Review TRENDS in Parasitology Vol.21 No.11 November 2005

Regulation of antigen gene expression in *Trypanosoma brucei*

Etienne Pays

Laboratory of Molecular Parasitology, Institut de Biologie Moléculaire et Médicale, Université Libre de Bruxelles, 12 rue des Professeurs Jeener et Brachet, B6041 Gosselies, Belgium

The trypanosome genome is organized into long polycistronic units that seem to be permanently transcribed in proliferative stages of the parasite. Cellular differentiation is controlled primarily at the level of individual mRNA maturation and stability. The transcription units of the two major stage-specific antigens, the variant surface glycoprotein (VSG) of the bloodstream form and procyclin of the procyclic form, are subject to an additional layer of control: the mutually exclusive activation of RNA elongation and processing. The high recombination frequency prevailing in the telomere that harbours the active *VSG* expression site has been exploited by the parasite to both drive antigenic variation and generate VSG-based adaptive proteins.

Twenty years on

Twenty years ago, with my late mentor and friend Maurice Steinert, I published a report in *Parasitology Today* about the current views regarding the stage-specific regulation of antigen gene expression in trypanosomes [1]. Since then, with the (almost) complete knowledge of the genome of these parasites, one would have expected major advances in identifying the factors and mechanisms involved. Unfortunately, despite undeniable progress in some areas, it is still not understood how the parasite switches antigen gene expression to adapt. In this article, I summarize where we are and where we are going in terms of this research.

Post-transcriptional regulation of gene expression

Now, it is generally accepted that the kinetoplastid genome is organized in long, polycistronic transcription units, with batteries of genes oriented in the same direction. These genes are usually separated by only a few hundred base pairs and, with a few exceptions, they do not contain introns. In *Trypanosoma brucei,* only four transcription promoters are known for certain. The promoter for the spliced leader RNA genes recruits RNA Pol II [2], whereas the three other known promoters recruit RNA Pol I. As expected, one of these promoters is that of the rDNA. Remarkably, the two others direct transcription of protein-coding genes, namely the genes for the two major stage-specific antigens, the variant surface glycoprotein (VSG) of bloodstream forms (there are probably ~20 *VSG* expression sites, one of which is active at any one time) and procyclins of procyclic forms (two *EP* units and two *GPEET* units, which are, probably, active simultaneously). The *VSG* and *procyclin* transcription units are unique because they contain protein-coding genes and use an rRNA polymerase [3], and because they are the only ones in which a beginning and a terminus has been identified. The *VSG* expression sites are insulated between arrays of repeats at the end of several chromosomes [4], whereas the *procyclin* units are embedded in regions that are transcribed by RNA Pol II. In *EP* units this results in overlapping segments that are transcribed on both strands [5], whereas, in *GPEET* units a long terminator sequence prevents this interference [6].

The general organization of genes in polycistronic units means that all mRNAs must be *trans*-spliced to be capped and that post-transcriptional regulation of the mRNA level is key to achieving developmental and conditional changes of gene expression [7]. This regulation is exerted through sequence elements in intergenic regions and evidence of the primary role of untranslated regions (UTRs) to determine the relative stage-specific mRNA abundance has accumulated in the past ten years [8–12]. Constitutive synthesis of the transcriptome and selection of the convenient messages only at the maturation step (Figure 1) probably enables the parasite to switch gene expression rapidly to survive and adapt to a new environment. This happens when the cell experiences a major stress that probably precludes the activation of energy-dependent mechanisms. Just changing the choice between pre-existing RNAs would be more efficient than switching promoter activity through the recruitment of specific silencers and/or activators as occurs in most eukaryotes. Although this system requires the permanent degradation of an important fraction of the transcriptome [10,13], trypanosomes have avoided the burden of encoding networks of specific transcription factors and target sequences. Regarding how the choice between transcripts is switched during the parasite life-cycle, the analysis of the developmental-stage intermediate between the proliferative bloodstream and procyclic forms, namely the quiescent stumpy form, indicates that a first uncoupling between RNA polymerase and the RNA-elongation and RNA-processing machinery occurs before modification of the latter [14] (Figure 1).

