Review

Antigenic variation in African trypanosomes

David Horn*

Division of Biological Chemistry & Drug Discovery, College of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK

ARTICLE INFO

ABSTRACT

Studies on Variant Surface Glycoproteins (VSGs) and antigenic variation in the African trypanosome, *Trypanosoma brucei*, have yielded a remarkable range of novel and important insights. The features first identified in *T. brucei* extend from unique to conserved-among-trypanosomatids to conserved-among-eukaryotes. Consequently, much of what we now know about trypanosomatid biology and much of the technology available has its origin in studies related to VSGs. *T. brucei* is now probably the most advanced early branched eukaryote in terms of experimental tractability and can be approached as a pathogen, as a model for studies on fundamental processes, as a model for studies on eukaryotic evolution or often all of the above. In terms of antigenic variation itself, substantial progress has been made in understanding the expression and switching of the VSG coat, while outstanding questions continue to stimulate innovative new approaches. There are large numbers of *VSG* genes in the genome but only one is expressed at a time, always immediately adjacent to a telomere. DNA repair processes allow a new *VSG* to be copied into the single transcribed locus. A coordinated transcriptional switch can also allow a new *VSG* gene to be activated without any detectable change in the DNA sequence, thereby maintaining singular expression, also known as allelic exclusion. I review the story behind VSGs; the genes, their expression and switching, their central role in *T. brucei* virulence, the discoveries that emerged along the way and the persistent questions relating to allelic exclusion in particular.

Contents

1. Introduction

Variant Surface Glycoproteins (*VSGs*) are developmentally regulated genes that mediate immune evasion. Activated in the tsetse-fly salivary-gland [1], and inactivated upon return to the tsetse-fly mid-gut, they produce a protective cell coat throughout the mammalian infective cycle. The coat must provide robust protection as *Trypanosoma brucei* occupy the bloodstream and tissue-spaces of their hosts and are fully exposed to immune surveillance in this hostile environment. Indeed, as an infection persists, the vast majority of the parasite population is periodically eliminated. The key features underlying successful immune evasion are clone-specific singular VSG expression combined with switching from one VSG to another. The metacyclic cells in the salivary gland are challenging to study since *VSG* expression is heterogeneous during this phase and the yield of *T. brucei* from flies is limiting. Most studies, therefore, have been conducted using bloodstream forms, more recently in axenic culture. Antigenic

* Tel.: +44 (0) 1382 386 286.
  E-mail address: d.horn@dundee.ac.uk

variation continues to operate in this environment [2] indicating that host antibodies are selective rather than a trigger for variation. An advantage here is that *VSG* switching operates at a frequency of approximately 1 switch/10^5 cells per population doubling, allowing the analysis of almost homogeneous but switchable populations.

Many seminal discoveries have emerged from studies on VSGs in *T. brucei* and the drive to understand VSGs and their expression has also led to the development of many of the tools and technologies now available for a range of other studies on trypanosomatids. Indeed, studies on *VSG* expression revealed much of what we now know about gene expression in trypanosomatids. Some features are specific to *VSG* gene expression sites, while others operate across the genome and are conserved in trypanosomatids that do not express *VSGs*. Thus, work on VSGs has informed studies on other important parasites, including *Trypanosoma cruzi* and *Leishmania* species. *VSGs* in *T. b. gambiense*, *T. b. rhodesiense*, *T. equiperdum*, *T. congolense* and *T. vivax* are not discussed in detail here but a similar system of gene expression and antigenic variation appears to operate in *T. brucei brucei* and in these other African trypanosomes.

What I present below is a somewhat historical perspective on antigenic variation in *T. brucei* and, in this regard, I recommend further reading of some of the older papers in particular, not often cited these days but often impressive when viewed in this historical context. It is also worth noting that few studies on antigenic variation in *T. brucei* have been or are currently specifically focussed on the prospect of a therapy in the short term. The central role of VSGs in virulence does mean that improved knowledge in this area is likely to present further opportunities for intervention, however.

2. A very brief early history – pre *VSG* gene-cloning

Sir David Bruce had read David Livingstone's reports on the tsetse fly diseases known as nagana in cattle and sleeping sickness in humans and, while searching for the cause more than 100 years ago, reported that" a rapidly moving object was seen lashing about among the red blood corpuscles . . . probably a trypanosome" [3]. Bruce also noted "the parasites come and go in the blood" and Franke & Ehrlich had deduced in 1905 that *T. brucei* acquired properties that conferred resistance to host "defensive substances". Ronald Ross and others then enumerated the relapsing parasitaemia in patients [4], albeit treated with several different drugs during monitoring in this case. A number of parasitic infections of mammals are now known to display relapses due to the emergence of new variants that are no longer susceptible to the latest host immune response (Fig. 1A).

In the 1960s, Keith Vickerman's work using electron microscopy revealed the dense *T. brucei* coat responsible for clone-specific relapses [5]. The identification and purification of the coat proteins by George Cross followed in the 1970s [6] and then the cloning and sequencing of the corresponding cDNA in the late 1970s and early 1980s (detailed below). The VSG responsible for clone-specific immunity or antigenic variation were found to be *Variant* throughout much of their length, they were known to represent the major component of the trypanosome *Surface* coat and they were *Glycoproteins*, decorated with multiple sugar residues [6].

3. The variant surface glycoprotein coat

The 15 nm thick VSG coat covers the entire cell and is an essential virulence factor. Formed from approximately 10 million molecules of approximately 60 kDa, the coat represents up to 20% of total cell protein [6], facilitating the production of antisera that recognise distinct VSGs [7]. VSGs are present at the cell surface as homodimers and, despite extreme sequence divergence, display remarkably similar structures [8], partly due to a conserved
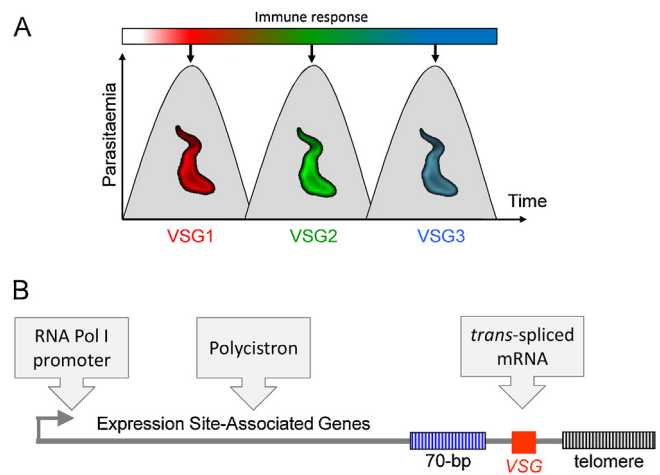

