *T. brucei* strains contain many common *VSG* variants [28], there is nonetheless strain-specific divergence of these *VSG* repertoires, presumably facilitating superinfection of a host by multiple strains of trypanosomes [29,30]. For antigenic variation to work, it is of course crucial that the silent *VSGs* are not transcribed. However, it is not known what 'barriers' prevent transcription of these huge arrays of silent *VSGs* via read-through transcription from neighbouring RNAPII (RNA polymerase II) transcription units. In addition, it is still unclear how the *VSG* ESs are transcriptionally controlled.

A highly surprising observation from analysis of the *T. brucei* genome sequence is that the vast majority of these silent *VSGs* appear to be pseudogenes, with only an estimated 5–15% predicted to be fully functional [19,31,32]. This obviously has very important implications for antigenic variation, and the textbook description of intact *VSG* cassettes being moved around is clearly too simplistic. Instead, it is evident that segmental gene conversion
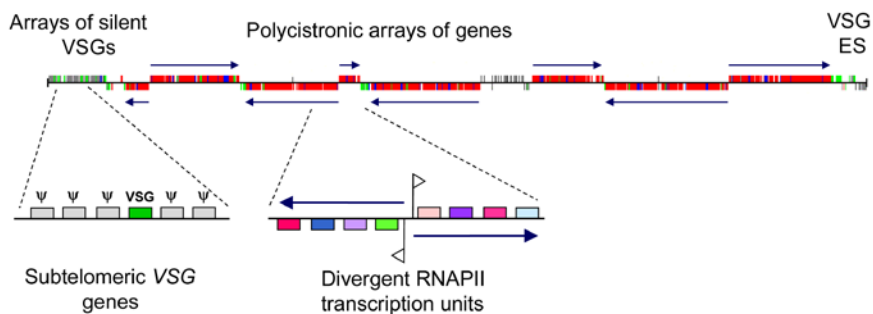

Figure 4. The genome of *T. brucei* is organized in polycistronic arrays of genes, with large subtelomeric arrays of silent *VSG* genes
A schematic representation of a single chromosome is shown in the upper section; genes are indicated with red lines and the direction of transcription is shown with blue arrows. Various genomic regions are expanded in the lower section. Extensive arrays of silent *VSGs* are located at many subtelomeres, with most of the *VSGs* present as pseudogenes (ψ) (indicated by light grey boxes). Chromosome internal genes are organized as polycistronic units transcribed from RNAPII promoters (white flags). *VSG* ESs are also located at some telomeres (see Figure 5). Modified from *Essays in Biochemistry*, vol. 48, Rudenko, G., Epigenetics and transcriptional control in African trypanosomes, pp. 201–219, © 2010 Biochemical Society.

must play a critical role in trypanosome antigenic variation (Figure 3B). In addition, it now appears that the sequence composition of the *VSG* repertoire might facilitate this particular switching event. A comprehensive analysis of a large proportion of the *VSG* repertoire of the *T. brucei* 927 genome strain has revealed that approximately 60% of *VSG*s within the repertoire are unique, with the rest present in very small families of two to four close homologues [32]. This structure of the *VSG* repertoire could facilitate segmental gene conversion events leading to the creation of mosaic *VSG*s.

Segmental gene conversion of *VSG*s was documented decades ago, and is particularly well analysed in *Trypanosoma equiperdum*. Here, *VSG* switch events were described involving bits and pieces of different *VSG*s including *VSG* pseudogenes, which appeared to have been 'sewn' together through multiple gene conversion events producing new mosaic *VSG*s [33]. However, it is now clear that, rather than being a mechanistic curiosity, segmental gene conversion must be key in allowing the African trypanosome to mount chronic infections lasting for years using a *VSG* repertoire that is predominantly 'broken'. This need to continuously create new *VSG*s might provide an advantage born from necessity. The process of creating new mosaic *VSG*s from pseudogenes allows the trypanosome to endlessly recombine its genetic information in different permutations and combinations. Rather than simply having 1500 VSG coats, *T. brucei* has the potential to create an almost infinite number of VSG coat variants.