An example of this process is the switch from VSG to procyclin when trypanosomes differentiate from bloodstream to procyclic forms. Even in this case, where the

Corresponding author: Pays, E. (epays@ulb.ac.be).
Available online 29 August 2005

 1471-4922/$ - see front matter doi:10.1016/j.pt.2005.08.016

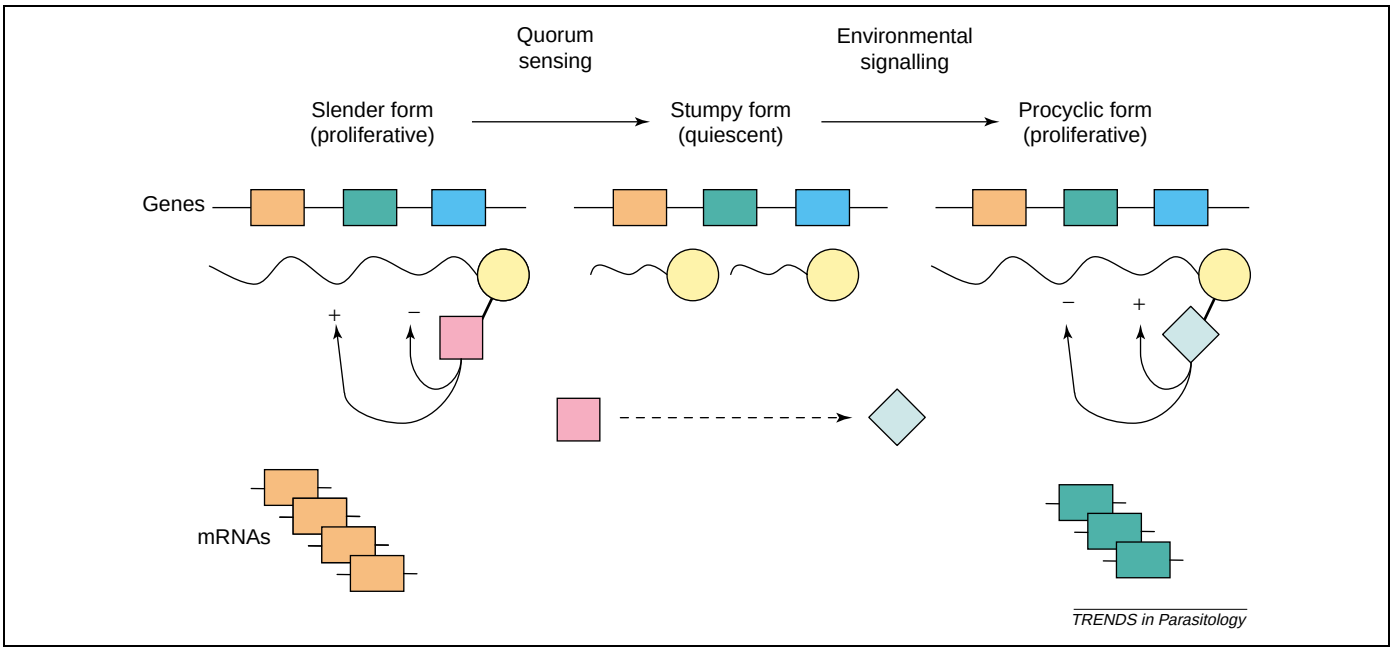

Figure 1. Control of gene expression in *Trypanosoma brucei*. In proliferative forms, most genes are transcribed permanently into polycistronic transcripts (wavy lines), and the stage-regulated control of transcript maturation (either + or −), which is exerted on intergenic sequences by RNA-elongating and RNA-processing machinery (represented here in two forms – by a pink square and light-green diamond, respectively), determines the steady-state amounts of mRNA. It is proposed that control is changed through dissociation of the machinery from RNA polymerase (yellow). This results in the strong transcriptional repression that is characteristic of quiescent stages and is followed by the recoupling of a form that is modified by environmental signalling. Modification of the RNA-elongating and RNA-processing machinery would also switch transcriptional activity between the mutually exclusive transcription units of the major surface antigens VSG and procyclin [17].