Fig. 1. *VSG* expression and switching. (A) *VSG* switching brings about antigenic variation. Combined with successive immune responses, this can generate a relapsing parasitaemia. Natural infections are more complex than this highly simplified schematic. (B) Studies on *VSG* expression revealed some unusual features. The single expressed *VSG* was found to be flanked by distinct repetitive sequences. Three further unusual features are indicated (boxes).

arrangement of disulphide bonds. Coat exchange during a VSG-switch appears to be primarily by dilution during cell division since cells divide approximately every 6 h with shedding and turnover being relatively much slower [9]. Cytokinesis is in fact dependent upon VSG supply as demonstrated by knockdown experiments [10]. The VSG is not a transmembrane protein but is rather anchored in the membrane by glycosylphosphatidylinositol (GPI). Indeed, GPI was discovered in *T. brucei* and this yielded the first GPI structure [11] as well as a description of GPI biosynthesis [12]. GPI anchors were subsequently discovered in mammals and in other cells [13] but the biosynthetic pathways do display differences [14].

The fluid nature of the VSG coat allows for a remarkably high rate of recycling involving endocytosis at the flagellar pocket [15]. This allows the coat to be cleansed of antibodies, at low titre at least, and this requires vigorous directional cell motility mediated by the flagellum [16]. VSG coats are highly immunogenic however, so once antibody titre increases, the vast majority of parasites are eliminated and only cells with distinct VSG coats survive. As well as this variable function, the VSG coat also serves to protect less variable or even invariant surface proteins from immune effectors. Comparison of the genome sequences of *T. brucei*, *T. congolense* and *T. vivax* revealed not only how *VSG* repertoires evolved [17] but also allowed reconstruction of a surface phylome [17]. This revealed a diversity of potential non-VSG surface proteins, nutrient receptors and other 'invariant' surface glycoproteins. It appears that the densely packed and thick VSG coat can physically obstruct access to these proteins by conventional immunoglobulins while selectively allowing access to smaller molecules such as nutrients [18]. Nanobodies also have the potential to access the less variable epitopes usually hidden within the coat by the size exclusion limit [19].

4. *VSG* genes and their subtelomeric environment

Access to the first cloned *VSG* genes provided probes to explore copy number, location and diversity within trypanosome genomes, especially because the 3′ terminal regions of *VSG* mRNAs were found to be conserved [20]. It was soon recognised that the single expressed *VSG* was adjacent to a telomere, or a chromosome discontinuity that was subject to *Bal*31 exonuclease digestion [21,22] and this strict association remains intact to date (Fig. 1B). This stimulated much interest in telomere biology in trypanosomes,

which revealed the addition (and occasional large deletions) of TTAGGG/CCCTAA-repeats to growing telomeres [23–25]; the same hexameric repeats were later found at human telomeres. *T. brucei* telomeres were also found to terminate in t-loops [26] and this was followed by the identification and characterisation of telomerase reverse transcriptase, responsible for telomere growth, and other telomere-repeat-binding proteins [27].

The *VSG*-telomere association actually goes much further. Each *T. brucei* genome contains around 250 telomeres, almost all of which may be closely linked to non-transcribed *VSG* genes, with their 3′-ends closest to the telomere. Around 80% of these telomeres reside on 50–100 kbp long minichromosomes [22,28,29], which appear to be entirely dedicated to *VSG* archiving. Another ten or so telomeres reside on 'intermediate' chromosomes and the remaining 44 reside on the eleven pairs of megabase-chromosomes that comprise the diploid genome. With many additional arrays of subtelomeric *VSGs*, it is currently thought that up to 30% of an African trypanosome genome is dedicated to archiving up to 2000 *VSG* genes and gene-fragments.

Although the size of the *VSG* archive and the telomeric environment continue to present challenges for complete genome assembly and functional analyses, important insights into *VSG* gene evolution and diversification have emerged from genome sequencing [30] and the cloning and sequencing of large intact telomeric fragments [31]. Extensive hemizygous subtelomeric domains on the megabase chromosomes are dedicated to arrays of archival *VSGs* [32], meaning that many *VSGs* are present as a single copy even in a diploid genome. Most of these *VSGs* are pseudogenes in *T. brucei* (749/804 analysed) and these *VSGs* are flanked by 70-bp repeats (Fig. 1B) upstream [33] and highly conserved elements within the 3′-untranslated region (3′-UTR); both of these sequences facilitate recombination (see below). The 3′-UTR is also involved in specific stabilisation of the *VSG* mRNA at the bloodstream stage, contributing to a half-life of 4.5 h [34].

5. *VSG* gene expression – *trans*-splicing and polycistronic Pol I transcription

The 'yeast to human' view of eukaryotic diversity is very narrow so it is not surprising that several dogmas have been overturned by work on the divergent trypanosomes. Studies on *VSGs* revealed some unusual features underlying gene-expression in trypanosomatids for example (Fig. 1B). S1 nuclease protection and RNA blotting experiments revealed a spliced segment at the 5′ end of the *VSG* mRNA and reverse transcription then showed this "mini-exon" or "spliced leader" sequence to be 35 nt long [35]. Intriguingly, the same sequence was found at the 5′-end of two other *VSG* mRNAs [36]. *Cis*-splicing was initially considered to be the most likely explanation but it turned out that this same sequence was present on *all* mRNAs. In fact, discontinuous mRNA synthesis through bimolecular splicing or *trans*-splicing allows mature mRNA to be derived from precursor RNAs transcribed from two different chromosomes [37]. *Trans*-splicing was also found to operate in other trypanosomatids [38] and in nematodes.