The unusual architecture of the trypanosome genome

The *T. brucei* genome is organized in an unorthodox fashion for a eukaryote. There is very little transcriptional control, and most genes are located in extremely large polycistronic arrays that are constitutively transcribed by RNAPII [31] (Figure 4). In general, most *T. brucei* genes have functionally unrelated neighbours, and there is little evidence for the presence of many co-regulated regulons [31,34]. The polycistronic precursor transcripts are co-transcriptionally processed into mature mRNA transcripts via coupled reactions of *trans*-splicing of a capped spliced leader RNA with polyadenylation [35]. Almost 6% of *T. brucei* genes are differentially expressed between bloodstream- and insect-form trypanosomes; however, this differential expression does not appear to be regulated by life-cycle-specific *trans*-splicing [36]. Instead, differential RNA stability mediated through the UTRs (untranslated regions) of the transcripts is thought to play a major role in regulating gene expression [37].

The promoters of these extensive RNAPII transcription units remain undefined. However, polyguanine stretches have been identified at strand-switch regions, i.e. regions that contain two transcription units pointing in opposite directions. These GC-rich regions are likely to be important for RNAPII promoter activity [38]. In agreement with the observed lack of RNAPII transcriptional control, the *T. brucei* genome contains very few

homologues for RNAPII basal transcription factors found in other eukary-otes [39]. However, epigenetic marks could play a role in the regulation of this polycistronically transcribed genome, and different histone modifications and histone variants have been mapped to putative *T. brucei* transcription start and termination sites [38,40]. In an extensive study using ChIP-seq (chro-matin immunoprecipitation and sequencing), the Cross laboratory [38] has mapped the histone modifications H4K10ac and H3K4me3 and the histone variants H2AZ and H2BV to RNAPII transcription start sites. The histone variants H3V and H4V were mapped to RNAPII transcription termin-ation sites. These data have led to the speculation that these different epigenetic marks rather than transcription factor binding could demarcate *T. brucei* tran-scription units [41]

An additional epigenetic feature, the modified DNA base called 'J' has been found to be preferentially located at these RNAPII strand-switch regions [42]. Base J is an unusual glucosylated base (β-D-glucopyranosyloxymethyl-uracil) found in kinetoplastid protozoa, where it is invariably present at telo-meres [43]. In *T. brucei*, base J was originally identified at telomeres containing silent *VSG*s [44], but is also present in a variety of DNA repeat sequences [45]. The mechanistic significance of the distribution of J at regions containing puta-tive RNAPII transcription initiation and termination sites is still unclear [42].

VSG expression sites

One of the main exceptions to the rule of constitutive polycistronic transcription concerns the genes involved in *T. brucei* antigenic variation. ESs are polycistronic transcription units containing a variety of different *ESAG*s (expression site-associated genes) in addition to the telomeric *VSG* (Figure 5). ESs could therefore be considered as containing 'kits' of *ESAG*s encoding proteins facilitating life in different species of mammalian host [46]. For example, *ESAG6* and *ESAG7* encode polymorphic transferrin receptor subunits involved in uptake of mammalian transferrin [47]. In addition, ESAG5 shows similarity to a family of lipid transfer/lipopolysaccharide-binding proteins including BPI (bactericidal permeability-increasing protein) [48]. Many of these *ESAG* families have large numbers of similar but non-identical gene copies in different genomic locations, and members are located in chromosome internal positions as well as in ESs. As *ESAG*s present in the active ES are transcribed at a high rate by RNAPI (RNA polymerase I), this allows for particularly high levels of expression of the *ESAG* variants located in the active ES. As the polymorphic *ESAG*s present in different ESs differ in sequence, switching between ESs possibly allows the trypanosome to modulate exactly which variant of a particular ESAG receptor is expressed at any time. Elucidating the function of more of these *ESAG* gene families will hopefully allow us to better understand why *ESAG*s are present in multiple genomic locations, including the telomeric ESs.
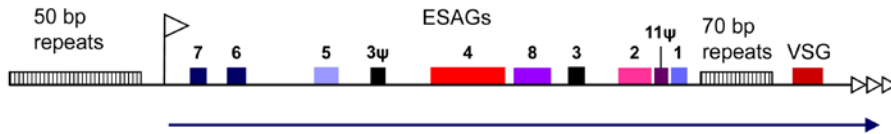
Figure 5. Schematic diagram of a typical *VSG* ES
The polycistronic ES transcription unit contains different *ESAGs* as well as the telomeric *VSG*. Members of these ES gene families are indicated by filled boxes, and the telomere repeats are indicated by horizontal arrowheads. The ES promoter is indicated by an open flag, and transcription is indicated by an arrow. Characteristic 50 bp or 70 bp simple sequence repeat arrays (vertically striped boxes) as well as pseudogenes (ψ) are also indicated.