transcription promoters are well defined, there is no evidence for strict promoter control. These promoters are of the ribosomal type, and seem to be extremely simple. Despite the typical stage-specificity of the *VSG* and *procyclin* transcription units, their promoters seem to be active permanently, although some modulation of the recruitment of RNA polymerase cannot be excluded. By contrast, the *VSG* expression sites and *procyclin* units clearly show regulated progression of the RNA polymerase (RNA-elongation control). Because *VSG* and *procyclin* are the last and first genes of their respective transcription units, RNA-elongation control targets these genes as a priority when the bloodstream conditions are altered: the *VSG* gene is the first to be silenced and the *procyclin* genes are the first to undergo an increase in transcription [15]. The permanent recruitment of RNA Pol I on both promoters enables the continuous synthesis of primary transcripts of the first genes of both units, namely *ESAG7* and *ESAG6* in *VSG* expression sites, and the *procyclin* genes in *procyclin* units. Recent data [16] highlight the importance of this last observation. The cellular differentiation from bloodstream to procyclic forms seems to require two independent, successive signals from the environment, a cold shock to 20°C and the presence of micromolar amounts of citrate or *cis*-aconitate [16]. Whereas the second signal might depend on the appearance of a stage-specific receptor at the cellular surface, the first operates on a restricted sequence of the 3′-UTR of *procyclin* transcripts that are still produced at basal levels in bloodstream forms. Normally, these basal levels do not enable detectable expression of *procyclin* at this stage but, owing to the mRNA-stabilizing effect induced by cold shock through the 3′-UTR, efficient production of EP procyclin can be achieved. This example *in vivo* illustrates the prevailing role of post-transcriptional over transcriptional controls, which parallels similar results in an experimental system using reporter gene constructs opposing post-transcriptional and transcriptional regulation of the same units [8]. In this case, the presence of a *VSG* 3′UTR downstream from the reporter gene abrogated transcriptional stimulation by the procyclin promoter during the parasite differentiation from bloodstream to procyclic forms.

The everlasting mystery of VSG switching

The puzzle of trypanosome antigenic variation has been tackled by increasingly sophisticated technology, but the number of reviews about the subject might be larger than the number of articles that report progress in understanding the mechanisms and factors involved (for review, see Ref. [17]). The basics were already established in 1986; of several genomic loci, all telomeric, only one *VSG* expression site at a time is active in bloodstream forms ('monoallelic' expression), and none is active in procyclic forms. VSG switching results either from homologous DNA recombination that is targeted to the unique active site, which is mainly gene conversion, or from transcriptional *in situ* switching between different *VSG* expression sites.

Concerning the enzymes that are involved in recombination, the trypanosomal homologue of yeast RAD51 seems to have a role, although it is not clear that this enzyme also affects switching *in situ* [18]. The recombination frequency in the active locus must be extremely high, given the accessibility of DNA in the open configuration of chromatin and the presence of homologous sequences that are repeated in many loci, particularly around silent *VSG* genes. Therefore, the active *VSG* expression site can be considered to be a recombination hot-spot that generates

genes for new VSGs and for novel VSG-based adaptive molecules [19]. The potential of this genomic environment to generate novelty has been highlighted by the recent discovery, following genome sequencing, that most *VSG*s are pseudogenes. Thus, during chronic infection, most genes that contribute to antigenic variation have to recombine partially with the *VSG* resident in the active expression site to reconstitute functional genes. In this way, chimeric *VSG*s must appear continuously in this site. As discussed elsewhere [19], *T. brucei* both generates novel *VSG* sequences and stores these sequences temporarily in the genome through transcriptional switching to other *VSG* expression sites. Therefore, the hyperevolution of the *VSG* repertoires seems to be linked closely to the mechanisms of antigenic variation.