The search for a *VSG* gene promoter extended further from the gene itself than expected. In most eukaryotes, each gene has its own promoter, while polycistrons are largely restricted to prokaryotes. Cloning and mapping upstream of an active *VSG* revealed a large expanse of imperfect 70-bp repeats [39] and then an Expression Site Associated Gene or *ESAG* [40] and then several more *ESAGs* [41] forming a polycistron. It was found that α-amanitin failed to inhibit ribosomal RNA (*rRNA*) transcription, as expected, but also *VSG* transcription [42] implicating RNA Pol I. These findings prompted a detailed analysis of RNA pol I and associated factors in *T. brucei* (see below). Polycistronic transcription was also found to

operate within Pol II transcription units elsewhere in the genome [43] and this has proven to be a pervasive feature of trypanosomatid genomes. As for the *VSG* expression site promoter, the first one was eventually found around 60 kbp from the *VSG* [44] and was indeed, following mounting evidence, confirmed to recruit RNA Pol I *in vitro* [45]. Nuclear 'run-on' assays combined with inhibition of transcription elongation using UV exposure were instrumental in locating the *VSG* expression site promoter [46], the sequence of which was unrelated to the more conventional *rDNA* promoter [47]. The metacyclic promoters responsible for *VSG* transcription in the tsetse-fly salivary gland were also found to recruit RNA Pol I but were distinct from the bloodstream promoters and in this case were located only a short distance upstream of the telomeric *VSGs* [48]; these are among only a few monocistronic transcription units in trypanosomatids.

The result of all this work revealed a somewhat surprising situation whereby RNA Pol I and RNA Pol II are required to produce mature *VSG* expression site transcripts; transcription of the spliced leader sequence was confirmed to be RNA Pol II dependent [49]. Further surprises were in store as assessment of other potential Pol II promoters in *T. brucei* or in other trypanosomatids failed to yield any further conventional examples. Rather, around sixty RNA Pol II 'transcription initiation sites' appear to depend upon a particular chromatin structure [50] that is not readily reconstituted on reporter constructs.

Fifteen copies of the highly conserved subtelomeric *VSG* expression sites active in the bloodstream have now been identified, cloned by recombination in yeast and sequenced from the most widely studies *T. brucei* strain [31]. Present on the diploid and intermediate-chromosomes, these contain intact promoters and are competent transcription units but are typically reversibly repressed due to monoallelic *VSG* expression control. Many *ESAGs* found within these polycistronic units remain to be characterised but the evidence so far points to roles in host-parasite interactions. For example, a novel heterodimeric transferrin receptor, encoded by *ESAG6* and *ESAG7* and related to the *N*-terminal domain of the VSG, has the capacity to bind transferrin from different hosts with different affinities [51]. A human serum resistance associated gene, or *SRA*, found in *T. brucei rhodesiense*, also resembles a *VSG* and is also an *ESAG* [52] while another gene resembling a *VSG*, known as *TgsGP*, confers human-serum resistance to *T. brucei gambiense* [53]. *ESAG4* genes are unrelated to *VSGs*, but also mediate host-parasite interactions. These genes encode adenylate cyclases which are released by lysed trypanosomes and inhibit the innate immune response [54]. Clearly, *VSGs* and their associated genes have been central to the 'arms-race' operating and evolving at the host-parasite interface. The vast reservoir available has allowed *VSG* genes to be co-opted to functions beyond classical antigenic variation. The relationships among these proteins could equally reflect an evolutionary origin of VSGs from ancient surface receptors.

6. Antigenic variation by *VSG* gene rearrangement

Subtelomeres are recombinogenic hotspots, plastic regions of genomes enriched in gene families that are most commonly involved in adaptation to different environments. Like many other cell types, African trypanosomes appear to have exploited these properties, in this case for the massive expansion and evolution of the *VSG* family and also for the *ESAGs* [55]. Antigenic variation in *T. brucei* involves switching to expression of a distinct *VSG* so the availability of *VSG* cDNA clones allowed researchers to look for changes associated with the activation and inactivation of those *VSGs*. Early analyses using Southern blotting [56,57], northern blotting [58], DNAse I digestion [59] or DNA sequencing [60] revealed switching by duplicative transposition of a non-telomeric 'basic copy'
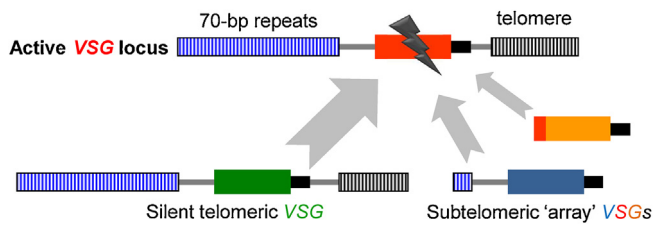
Fig. 2. DNA recombination is central to *VSG* switching. The active subtelomeric *VSG* locus is prone to spontaneous DNA breaks. Three classes of repair templates are shown with *VSGs* represented as coloured boxes; those sharing more (flanking) homology with the active site are used more frequently (grey arrows). The homologous sequences indicated are the 70-bp repeats (blue stripes), the telomeric repeats (black stripes), the *VSG* 3′-UTRs (thick black bars) and a portion of the *VSG* coding sequence (red). A break in or around the active *VSG* is followed by DNA resection extending towards the 70-bp repeats, often initiating recombination in this region.

VSG and replacement of the old *VSG* by the new 'expression-linked copy' at a single transcribed locus. Silent telomeric *VSG* cassettes appeared to be copied all the way to the end of the chromosome [61,62] by a mechanism now known as 'break-induced replication' (BIR).

Subsequent analyses revealed a range of variations on the recombination theme and DNA homology emerged as the key driver of these *VSG* rearrangements [63] stimulating further studies on a range of *T. brucei* DNA-repair factors. Most *VSGs* are flanked upstream by common 70-bp repeats and even non-telomeric genes share sequences at the 3′-end [64], sometimes within the region encoding the VSG *C*-terminal domain and almost always in the conserved *VSG* 3′-UTR (Fig. 2). Replacement of larger parts of the expression site or non-duplicative telomere exchange [65] can bring about a switch but these events appear to be relatively rare and even actively suppressed [66]. Translocation to the active site is preferentially initiated by the long tracts of 70-bp repeats found upstream of the active *VSG* [67,68] so the probability of being 'selected' as a template for repair depends upon these and other sequences shared with the active *VSG* locus [69]. Indeed, 70-bp repeat recombination is primarily responsible for the high-efficiency duplication of telomeric *VSGs* [61,62], which can employ the BIR mechanism, while chromosome-internal *VSGs* must use a 'gene-conversion' mechanism, similar to BIR but with a second recombination junction, most often within the *VSG* 3′-UTR [39,70]. Conversion of even shorter *VSG* segments can generate 'mosaic' *VSGs* that, because of their lower-frequency emergence, become increasingly important for parasite persistence in a chronic infection [64,71–73], as immunity to more-frequently activated VSGs builds [67].