Analysis of the genetic diversity of these telomeric ES transcription units has been complicated by the difficulties posed with cloning chromosome ends. Using a modified technique of TAR (transformation-associated recombination) cloning adapted to resolve this problem, the repertoire of ESs from the *T. brucei brucei* 427 laboratory strain was cloned in yeast [49]. Sequence analysis of this complete repertoire of ESs showed that they display a surprisingly conserved architecture considering the recombinogenicity of telomere ends [20]. TAR cloning was also used to analyse the genetic diversity of ESs from *T. brucei gambiense*, which causes disease in humans, and the horse pathogen *Trypanosoma equiperdum*, in addition to ESs from *T. brucei brucei* [50]. Despite the proposed role of *ESAG*s in facilitating adaptation of *T. brucei* for life in different species of mammalian host, no immediate correlation was observed between the size of the trypanosome host range and *ESAG* sequence diversity.

A striking feature of ES transcription is that it is mediated by RNAPI rather than RNAPII [51]. In addition to the active ES, the major surface protein of procyclic *T. brucei* (procyclin) is also transcribed by RNAPI [52]. Trypanosomes are the only eukaryotes known to transcribe a subset of their protein-coding genes using RNAPI, which normally exclusively transcribes rDNA (ribosomal DNA) [51,53]. As RNAPI transcription units appear to be transcribed at a particularly high rate [53], trypanosomes appear to have recruited their strongest promoters to direct expression of their surface-coat-encoding genes. For example, VSG, the most abundant protein in the cell (at approximately 10% of total protein), is encoded by a single active *VSG* gene. It is likely that all aspects of VSG expression from its transcription, to the RNA processing, the stability of the *VSG* transcript and its translation into protein are optimized for maximal expression.

Epigenetics and *VSG* expression site control

For antigenic variation to work, and to ensure that the trypanosome does not run through its repertoire of VSG coats too quickly, it is vital that only a single *VSG* is expressed at a time. *VSG* ESs are therefore expressed in a strictly mono-allelic fashion [54,55]. It is still unclear how the ES 'counting'

mechanism operates, and which epigenetic marks differentiate the one active ES from the many silent ones. The glucosylated base J is one epigenetic mark that is enriched on silent ESs, where it is distributed in a gradient that increases towards the telomere [56]. J is not only present in silent ESs, but also enriched in the 50 bp simple sequence repeats extending for tens of kilobases upstream of all known ESs [45]. In addition, J is enriched in the telomere repeats of both active and silent ESs. However, blocking synthesis of J appears to have a relatively minor effect on disrupting ES silencing [44]. One possible function for J could be in suppressing DNA recombination, but this issue is still under investigation.

What is becoming increasingly clear is that epigenetic processes such as chromatin remodelling and histone modification must play important roles for ES control. The active ES is highly depleted of nucleosomes compared with silent ESs [57,58]. In addition, a number of chromatin remodelling or modification enzymes have been shown to play a role in regulating the mono-allelic expression of a single ES. These include the chromatin remodeller TbISWI (a member of the SWI2/SNF2 superfamily of DNA-dependent ATPases) [59], the telomere-binding protein RAP1 [60] and the histone methyltransferase DOT1B [61].