The nature of the mono-allelic control is even less clear. The site where transcription of the active *VSG* unit occurs is distinct from the nucleolus despite the use of a ribosomal polymerase [20]. Whether this site contains special structural elements or whether it reflects the presence of the necessary machinery for RNA synthesis, elongation and processing (usually termed the 'RNA factory') remains to be determined. DNA repair factors might be involved in the control of the 'silent' *VSG* expression sites because treatments that affect either DNA synthesis or DNA damage lead to upregulation of these units [21]. Unpublished data from my laboratory further support this hypothesis and extend the observation to include transcriptional factors that are involved in both DNA repair and RNA elongation (see model in [17]). It is probable that factors affecting chromatin structure also have a role, but telomere silencing, which exists in trypanosomes as in yeast, is unlikely to be the key because (i) the *VSG* promoters are too distant from the telomere, (ii) the knockdown of genes that are likely to be involved in telomere silencing do not influence antigenic variation, and (iii) the absence of telomeric repeats does not affect the stage-regulated control of the VSG promoter (discussed in [17]). Heterochromatinization of the 'silent' sites, possibly secured by unusual DNA modification that recruits chromatin-remodeling proteins [22,23], might be a consequence rather than the cause of transcription inhibition. Whether this system is driven by active transcriptional (de)repression or by the requirement for stage-regulated limiting factors is still an open question.

Procyclin regulation: more complex than thought initially

Since 1986, the procyclin question has gained complexity. It is now clear that several procyclins exist, which are subdivided into two major classes and several isoforms. Although these have different expression controls and, probably, different functions related to the journey of the parasite in the fly [24,25], the requirement of procyclin for cyclical transmission of the parasite has not been demonstrated. Whereas the developmental regulation of the *procyclin* units is as obscure as that of the *VSG* units, there has been major progress in characterizing the post-transcriptional control of the *procyclin* genes [10–12,16]. It seems that the 3′-UTR of *procyclin* mRNA contains discrete elements that are responsible for regulation by environmental factors such as glucose, hypoxic conditions and glycerol (repression of GPEET) [11,12], and cold shock (induction of EP) [16]. Moreover, it is possible that both *procyclin* mRNAs have responsive elements for factors that are induced by citrate and/or cis-aconitate. In all cases the signalling pathways and UTR-interacting factors are totally unknown.

Crosstalk between VSG and procyclin

Factors that upregulate the production of *procyclin* mRNA also systematically downregulate transcription in the active *VSG* expression site, and this opposite control acts on 3′-UTRs and RNA elongation [8,15]. Interestingly, a similar transcriptional crosstalk seems to operate in cell-cycle-arrested stumpy bloodstream forms, in which transcription of *VSG* and *procyclin* is downregulated and upregulated, respectively [14]. These observations indicate the involvement of common factors that are recruited differentially to the two stage-specific units depending on the environmental conditions [17]. In this respect, it might be relevant to note that a common protein seems to bind to the *VSG* and *procyclin* promoters, as well as to the *procyclin* terminator [26]. More generally, it is striking that the genomic organization of the *VSG* and *procyclin* transcription units has common features: they are transcribed by the same, unusual, RNA polymerase, and are associated with genes of the same families. The *procyclin* units contain the *VSG*-like genes *PAG1*, *PAG2* and *PAG4* [27] as well as *ESAG2*-like genes [28] and, as revealed in the genome database, they are also linked closely to *VSG*-like, *ESAG4*-like and *ESAG11*-like genes.

Emerging new functions of the VSG

It is remarkable that, apart from VSG and procyclin, the number of known trypanosome antigens and surface receptors is limited and most are stage-specific [29]. Only some have an identified function, namely transferrin receptor for ESAG7/6, and transmembrane adenylate cyclase for the ESAG4 family [19]. Several antigens share the VSG folding [30], and the transferrin receptor seems to be constituted with two VSG N-terminal domains [31]. This last observation might represent the tip of the iceberg because there are additional examples of putative VSG-based receptors in both developmental forms [27,30,32]. It possible that different receptors are constructed from VSGs, probably following DNA-recombination events that occur in the *VSG* expression site, as might be the case for the transferrin receptor. In addition, the VSG itself might have additional functions other than antigenic variation. The exceptional rate of VSG and surface-membrane trafficking [33] is probably essential to ensure that the parasite escapes from the immune response by the rapid internalization of bound antibodies, but it might also be involved in the uptake of other macromolecules such as complement elements and cytokines [34]. Finally, the recent example of the adaptation of *T. b. rhodesiense* to humans, where a truncated VSG is responsible for resistance of the parasite to the trypanolytic factor apolipoprotein L-I [35], highlights the potential of the VSG system in generating factors that enable escape from innate immunity of the host.