This understanding of the shared sequences that drive recombination can now explain why the expression of *VSGs* is 'semi-ordered' or somewhat predictable. Any order though will be isolate-dependent, less predictable as the infection progresses and will ultimately be highly dynamic and variable in different infections and also from one epidemic to the next, meaning that herd-immunity is unlikely to be achieved. The potential *VSG* repertoire is ultimately larger than a single genomic repertoire due to segmental gene conversion and inter-strain mating in the tsetse fly [74]. This genome plasticity also impacts the *VSGs* expressed in the tsetse fly salivary gland and presents stark challenges in terms of any vaccine strategy that targets VSGs.

VSG recombination does not appear to be naturally triggered by a site-specific nuclease but rather appears to depend upon the inherent instability of subtelomeres [75,76]. Breaks were shown to arise naturally at telomeric *VSG* loci, probably due to replication fork collapse, and an artificial, meganuclease-induced DNA break at the active site can trigger a switch [75,77]. These breaks initiate DNA resection producing ssDNA and triggering a homology search.

Once a suitable template is found, DNA can be copied from the template to repair the lesion [76]. Non-homologous DNA repair does not appear to operate in *T. brucei*, placing an emphasis on RAD51-dependent homologous recombination [63,78]. This has had a major impact on our ability to manipulate the *T. brucei* genome and also makes an important contribution to *VSG* gene rearrangements. There is an alternative form of microhomology-mediated end-joining, however. This end-joining is RAD51-independent and may be particularly effective within 70-bp repeats, thereby making a substantial contribution to the duplicative transposition of *VSGs* [76].

Nuclear positioning and the chromatin environment of *VSGs* may be important for *VSG* recombination [77]. In the bloodstream form, telomeres [79], active [80] and silent *VSG* expression sites [81] are distributed throughout the nuclear space rather than sequestered at the periphery and this may facilitate homology searching during DNA repair. Notably, the active *VSG* expression site specifically migrates to the nuclear periphery during differentiation to the insect stage [82].

Switching occurs in only approximately 0.001% of cells per cell division cycle in experimental *in vitro* culture or during frequent syringe passage but appears to be much higher naturally; switch rate returns to >0.2% following transmission through flies [83] and also apparently increases during growth *in vivo* in mammals [68,84]. This transition is not understood but could involve the acquisition of a hyper-labile or hyper-recombinogenic state at the active *VSG* locus [85]. Rapid switching also operates in metacyclic cells obtained from tsetse-fly salivary glands [86] but in this case, switching does not involve recombination [87].

The vast and dynamic *VSG* gene family, the large number of subtelomeres in *T. brucei* and the incomplete sequence coverage currently available for these regions presents challenges but current techniques should now allow some longstanding questions to be addressed. For example, where in the genome are often entirely pseudogene-derived mosaic-*VSGs* assembled and presumably selected for? Do non-templated *VSG* mutations [88] naturally contribute to immune evasion? Are novel *VSGs* typically preserved or permanently lost once successfully used for immune evasion? In terms of this last question, the presence of intact *VSGs* on minichromosomes [29] could reflect an effective archiving mechanism for novel *VSGs*. It will also be of interest to determine whether *VSG* recombination requires or exploits dedicated or specifically modified DNA repair factors, such as BRCA2 with an expanded set of RAD51-interacting repeats [89].

7. Antigenic variation by *VSG* transcription (in)activation – allelic exclusion

Studies on *T. brucei* clones with switched VSGs also revealed that some switching events were not associated with duplicative *VSG* transposition [90]. Pulsed-field gel-electrophoresis, combined with Southern blotting, which had also facilitated 'mapping' of the *VSG* rearrangements described above, provided confirmation of coordinated on and off switching, known as 'in situ' switching [28]. This showed that certain telomeric *VSGs* were reversibly active or repressed and could be subject to transcriptional switching as well as recombination [61,91]. The *in situ* switches are classical epigenetic switches since they occur without changes in the DNA sequence [44]. Thus, although gene expression control in trypanosomatids is primarily post-transcriptional, *VSG* exclusion represents a prominent exception (Fig. 3). Although simultaneous expression of two VSGs from distinct telomeric sites has been reported in *T. equiperdum* [92], VSG double-expressors arose rarely and were unstable in *T. brucei* [93]. Thus, *VSG* allelic exclusion is generally strictly maintained.
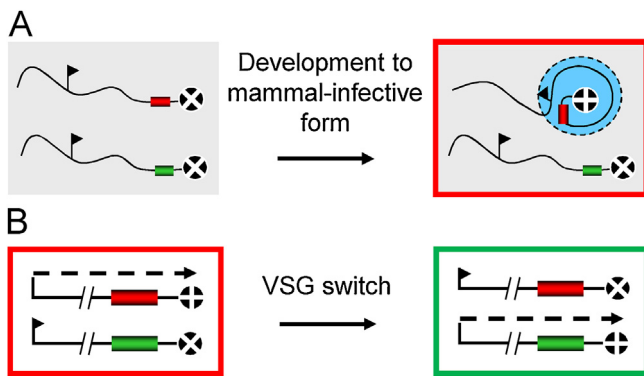
Fig. 3. *VSG* allelic exclusion is not yet understood. (A) Only one of the telomeric *VSGs* (red and green) is active and only in the mammalian stages. This associates with a focus of extranucleolar RNA Pol I and other factors (light blue) and produces the VSG coat (outer red box). The '+' and 'x' symbols indicate the active and silent telomeres, respectively. (B) Transcription can occasionally switch from one telomeric *VSG* to another; a coordinated epigenetic switch maintains allelic exclusion. The dashed arrow indicates transcription. Other details as in A.

The mechanism that distinguishes among almost identical *VSG* expression sites in the bloodstream form suppresses both transcription initiation [94] and elongation [95] and this results in a remarkable 10,000-fold abundance differential among active and silent *VSG* mRNAs [96]. Despite the distinct sequences of *rDNA* promoters, these promoters, inserted at *VSG* expression sites, adopt the transcriptional status of that site [97,98]. A second *VSG* inserted at an active *VSG* expression site is similarly also active [99,100]. In contrast to this locus-specific control in the bloodstream form, it seems that *VSG* promoter-specific elements, distinct from *rDNA* promoters, allow selective down-regulation of all *VSG* transcription during differentiation to the insect-stage [101].