An additional factor that could play a critical role in ES expression is its exact location within the nucleus. The active ES is transcribed at a very high rate by RNAPI in an extranucleolar ESB (expression-site body) structure [62]. It is likely that the ESB contains not only a high concentration of RNAPI transcription initiation and elongation factors, but also the RNA-processing machinery necessary for efficient production of the highly abundant *VSG* transcript. Attempts to force a trypanosome to activate two ESs at a time have not been successful, but have resulted in the selection for trypanosomes that appear to be rapidly switching between the two [54]. These two ESs become positioned relatively close to each other within the cell, arguing that they could be sharing the same ESB. Molecular markers defining the ESB still need to be discovered.

Intriguingly, a link has been discovered between DNA replication and ES control, as knockdown of the cohesin subunit SCC1 results in an increase in *VSG* ES switching [63]. In this study it was shown that there is a cohesin-mediated delay in the separation of sister chromatids containing the active ES. The increased VSG switching observed after cohesin knockdown could argue that the newly replicated ES needs to interact with the single ESB in order to accurately inherit the epigenetic state of the active ES. Disruption of this interaction through cohesin knockdown could possibly result in ES 'resetting', allowing a switch to another ES. Alternatively, cohesin could influence the chromatin remodelling necessary to ensure that the newly replicated ES is transcriptionally active [63]. Further experiments will be necessary to unravel how the epigenetic state of the active ES is faithfully inherited by the daughter cells with relatively infrequent switches to a new ES.

Conclusions and future goals

Research into African trypanosomes has recently been greatly facilitated by the development of new methods. Electron tomography has made three-dimensional reconstructions of the *T. brucei* flagellar pocket possible, thereby allowing the discovery of new cellular structures [8,12]. ChIP-seq has facilitated the global mapping of epigenetic marks such as histone variants and modifications, or distribution of the unusual DNA base J over the entire *T. brucei* genome [38,40,42]. In addition, microarray experiments and transcriptome analysis have allowed the analysis of global differential levels of transcription in different trypanosome life-cycle stages, and have given us better insight into *trans*-splicing and polyadenylation signals [34,36]. These whole-genome approaches are giving us a much better idea of how the entire *T. brucei* genome is epigenetically modified and expressed.

However, trypanosomes are still mysterious parasites in that the central questions of antigenic variation remain unanswered. One major challenge will come in trying to understand how VSG switching events late in an infection utilizing segmental gene conversions are mediated. These experiments are not easy, as chronic infections of *T. brucei* can be hard to interpret, and it can be difficult to prove that different VSG variants have arisen sequentially from each other. Another key question will be to understand how the *VSG* ESs are controlled in such a stringently mono-allelic fashion within the context of a genome that is constitutively transcribed.

Future goals include answering some of the following questions:
- How does the counting mechanism ensuring mono-allelic expression of one out of 15 highly similar ESs operate?
- What epigenetic marks distinguish active ESs from silent ones?
- Which of these epigenetic features are important for ES control?
- What are the molecular components defining the ESB?
- How does the mechanism allowing novel mosaic *VSG*s to be created through segmental gene conversion of multiple *VSG* pseudogenes operate?
- Where in the *T. brucei* genome are these new *VSG* chimaeras created?
- How is the observed diversity in the *VSG* repertoires created and maintained?

Hopefully, answers to these questions, as well as elucidating some interesting molecular mechanisms, will allow us to identify trypanosome biochemical pathways that differ enough from those of the host to provide us with an Achilles heel for tackling this evasive pathogen.

Summary

- *African trypanosomes are highly adapted for life as extracellular parasites of the mammalian bloodstream and have various features that facilitate immune evasion.*

- *A specialized morphology restricts most invariant cell-surface receptors to the sheltered location of the flagellar pocket.*
- *Extremely high rates of recycling of a protective VSG coat allow the trypanosome to strip off host anti-VSG antibodies.*
- *A highly sophisticated strategy of antigenic variation of the VSG coat allows the trypanosome to switch between many hundreds of VSGs, thereby temporarily avoiding recognition by host anti-VSG antibodies.*
- *The ability of the trypanosome to assemble fragments of VSG pseudo-genes into new functional VSGs allows an almost endless number of new mosaic VSG coats to be produced during a chronic infection.*
- *Hopefully, a better understanding of these different trypanosome immune-evasion strategies will eventually allow us to develop the means for disrupting them.*