Future perspectives

Obviously, much remains to be done in this field and, with the trypanosome genome in hand, the pace of research should be improved greatly. In this particular organism, in which transcription initiation seems to be barely controlled, if at all, trying to identify transcriptional activators and repressors does not seem to be the most promising approach. Accordingly, the genome database has not revealed clear candidates for these functions. By contrast, basic characterization of the components of the transcriptional machineries seems to be both feasible and useful, two essential criteria of any sound undertaking. Knowing the factors that are involved in RNA synthesis should enable the subsequent identification of downstream control elements. Not surprisingly, the available data point to environmental stress as a driving force for controlling the system of antigen switching through changes in mRNA maturation and stability, and transcriptional activation of genomic loci for major stage-specific antigens. The possible involvement of DNA-repair factors in the control of VSG expression sites is in keeping with this suggestion. Therefore, deciphering the stress signalling pathways of *T. brucei* might constitute an interesting approach for clarifying the mechanisms that regulate antigen gene expression in these parasites.

Acknowledgements

This article was inspired by and is dedicated to the memory of Maurice Steinert. I thank Derek Nolan for comments about the manuscript. Work in my laboratory is funded by the Belgian National Fund for Scientific Research (FNRS), the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) and the Interuniversity Attraction Poles Programme – Belgian Science Policy.