A novel DNA base, the hydroxylated and glucosylated derivative of thymidine or base-J [102], was first discovered at silent *VSG* loci. Evidence of J-base came from blocked cleavage by certain restriction enzymes such as *Pst*I as revealed on Southern-blots. Base-J is notably absent from insect-stage *T. brucei* but, although implicated for many years, no role in antigenic variation has been demonstrated [103]. What is established is that J-base is required for RNA Pol II transcription termination in *L. major* [104] so it remains possible that J-base also presents a barrier to certain RNA polymerases or DNA polymerases in *T. brucei*, potentially impacting transcription or DNA recombination and repair, respectively.

Transcription of the single active *VSG* by RNA Pol I, rather than RNA Pol II, suggested a potential 'privileged domain' model for activation based on association with the nucleolus. As it turned out, active *VSG* transcripts and the active *VSG* locus were found to be extranucleolar [81]. Like the nucleolus though, this region, known as the expression-site body, is associated with an accumulation of RNA Pol I [80], a trypanosomatid-specific transcription initiation complex known as class I transcription factor A [105] and also a pol I-associated high-mobility group factor [106].

It remains likely that the telomeric environment is important for *VSG* allelic exclusion as well as for *VSG* recombination. There is evidence of a role for the telomere-binding protein, RAP1, in *VSG* silencing [96] and this silencing is also dependent upon chromatin structure [107,108]; since this topic is covered in a recent review [77] it is not covered in any detail here. Briefly, the histones, histone chaperones, chromatin remodelers, chromatin modifiers, cohesins and nuclear lamina, as well as other chromatin-associated factors, contribute to repression. DOT1B in particular, has a major role in establishing the silent state during *in situ* switching [109]. Thus, transcription is clearly repressed or attenuated at telomeric *VSG* loci. Understanding allelic exclusion will remain challenging until

we know more about the selection of a single *VSG* for expression and how this is coordinated with the silencing of all other *VSGs*.

8. Concluding remarks

The abundance of the active VSG, quite stable yet reversible *VSG* repression and the ease of genetic manipulation and cell culture mean that *T. brucei* provides a highly tractable experimental system for the study of monoallelic expression and antigenic variation. There has been tremendous progress in our understanding in this area, how the *VSGs* are organised and expressed and how expression is switched, through recombination in particular. An ancient and ongoing 'arms race' between host immunity and parasite immune evasion has been illuminated through studies on VSGs. The set of monocistronic VSGs first expressed in the tsetse-fly salivary gland facilitates the establishment of a mammalian infection following a blood-meal. The multiplicity of telomeric *VSGs* with alternative collections of *ESAGs* may then provide an opportunity to select the optimal expression site for effective nutrition and growth in distinct host environments [51]. Recombination can then allow for *VSG* switching compatible with continued expression of a favoured set of *ESAGs*. The vast reservoir of *VSG* genes allows the presentation of constantly changing epitopes at the cell surface to counter the hosts' capacity for adaptive immunity.

Access to genome sequence data changed the research landscape, allowing easy access to a vast number of *VSG* sequences, factors involved in transcription, telomere-binding, recombination and repair and chromatin-based control. There has been an inevitable focus on factors related to those with known functions in other eukaryotes, however, meaning that a lot of territory still remains uncharted in trypanosomatid research. Recent technical advances in areas such as forward genetics [110], proteomics [111], improved access to *T. brucei* developmental stages in *in vitro* culture [112] and new and improved technologies to come, will surely now help to deliver answers to some of the outstanding questions.

References

[1] Tetley L, Turner CM, Barry JD, Crowe JS, Vickerman K. Onset of expression of the variant surface glycoproteins of *Trypanosoma brucei* in the tsetse fly studied using immunoelectron microscopy. J Cell Sci 1987;87(Pt 2):363–72.

[2] Doyle JJ, Hirumi H, Hirumi K, Lupton EN, Cross GA. Antigenic variation in clones of animal-infective *Trypanosoma brucei* derived and maintained *in vitro*. Parasitology 1980;80:359–69.

[3] Bruce D. The Croonian Lectures on trypanosomes causing disease in man and domestic animals in central Africa. Br Med J 1915;2:5–10.

[4] Ross R, Thomson D. A case of sleeping sickness showing regular periodical increase of the parasites disclosed. Br Med J 1910;1:1544–5.

[5] Vickerman K. On the surface coat and flagellar adhesion in trypanosomes. J Cell Sci 1969;5:163–93.

[6] Cross GA, Identification. purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. Parasitology 1975;71:393–417.

[7] Cross GA. Cross-reacting determinants in the *C*-terminal region of trypanosome variant surface antigens. Nature 1979;277:310–2.

[8] Blum ML, Down JA, Gurnett AM, Carrington M, Turner MJ, Wiley DC. A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. Nature 1993;362:603–9.

[9] Seyfang A, Mecke D, Duszenko M. Degradation, recycling, and shedding of *Trypanosoma brucei* variant surface glycoprotein. J Protozool 1990;37:546–52.

[10] Sheader K, Vaughan S, Minchin J, Hughes K, Gull K, Rudenko G. Variant surface glycoprotein RNA interference triggers a precytokinesis cell cycle arrest in African trypanosomes. Proc Natl Acad Sci USA 2005;102:8716–21.

[11] Ferguson MA, Homans SW, Dwek RA, Rademacher TW. Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. Science 1988;239:753–9.

[12] Masterson WJ, Doering TL, Hart GW, Englund PT. A novel pathway for gly-can assembly: biosynthesis of the glycosyl-phosphatidylinositol anchor of the trypanosome variant surface glycoprotein. Cell 1989;56:793–800.

[13] Ferguson MA, Williams AF. Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. Annu Rev Biochem 1988;57:285–320.

[14] Ferguson MA. Glycosylphosphatidylinositol biosynthesis validated as a drug target for African sleeping sickness. Proc Natl Acad Sci USA 2000;97:10673–5.

[15] Field MC, Carrington M. The trypanosome flagellar pocket. Nat Rev Microbiol 2009;7:775–86.

[16] Engstler M, Pfohl T, Herminghaus S, Boshart M, Wiegertjes G, Heddergott N, et al. Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. Cell 2007;131:505–15.