References

1. Hide, G. and Tait, A. (2009) Molecular epidemiology of African sleeping sickness. Parasitology **136**, 1491–1500
2. Isobe, T., Holmes, E.C. and Rudenko, G. (2003) The transferrin receptor genes of *Trypanosoma equiperdum* are less diverse in their transferrin binding site than those of the broad-host range *Trypanosoma brucei*. J. Mol. Evol. **56**, 377–386
3. Taylor, J.E. and Rudenko, G. (2006) Switching trypanosome coats: what's in the wardrobe? Trends Genet. **22**, 614–620
4. Morrison, L.J., Marcello, L. and McCulloch, R. (2009) Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. Cell. Microbiol. **11**, 1724–1734
5. Vaughan, S. (2010) Assembly of the flagellum and its role in cell morphogenesis in *Trypanosoma brucei*. Curr. Opin. Microbiol. **13**, 453–458
6. Broadhead, R., Dawe, H.R., Farr, H., Griffiths, S., Hart, S.R., Portman, N., Shaw, M.K., Ginger, M.L., Gaskell, S.J., McKean, P.G. and Gull, K. (2006) Flagellar motility is required for the viability of the bloodstream trypanosome. Nature **440**, 224–227
7. Griffiths, S., Portman, N., Taylor, P.R., Gordon, S., Ginger, M.L. and Gull, K. (2007) RNA interference mutant induction *in vivo* demonstrates the essential nature of trypanosome flagellar function during mammalian infection. Eukaryotic Cell **6**, 1248–1250
8. Lacomble, S., Vaughan, S., Gadelha, C., Morphew, M.K., Shaw, M.K., McIntosh, J.R. and Gull, K. (2009) Three-dimensional cellular architecture of the flagellar pocket and associated cytoskeleton in trypanosomes revealed by electron microscope tomography. J. Cell Sci. **122**, 1081–1090
9. Field, M.C. and Carrington, M. (2009) The trypanosome flagellar pocket. Nat. Rev. Microbiol. **7**, 775–786
10. Overath, P. and Engstler, M. (2004) Endocytosis, membrane recycling and sorting of GPI-anchored proteins: *Trypanosoma brucei* as a model system. Mol. Microbiol. **53**, 735–744
11. Bonhivers, M., Nowacki, S., Landrein, N. and Robinson, D.R. (2008) Biogenesis of the trypanosome endo-exocytotic organelle is cytoskeleton mediated. PLoS Biol. **6**, e105

12. Gadelha, C., Rothery, S., Morphew, M., McIntosh, J.R., Severs, N.J. and Gull, K. (2009)
     Membrane domains and flagellar pocket boundaries are influenced by the cytoskeleton in African
     trypanosomes. Proc. Natl. Acad. Sci. U.S.A. **106**, 17425–17430

13. Schwede, A. and Carrington, M. (2010) Bloodstream form trypanosome plasma membrane
     proteins: antigenic variation and invariant antigens. Parasitology **137**, 2029–2039

14. Ferguson, M.A. (1999) The structure, biosynthesis and functions of glycosylphosphatidylinositol
     anchors, and the contributions of trypanosome research. J. Cell Sci. **112**, 2799–2809

15. Sheader, K., Vaughan, S., Minchin, J., Hughes, K., Gull, K. and Rudenko, G. (2005) Variant surface
     glycoprotein RNA interference triggers a precytokinesis cell cycle arrest in African trypanosomes.
     Proc. Natl. Acad. Sci. U.S.A. **102**, 8716–8721

16. Smith, T.K., Vasileva, N., Gluenz, E., Terry, S., Portman, N., Kramer, S., Carrington, M., Michaeli,
     S., Gull, K. and Rudenko, G. (2009) Blocking variant surface glycoprotein synthesis in *Trypanosoma
     brucei* triggers a general arrest in translation initiation. PLoS ONE **4**, e7532