References

1 Steinert, M. and Pays, E. (1986) Selective expression of surface antigen genes in African trypanosomes. *Parasitol. Today* 2, 15–19
2 Das, A. and Bellofatto, V. (2003) RNA polymerase II-dependent transcription in trypanosomes is associated with a SNAP complex-like transcription factor. *Proc. Natl. Acad. Sci. U. S. A.* 100, 80–85
3 Günzl, A. *et al*. (2003) RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryot. Cell* 2, 542–551
4 Sheader, K. *et al*. (2003) Delineation of the regulated Variant Surface Glycoprotein gene expression site domain of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 128, 147–156
5 Liniger, M. *et al*. (2001) Overlapping sense and antisense transcription units in *Trypanosoma brucei*. *Mol. Microbiol.* 40, 869–878
6 Berberof, M. *et al*. (1996) Characterization of a transcription terminator of the procyclin PARP A unit of *Trypanosoma brucei*. *Mol. Cell. Biol.* 16, 914–924
7 Clayton, C.E. (2002) Life without transcriptional control? From fly to man and back again. *EMBO J.* 21, 1881–1888
8 Berberof, M. *et al*. (1995) The 3′-terminal region of the mRNAs for VSG and procyclin can confer stage-specificity to gene expression in *Trypanosoma brucei*. *EMBO J.* 14, 2925–2934
9 Drozdz, M. and Clayton, C.E. (1999) Structure of a regulatory 3′ untranslated region from *Trypanosoma brucei*. *RNA* 5, 1632–1644
10 Irmer, H. and Clayton, C. (2001) Degradation of the unstable EP1 mRNA in *Trypanosoma brucei* involves initial destruction of the 3′-untranslated region. *Nucleic Acids Res.* 29, 4707–4715
11 Morris, J.C. *et al*. (2002) Glycolysis modulates trypanosome glycoprotein expression as revealed by an RNAi library. *EMBO J.* 21, 4429–4438
12 Vassella, E. *et al*. (2004) Expression of a major surface protein of Trypanosoma brucei insect forms is controlled by the activity of mitochondrial enzymes. *Mol. Biol. Cell* 15, 3986–3993
13 Brems, S. *et al*. (2005) The transcriptomes of *Trypanosoma brucei* Lister 427 and TREU927 bloodstream and procyclic trypomastigotes. *Mol. Biochem. Parasitol.* 139, 163–172
14 Amiguet-Vercher, A. *et al*. (2004) Loss of the mono-allelic control of the VSG expression sites during the development of *Trypanosoma brucei* in the bloodstream. *Mol. Microbiol.* 51, 1577–1588
15 Vanhamme, L. *et al*. (1995) Stimuli of differentiation regulate RNA elongation in the transcription units for the major stage-specific antigens of *Trypanosoma brucei*. *Nucleic Acids Res.* 23, 1862–1869
16 Engstler, M. and Boshart, M. (2004) Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in *Trypanosoma brucei*. *Genes Dev.* 18, 2798–2811
17 Pays, E. *et al*. (2004) Antigenic variation in *Trypanosoma brucei*. Facts, challenges and mysteries. *Curr. Opin. Microbiol.* 7, 369–374
18 McCulloch, R. and Barry, J.D. (1999) A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes Dev.* 13, 2875–2888
19 Pays, E. *et al*. (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
20 Navarro, M. and Gull, K. (2001) A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. *Nature* 414, 759–763
21 Sheader, K. *et al*. (2004) Bloodstream form specific upregulation of silent VSG expression sites and procyclin in *Trypanosoma brucei* after inhibition of DNA synthesis or DNA damage. *J. Biol. Chem.* 279, 13363–13374
22 Cross, M. *et al*. (2002) J-binding protein increases the level and retention of the unusual base J in trypanosome DNA. *Mol. Microbiol.* 46, 37–47
23 DiPaolo, C. *et al*. (2005) Regulation of trypanosome DNA glycosylation by a SWI2/SNF2-like protein. *Mol. Cell* 17, 441–451
24 Vassella, E. *et al*. (2001) Multiple procyclin isoforms are expressed differentially during the development of insect forms of *Trypanosoma brucei*. *J. Mol. Biol.* 312, 597–607
25 Chandra, M. *et al*. (2004) TsetseEP, a gut protein from the tsetse *Glossina morsitans*, is related to a major surface glycoprotein of trypanosomes transmitted by the fly and to the products of a *Drosophila* gene family. *Insect Biochem. Mol. Biol.* 34, 1163–1173
26 Berberof, M. *et al*. (2000) A single-stranded DNA-binding protein shared by telomeric repeats, the VSG transcription promoter and the procyclin transcription terminator of *Trypanosoma brucei*. *Nucleic Acids Res.* 28, 597–604
27 Koenig-Martin, E. *et al*. (1992) A procyclin-associated gene in *Trypanosoma brucei* encodes a polypeptide related to ESAG 6 and 7 proteins. *Mol. Biochem. Parasitol.* 55, 135–145
28 Berberof, M. *et al*. (1991) A similar gene is shared by both the VSG and procyclin gene transcription units of *Trypanosoma brucei*. *Mol. Cell. Biol.* 11, 1473–1479
29 Pays, E. and Nolan, D.P. (1998) Expression and function of surface proteins in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 91, 3–36
30 Carrington, M. and Boothroyd, J. (1996) Implications of conserved structural motifs in disparate trypanosome surface proteins. *Mol. Biochem. Parasitol.* 81, 119–126
31 Salmon, D. *et al*. (1997) Characterization of the ligand-binding site of the transferrin receptor in *Trypanosoma brucei* demonstrates a structural relationship with the N-terminal domain of the variant surface glycoprotein. *EMBO J.* 16, 7272–7278
32 Berberof, M. *et al*. (2001) A receptor-like flagellar pocket glycoprotein specific to *Trypanosoma brucei gambiense*. *Mol. Biochem. Parasitol.* 113, 109–120
33 Engstler, M. *et al*. (2004) Kinetics of endocytosis and recycling of the GPI-anchored variant surface glycoprotein in *Trypanosoma brucei*. *J. Cell Sci.* 117, 1105–1115
34 Magez, S. *et al*. (2001) A conserved flagellar pocket exposed high mannose moiety is used by African trypanosomes as a host cytokine-binding molecule. *J. Biol. Chem.* 276, 33458–33464
35 Vanhamme, L. *et al*. (2003) Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature* 422, 83–87