[17] Jackson AP, Allison HC, Barry JD, Field MC, Berriman M. A cell-surface phylome for African trypanosomes. PLoS Negl Trop Dis 2013;7:e2121.

[18] Mehlert A, Wormald MR, Ferguson MA. Modeling of the *N*-glycosylated trans-ferrin receptor suggests how transferrin binding can occur within the surface coat of *Trypanosoma brucei*. PLoS Pathog 2012;8:e1002618.

[19] Stijlemans B, Caljon G, Natesan SK, Saerens D, Conrath K, Perez-Morga D, et al. High affinity nanobodies against the *Trypanosoma brucei* VSG are potent trypanolytic agents that block endocytosis. PLoS Pathog 2011;7:e1002072.

[20] Majumder HK, Boothroyd JC, Weber H. Homologous 3′-terminal regions of mRNAs for surface antigens of different antigenic variants of *Trypanosoma brucei*. Nucleic Acids Res 1981;9:4745–53.

[21] De Lange T, Borst P. Genomic environment of the expression-linked extra copies of genes for surface antigens of *Trypanosoma brucei* resembles the end of a chromosome. Nature 1982;299:451–3.

[22] Williams RO, Young JR, Majiwa PA. Genomic environment of *T. brucei VSG* genes: presence of a minichromosome. Nature 1982;299:417–21.

[23] Bernards A, Michels PA, Lincke CR, Borst P. Growth of chromosome ends in multiplying trypanosomes. Nature 1983;303:592–7.

[24] Blackburn EH, Challoner PB. Identification of a telomeric DNA sequence in *Trypanosoma brucei*. Cell 1984;36:447–57.

[25] Van der Ploeg LH, Liu AY, Borst P. Structure of the growing telomeres of trypanosomes. Cell 1984;36:459–68.

[26] Munoz-Jordan JL, Cross GA, de Lange T, Griffith JD. t-loops at trypanosome telomeres. Embo J 2001;20:579–88.

[27] Dreesen O, Li B, Cross GA. Telomere structure and function in trypanosomes: a proposal. Nat Rev Microbiol 2007;5:70–5.

[28] Van der Ploeg LH, Schwartz DC, Cantor CR, Borst P. Antigenic variation in *Trypanosoma brucei* analyzed by electrophoretic separation of chromosome-sized DNA molecules. Cell 1984;37:77–84.

[29] Wickstead B, Ersfeld K, Gull K. The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palin-dromes. Genome Res 2004;14:1014–24.

[30] Jackson AP, Berry A, Aslett M, Allison HC, Burton P, Vavrova-Anderson J, et al. Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species. Proc Natl Acad Sci USA 2012;109:3416–21.

[31] Hertz-Fowler C, Figueiredo LM, Quail MA, Becker M, Jackson A, Bason N, et al. Telomeric expression sites are highly conserved in *Trypanosoma brucei*. PLoS ONE 2008;3:e3527.

[32] Callejas S, Leech V, Reitter C, Melville S. Hemizygous subtelomeres of an African trypanosome chromosome may account for over 75% of chromosome length. Genome Res 2006;16:1109–18.

[33] Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, et al. The genome of the African trypanosome *Trypanosoma brucei*. Science 2005;309:416–22.

[34] Berberof M, Vanhamme L, Tebabi P, Pays A, Jefferies D, Welburn S, et al. The 3′-terminal region of the mRNAs for VSG and procyclin can confer stage specificity to gene expression in *Trypanosoma brucei*. Embo J 1995;14:2925–34.

[35] Van der Ploeg LH, Liu AY, Michels PA, De Lange T, Borst P, Majumder HK, et al. RNA splicing is required to make the messenger RNA for a variant surface antigen in trypanosomes. Nucleic Acids Res 1982;10:3591–604.

[36] Boothroyd JC, Cross GA. Transcripts coding for variant surface glycoproteins of *Trypanosoma brucei* have a short, identical exon at their 5′ end. Gene 1982;20:281–9.

[37] Kooter JM, De Lange T, Borst P. Discontinuous synthesis of mRNA in try-panosomes. Embo J 1984;3:2387–92.

[38] Borst P. Discontinuous transcription and antigenic variation in trypanosomes. Annu Rev Biochem 1986;55:701–32.

[39] Liu AY, Van der Ploeg LH, Rijsewijk FA, Borst P. The transposition unit of variant surface glycoprotein gene 118 of *Trypanosoma brucei*. Presence of repeated elements at its border and absence of promoter-associated sequences. J Mol Biol 1983;167:57–75.

[40] Cully DF, Ip HS, Cross GA. Coordinate transcription of variant surface glyco-protein genes and an expression site associated gene family in *Trypanosoma brucei*. Cell 1985;42:173–82.

[41] Pays E, Tebabi P, Pays A, Coquelet H, Revelard P, Salmon D, et al. The genes and transcripts of an antigen gene expression site from *T. brucei*. Cell 1989;57:835–45.

[42] Kooter JM, Borst P. α-Amanitin-insensitive transcription of variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. Nucleic Acids Res 1984;12:9457–72.

[43] Imboden MA, Laird PW, Affolter M, Seebeck T. Transcription of the inter-genic regions of the tubulin gene cluster of *Trypanosoma brucei*: evidence for a polycistronic transcription unit in a eukaryote. Nucleic Acids Res 1987;15:7357–68.

[44] Zomerdijk JC, Ouellette M, ten Asbroek AL, Kieft R, Bommer AM, Clayton CE, et al. The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. Embo J 1990;9:2791–801.

[45] Gunzl A, Bruderer T, Laufer G, Schimanski B, Tu LC, Chung HM, et al. RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. Eukaryot Cell 2003;2:542–51.

[46] Johnson PJ, Kooter JM, Borst P. Inactivation of transcription by UV irradiation of *T. brucei* provides evidence for a multicistronic transcription unit including a *VSG* gene. Cell 1987;51:273–81.

[47] Zomerdijk JC, Kieft R, Shiels PG, Borst P. α-Amanitin-resistant transcrip-tion units in trypanosomes: a comparison of promoter sequences for a *VSG* gene expression site and for the ribosomal RNA genes. Nucleic Acids Res 1991;19:5153–8.

[48] Alarcon CM, Son HJ, Hall T, Donelson JE. A monocistronic transcript for a trypanosome variant surface glycoprotein. Mol Cell Biol 1994;14:5579–91.