17. Engstler, M., Thilo, L., Weise, F., Grunfelder, C.G., Schwarz, H., Boshart, M. and Overath, P.
     (2004) Kinetics of endocytosis and recycling of the GPI-anchored variant surface glycoprotein in
     Trypanosoma brucei. J. Cell Sci. **117**, 1105–1115

18. Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N. and Overath,
     P. (2007) Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. Cell
     131, 505–515

19. Marcello, L., Menon, S., Ward, P., Wilkes, J.M., Jones, N.G., Carrington, M. and Barry, J.D. (2007)
     VSGdb: a database for trypanosome variant surface glycoproteins, a large and diverse family of
     coiled coil proteins. BMC Bioinformatics **8**, 143

20. 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

21. McCulloch, R., Rudenko, G. and Borst, P. (1997) Gene conversions mediating antigenic
     variation in *Trypanosoma brucei* can occur in variant surface glycoprotein expression sites lacking
     70-base-pair repeat sequences. Mol. Cell. Biol. **17**, 833–843

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. Hartley, C.L. and McCulloch, R. (2008) *Trypanosoma brucei* BRCA2 acts in antigenic variation and
     has undergone a recent expansion in BRC repeat number that is important during homologous
     recombination. Mol. Microbiol. **68**, 1237–1251

24. Proudfoot, C. and McCulloch, R. (2005) Distinct roles for two RAD51-related genes in
     Trypanosoma brucei antigenic variation. Nucleic Acids Res. **33**, 6906–6919

25. 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

26. 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

27. 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

28. Jackson, A.P., Sanders, M., Berry, A., McQuillan, J., Aslett, M.A., Quail, M.A., Chukualim, B.,
     Capewell, P., MacLeod, A., Melville, S.E. et al. (2010) The genome sequence of *Trypanosoma brucei
     gambiense*, causative agent of chronic human african trypanosomiasis. PLoS Negl. Trop. Dis. **4**,
     e658

29. Balmer, O. and Caccone, A. (2008) Multiple-strain infections of *Trypanosoma brucei* across Africa.
     Acta Trop. **107**, 275–279

30. Hutchinson, O.C., Picozzi, K., Jones, N.G., Mott, H., Sharma, R., Welburn, S.C. and Carrington,
     M. (2007) Variant surface glycoprotein gene repertoires in *Trypanosoma brucei* have diverged to
     become strain-specific. BMC Genomics **8**, 234

31. Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B. et al. (2005) The genome of the African trypanosome *Trypanosoma brucei*. Science **309**, 416–422

32. Marcello, L. and Barry, J.D. (2007) Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. Genome Res. **17**, 1344–1352

33. Roth, C., Bringaud, F., Layden, R.E., Baltz, T. and Eisen, H. (1989) Active late-appearing variable surface antigen genes in *Trypanosoma equiperdum* are constructed entirely from pseudogenes. Proc. Natl. Acad. Sci. U.S.A. **86**, 9375–9379

34. Veitch, N.J., Johnson, P.C., Trivedi, U., Terry, S., Wildridge, D. and MacLeod, A. (2010) Digital gene expression analysis of two life cycle stages of the human-infective parasite, *Trypanosoma brucei* gambiense reveals differentially expressed clusters of co-regulated genes. BMC Genomics **11**, 124

35. Liang, X.H., Haritan, A., Uliel, S. and Michaeli, S. (2003) *trans* and *cis* splicing in trypanosomatids: mechanism, factors, and regulation. Eukaryotic Cell **2**, 830–840

36. Siegel, T.N., Hekstra, D.R., Wang, X., Dewell, S. and Cross, G.A. (2010) Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. Nucleic Acids Res. **38**, 4946–4957

37. Clayton, C.E. (2002) Life without transcriptional control? From fly to man and back again. EMBO J. **21**, 1881–1888

38. Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S. and Cross, G.A. (2009) Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes Dev. **23**, 1063–1076

39. Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R. et al. (2005) The genome of the kinetoplastid parasite, *Leishmania major*. Science **309**, 436–442

40. Wright, J.R., Siegel, T.N. and Cross, G.A. (2010) Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*. Mol. Biochem. Parasitol. **172**, 141–144