[49] Gilinger G, Bellofatto V. Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. Nucleic Acids Res 2001;29:1556–64.

[50] Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, et al. Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes Dev 2009;23:1063–76.

[51] Bitter W, Gerrits H, Kieft R, Borst P. The role of transferrin-receptor variation in the host range of *Trypanosoma brucei*. Nature 1998;391:499–502.

[52] Xong HV, Vanhamme L, Chamekh M, Chimfwembe CE, Van Den Abbeele J, Pays A, et al. A *VSG* expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. Cell 1998;95:839–46.

[53] Uzureau P, Uzureau S, Lecordier L, Fontaine F, Tebabi P, Homble F, et al. Mech-anism of *Trypanosoma brucei gambiense* resistance to human serum. Nature 2013;501:430–4.

[54] Salmon D, Vanwalleghem G, Morias Y, Denoeud J, Krumbholz C, Lhomme F, et al. Adenylate cyclases of *Trypanosoma brucei* inhibit the innate immune response of the host. Science 2012;337:463–6.

[55] Horn D, McCulloch R. Molecular mechanisms underlying the control of anti-genic variation in African trypanosomes. Curr Opin Microbiol 2010;13:700–5.

[56] Hoeijmakers JH, Frasch AC, Bernards A, Borst P, Cross GA. Novel expression-linked copies of the genes for variant surface antigens in trypanosomes. Nature 1980;284:78–80.

[57] Williams RO, Young JR, Majiwa PA. Genomic rearrangements correlated with antigenic variation in *Trypanosoma brucei*. Nature 1979;282:847–9.

[58] Pays E, Delronche M, Lheureux M, Vervoort T, Bloch J, Gannon F, et al. Cloning and characterization of DNA sequences complementary to messenger ribonu-cleic acids coding for the synthesis of two surface antigens of *Trypanosoma brucei*. Nucleic Acids Res 1980;8:5965–81.

[59] Pays E, Lheureux M, Steinert M. The expression-linked copy of a sur-face antigen gene in *Trypanosoma* is probably the one transcribed. Nature 1981;292:265–7.

[60] Borst P, Frasch AC, Bernards A, Van der Ploeg LH, Hoeijmakers JH, Arn-berg AC, et al. DNA rearrangements involving the genes for variant antigens in *Trypanosoma brucei*. Cold Spring Harb Symp Quant Biol 1981;45(Pt 2):935–43.

[61] Bernards A, De Lange T, Michels PA, Liu AY, Huisman MJ, Borst P. Two modes of activation of a single surface antigen gene of *Trypanosoma brucei*. Cell 1984;36:163–70.

[62] Scholler JK, Myler PJ, Stuart KD. A novel telomeric gene conversion in *Try-panosoma brucei*. Mol Biochem Parasitol 1989;35:11–9.

[63] McCulloch R, Barry JD. A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. Genes Dev 1999;13:2875–88.

[64] Marcello L, Barry JD. 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 2007;17:1344–52.

[65] Pays E, Guyaux M, Aerts D, Van Meirvenne N, Steinert M. Telomeric recip-rocal recombination as a possible mechanism for antigenic variation in trypanosomes. Nature 1985;316:562–4.

[66] Kim HS, Cross GA. TOPO3α influences antigenic variation by monitoring expression-site-associated *VSG* switching in *Trypanosoma brucei*. PLoS Pathog 2010;6:e1000992.

[67] Morrison LJ, Majiwa P, Read AF, Barry JD. Probabilistic order in antigenic variation of *Trypanosoma brucei*. Int J Parasitol 2005;35:961–72.

[68] Timmers HT, de Lange T, Kooter JM, Borst P. Coincident multiple acti-vations of the same surface antigen gene in *Trypanosoma brucei*. J Mol Biol 1987;194:81–90.

[69] Laurent M, Pays E, Van der Werf A, Aerts D, Magnus E, Van Meirvenne N, et al. Translocation alters the activation rate of a trypanosome surface antigen gene. Nucleic Acids Res 1984;12:8319–28.

[70] Bernards A, Van der Ploeg LH, Frasch AC, Borst P, Boothroyd JC, Coleman S, et al. Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3′ end. Cell 1981;27:497–505.

[71] Hall JP, Wang H, Barry JD. Mosaic VSGs and the scale of *Trypanosoma brucei* antigenic variation. PLoS Pathog 2013;9:e1003502.

[72] Pays E, Van Assel S, Laurent M, Darville M, Vervoort T, Van Meirvenne N, et al. Gene conversion as a mechanism for antigenic variation in trypanosomes. Cell 1983;34:371–81.

[73] Roth CW, Longacre S, Raibaud A, Baltz T, Eisen H. The use of incomplete genes for the construction of a Trypanosoma equiperdum variant surface glycoprotein gene. Embo J 1986;5:1065–70.

[74] Peacock L, Bailey M, Carrington M, Gibson W. Meiosis and haploid gametes in the pathogen Trypanosoma brucei. Curr Biol 2014;24:181–6.

[75] Boothroyd CE, Dreesen O, Leonova T, Ly KI, Figueiredo LM, Cross GA, et al. A yeast-endonuclease-generated DNA break induces antigenic switching in Trypanosoma brucei. Nature 2009;459:278–81.

[76] Glover L, Alsford S, Horn D. DNA break site at fragile subtelomeres determines probability and mechanism of antigenic variation in African trypanosomes. PLoS Pathog 2013;9:e1003260.

[77] Glover L, Hutchinson S, Alsford S, McCulloch R, Field MC, Horn D. Antigenic variation in African trypanosomes: the importance of chromosomal and nuclear context in VSG expression control. Cell Microbiol 2013;15:1984–93.

[78] Glover L, McCulloch R, Horn D. Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. Nucleic Acids Res 2008;36:2608–18.

[79] Perez-Morga D, Amiguet-Vercher A, Vermijlen D, Pays E. Organization of telomeres during the cell and life cycles of Trypanosoma brucei. J Eukaryot Microbiol 2001;48:221–6.

[80] Navarro M, Gull K. A pol I transcriptional body associated with VSG monoallelic expression in Trypanosoma brucei. Nature 2001;414:759–63.

[81] Chaves I, Zomerdijk J, Dirks-Mulder A, Dirks RW, Raap AK, Borst P. Subnuclear localization of the active variant surface glycoprotein gene expression site in Trypanosoma brucei. Proc Natl Acad Sci USA 1998;95:12328–33.

[82] Landeira D, Navarro M. Nuclear repositioning of the VSG promoter during developmental silencing in Trypanosoma brucei. J Cell Biol 2007;176:133–9.

[83] Turner CM. The rate of antigenic variation in fly-transmitted and syringe-passaged infections of Trypanosoma brucei. FEMS Microbiol Lett 1997;153:227–31.

[84] Lee MG, Van der Ploeg LH. Frequent independent duplicative transpositions activate a single VSG gene. Mol Cell Biol 1987;7:357–64.

[85] Hovel-Miner GA, Boothroyd CE, Mugnier M, Dreesen O, Cross GA, Papavasiliou FN. Telomere length affects the frequency and mechanism of antigenic variation in Trypanosoma brucei. PLoS Pathog 2012;8:e1002900.

[86] Esser KM, Schoenbechler MJ. Expression of two variant surface glycoproteins on individual African trypanosomes during antigen switching. Science 1985;229:190–3.

[87] Lenardo MJ, Esser KM, Moon AM, Van der Ploeg LH, Donelson JE. Metacyclic variant surface glycoprotein genes of Trypanosoma brucei subsp. rhodesiense are activated in situ, and their expression is transcriptionally regulated. Mol Cell Biol 1986;6:1991–7.

[88] Lu Y, Alarcon CM, Hall T, Reddy LV, Donelson JE. A strand bias occurs in point mutations associated with variant surface glycoprotein gene conversion in Trypanosoma rhodesiense. Mol Cell Biol 1994;14:3971–80.

[89] Hartley CL, McCulloch R. Trypanosoma brucei BRC2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number that is important during homologous recombination. Mol Microbiol 2008;68:1237–51.

[90] Young JR, Donelson JE, Majiwa PA, Shapiro SZ, Williams RO. analysis of genomic rearrangements associated with two variable antigen genes of Trypanosoma brucei. Nucleic Acids Res 1982;10:803–19.

[91] Myler P, Nelson RG, Agabian N, Stuart K. Two mechanisms of expression of a predominant variant antigen gene of Trypanosoma brucei. Nature 1984;309:282–4.

[92] Baltz T, Giroud C, Baltz D, Roth C, Raibaud A, Eisen H. Stable expression of two variable surface glycoproteins by cloned Trypanosoma equiperdum. Nature 1986;319:602–4.

[93] Chaves I, Rudenko G, Dirks-Mulder A, Cross M, Borst P. Control of variant surface glycoprotein gene-expression sites in Trypanosoma brucei. Embo J 1999;18:4846–55.

[94] Nguyen TN, Muller LS, Park SH, Siegel TN, Gunzl A. Promoter occupancy of the basal class I transcription factor A differs strongly between active and silent VSG expression sites in Trypanosoma brucei. Nucleic Acids Res 2014;42:3164–76.

[95] Vanhamme L, Poelvoorde P, Pays A, Tebabi P, Van Xong H, Pays E. Differential RNA elongation controls the variant surface glycoprotein gene expression sites of Trypanosoma brucei. Mol Microbiol 2000;36:328–40.

[96] Yang X, Figueiredo LM, Espinal A, Okubo E, Li B. RAP1 is essential for silencing telomeric variant surface glycoprotein genes in Trypanosoma brucei. Cell 2009;137:99–109.

[97] Horn D, Cross GA. A developmentally regulated position effect at a telomeric locus in Trypanosoma brucei. Cell 1995;83:555–61.

[98] Rudenko G, Blundell PA, Dirks-Mulder A, Kieft R, Borst P. 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 1995;83:547–53.

[99] Munoz-Jordan JL, Davies KP, Cross GA. Stable expression of mosaic coats of variant surface glycoproteins in Trypanosoma brucei. Science 1996;272:1795–7.

[100] Smith TK, Vasileva N, Gluenz E, Terry S, Portman N, Kramer S, et al. Blocking variant surface glycoprotein synthesis in Trypanosoma brucei triggers a general arrest in translation initiation. PLoS ONE 2009;4:e7532.

[101] Horn D, Cross GA. Position-dependent and promoter-specific regulation of gene expression in Trypanosoma brucei. Embo J 1997;16:7422–31.

[102] Borst P, Sabatini R. Base J: discovery, biosynthesis, and possible functions. Annu Rev Microbiol 2008;62:235–51.

[103] Cliffe LJ, Siegel TN, Marshall M, Cross GA, Sabatini R. 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 2010;38:3923–35.

[104] van Luenen HG, Farris C, Jan S, Genest PA, Tripathi P, Velds A, et al. Glucosylated hydroxymethyluracil, DNA base J, prevents transcriptional readthrough in Leishmania. Cell 2012;150:909–21.

[105] Brandenburg J, Schimanski B, Nogoceke E, Nguyen TN, Padovan JC, Chait BT, et al. Multifunctional class I transcription in Trypanosoma brucei depends on a novel protein complex. Embo J 2007;26:4856–66.

[106] Narayanan MS, Rudenko G. TDP1 is an HMG chromatin protein facilitating RNA polymerase I transcription in African trypanosomes. Nucleic Acids Res 2013;41:2981–92.

[107] Figueiredo LM, Cross GA. Nucleosomes are depleted at the VSG expression site transcribed by RNA polymerase I in African trypanosomes. Eukaryot Cell 2010;9:148–54.

[108] Stanne TM, Rudenko G. Active VSG expression sites in Trypanosoma brucei are depleted of nucleosomes. Eukaryot Cell 2010;9:136–47.

[109] Figueiredo LM, Janzen CJ, Cross GA. A histone methyltransferase modulates antigenic variation in African trypanosomes. PLoS Biol 2008;6:e161.

[110] Alsford S, Eckert S, Baker N, Glover L, Sanchez-Flores A, Leung KF, et al. High-throughput decoding of antitrypanosomal drug efficacy and resistance. Nature 2012;482:232–6.

[111] Urbaniak MD, Guther ML, Ferguson MA. Comparative SILAC proteomic analysis of Trypanosoma brucei bloodstream and procyclic lifecycle stages. PLoS ONE 2012;7:e36619.

[112] Kolev NG, Ramey-Butler K, Cross GA, Ullu E, Tschudi C. Developmental progression to infectivity in Trypanosoma brucei triggered by an RNA-binding protein. Science 2012;338:1352–3.