41. Talbert, P.B. and Henikoff, S. (2009) Chromatin-based transcriptional punctuation. Genes Dev. **23**, 1037–1041

42. Cliffe, L.J., Siegel, T.N., Marshall, M., Cross, G.A. and Sabatini, R. (2010) Two thymidine hydroxylases differentially regulate the formation of glucosylated DNA at regions flanking polymerase II polycistronic transcription units throughout the genome of *Trypanosoma brucei*. Nucleic Acids Res. **38**, 3923–3935

43. Borst, P. and Sabatini, R. (2008) Base J: discovery, biosynthesis, and possible functions. Annu. Rev. Microbiol. **62**, 235–251

44. van Leeuwen, F., Kieft, R., Cross, M. and Borst, P. (1998) Biosynthesis and function of the modified DNA base β-D-glucosyl-hydroxymethyluracil in *Trypanosoma brucei*. Mol. Cell. Biol. **18**, 5643–5651

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

46. Pays, E., Lips, S., Nolan, D., Vanhamme, L. and Perez-Morga, D. (2001) The VSG expression sites of *Trypanosoma brucei*: multipurpose tools for the adaptation of the parasite to mammalian hosts. Mol. Biochem. Parasitol. **114**, 1–16

47. Bitter, W., Gerrits, H., Kieft, R. and Borst, P. (1998) The role of transferrin-receptor variation in the host range of *Trypanosoma brucei*. Nature **391**, 499–502

48. Barker, A.R., Wickstead, B., Gluenz, E. and Gull, K. (2008) Bioinformatic insights to the *ESAG5* and *GRESAG5* gene families in kinetoplastid parasites. Mol. Biochem. Parasitol. **162**, 112–122

49. Becker, M., Aitcheson, N., Byles, E., Wickstead, B., Louis, E. and Rudenko, G. (2004) Isolation of the repertoire of VSG expression site containing telomeres of *Trypanosoma brucei* 427 using transformation-associated recombination in yeast. Genome Res. **14**, 2319–2329

50. Young, R., Taylor, J.E., Kurioka, A., Becker, M., Louis, E.J. and Rudenko, G. (2008) Isolation and analysis of the genetic diversity of repertoires of VSG expression site containing telomeres from *Trypanosoma brucei gambiense*, *T.b. brucei* and *T. equiperdum*. BMC Genomics **9**, 385

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

52. Roditi, I., Furger, A., Ruepp, S., Schurch, N. and Butikofer, P. (1998) Unravelling the procyclin coat of *Trypanosoma brucei*. Mol. Biochem. Parasitol. **91**, 117–130

53. McStay, B. and Grummt, I. (2008) The epigenetics of rRNA genes: from molecular to chromosome biology. Annu. Rev. Cell. Dev. Biol. **24**, 131–157

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

55. Borst, P. (2002) Antigenic variation and allelic exclusion. Cell **109**, 5–8

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

57. Stanne, T.M. and Rudenko, G. (2010) Active VSG expression sites in *Trypanosoma brucei* are depleted of nucleosomes. Eukaryotic Cell **9**, 136–147

58. Figueiredo, L.M. and Cross, G.A. (2010) Nucleosomes are depleted at the VSG expression site transcribed by RNA polymerase I in African trypanosomes. Eukaryotic Cell **9**, 148–154

59. Hughes, K., Wand, M., Foulston, L., Young, R., Harley, K., Terry, S., Ersfeld, K. and Rudenko, G. (2007) A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. EMBO J. **26**, 2400–2410

60. Yang, X., Figueiredo, L.M., Espinal, A., Okubo, E. and Li, B. (2009) RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. Cell **137**, 99–109

61. Figueiredo, L.M., Janzen, C.J. and Cross, G.A. (2008) A histone methyltransferase modulates antigenic variation in African trypanosomes. PLoS Biol. **6**, e161

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

63. Landeira, D., Bart, J.M., Van Tyne, D. and Navarro, M. (2009) Cohesin regulates VSG monoallelic expression in trypanosomes. J. Cell Biol. **186**, 243–254