Antigenic variation allows African trypanosomes to develop chronic infections in mammalian hosts. This process results from the alternative occurrence of transcriptional switching and DNA recombination targeted to a telomeric locus that contains the gene of the variant antigen and is subjected to mono-allelic expression control. So far, the identification of mechanisms and factors involved still resists technological developments and genome sequencing.

Abbreviations
ES  expression site
ESAG expression site-associated gene
ESB expression site body
VSG variant surface glycoprotein

Introduction
African trypanosomes, the causative agents of human sleeping sickness, are favorite model organisms for molecular parasitologists, particularly for the spectacular mechanism of antigenic variation that develops when these parasites divide in the blood of their mammalian hosts. The entire trypanosome surface is covered with a homogeneous and dense coat made of approximately five million dimers of a single antigen termed variant surface glycoprotein (VSG) that is repeatedly changed in a fraction of the population. This allows these trypanosomes to escape antibody-mediated killing and repopulate the host, resulting in the development of long-lasting chronic infection. The basic rules of this system are now well defined [1–5]. The trypanosome genome contains hundreds of VSG genes (VSGs), only one of which is expressed at a time. This occurs in one of several polycistronic transcription units contained in specialised loci termed VSG expression sites (ESs), which are all telomeric. Thus, antigenic variation occurs either by transcriptional switching between different ESs (in situ activation), or by VSG recombination within the active ES. However, despite numerous high-tech developments, none of the molecular players involved in these processes have been characterised so far, except RNA polymerase, which unexpectedly turned out to be RNA Pol I. In particular, a central question still remains unanswered: how is only a single ES active at a time? Interestingly, this question may also include the genes encoding the major surface antigen replacing the VSG in the insect-specific procyclic form of the parasite, termed procyclin. Indeed, not only are the expression of VSG and procyclin strictly mutually exclusive, but the procyclin genes are also transcribed by RNA Pol I. This review focuses on the most recent advances on these questions.

Subnuclear localisation
In eukaryotic organisms, the interphase nucleus is organised in functional subcompartments reflecting the steady-state association of its residents within the nucleoplasmic space. In Trypanosoma brucei, the active ES of long slender (rapidly dividing) bloodstream forms was found to be present in a specialised region termed expression site body (ESB) [6], which was clearly distinct from the nucleolus and had no preferential location in the nucleus. No such structure was detected for silent ESs. In slender forms, telomeres are clustered into several discrete spots distributed within the nucleoplasm, although a single ESB appears to recruit the full transcriptional machinery [7]. In bloodstream short stumpy (non-dividing) and procyclic forms in which the VSG is inactive, telomeres redistributed towards the nuclear periphery [8]. Therefore, it is possible that the central position of telomeres in long slender forms reflects the association of those regions with nuclear subcompartments active in transcription. According to a working hypothesis developed below, one may envisage that a competition for transcription factors establishes the unique status of the ESB. In both stumpy and procyclic forms, the telomeres, including the ESs, would not have access to the transcriptional machinery. Such a developmentally regulated global change of nuclear organisation was also observed during the differentiation of T. cruzi proliferative forms into non-dividing forms, where it was linked to a general inhibition of transcription [9].

Telomere silencing: unlikely
A particular subnuclear localisation, close to the nuclear envelope, has been associated with telomere silencing in
were devoid of telomeric repeats [14] in circular bacterial artificial chromosomes (BACs) that were devoid of telomeric repeats [14*].

Chromatin modification: nothing relevant so far

Genes for enzymes able to modify chromatin proteins, such as histone deacetylases or acetyl transferases, are present in the trypanosome genome [15–17]. However, analysis of sensitivity to DNase and nucleosome spacing did not reveal chromatin difference between promoter regions from active and silent ESs [18], although the conformation of chromatin was clearly more open along the active ES than in silent ESs [19,20]. On the other hand, general differences in chromatin structure between bloodstream and procyclic forms [21,22] may possibly account for the transient release of ES repression that was observed during cellular differentiation between bloodstream and procyclic forms [23]. This also points to different repression mechanisms on ESs in bloodstream and procyclic stages. Finally, extensive studies of a modified base (β-3-glucosyl-hydroxymethyl-uracil) termed J have so far not pointed to linkage to antigenic variation, despite the presence of this base in inactive but not active ESs. Instead, the available evidence has suggested a role as an epigenetic label of repeated DNA or heterochromatin [24,25].

Gene recombination: still no full explanation

The genomic rearrangements associated with antigenic variation [1] are clearly driven by homologous recombination, most frequently gene conversion. Several recombination candidate genes have been identified by homology searches in the current databases. So far, only the knockout of a RAD51 homologue had an effect on antigenic variation [26]. In yeast, RAD51 catalyses the transfer of single-stranded DNA ends into double-stranded DNA. In trypanosomes, this knockout reduced the rate of antigenic variation by both DNA recombination and in situ activation (transcriptional switching) [27]. These findings suggested that RAD51 is involved in the system, but is not essential. Moreover, its effect on the transcriptional switching between ESs is puzzling and suggests, among other possibilities, that cryptic DNA recombination could be involved in transcriptional switching. Analysis of other putative players was disappointing. Thus, ablation of homologues of enzymes involved in double-strand break repair (MRE11) [27,28], mismatch repair (MSH2, MLH1) [29] or the non-homologous end-joining pathway of DNA repair (KU70, KU80) [13] did not affect antigenic variation, even though they had various effects on recombination, telomere length and resistance to DNA-damaging agents.

ES and promoter structure

The ESs contain polycistronic transcription units whose structure had been initially modeled on only a few examples [1–5]. The recent characterisation of more ESs uncovered some variation [30], suggesting so far that only two expression-site-associated genes (ESAGs), namely ESAG6 and ESAG7, are indispensable. This finding is in keeping with the important function of the products of these genes, namely the two subunits of the transferrin receptor, and with the fact that these genes seem to be present only in ESs. It was also established that the 50 bp repeats preceding all ESs act as a boundary insulating these particular transcription units from the rest of the chromosome [31].

The ES promoters were known to contain a 70 bp sequence necessary for basal transcription [32]. In situ deletional analysis now indicated that this sequence is sufficient for the control of antigenic variation and that no extra cis-acting element is present. Thus, these promoters are very simple and do not appear to contain regulatory elements [18,33], as also supported by their efficient replacement by a ribosomal promoter [34].

Transcription: RNA processing/elongation is the key

Since the formal demonstration that RNA Pol I is active on the ESs [35*], it is clear that transcription on these sites has the unique feature of pulling together, at least functionally, the polymerase normally used to synthesize rRNA and the RNA elongation/processing machinery usually associated with RNA Pol II. The controls operating on this transcription remain controversial, but so far not a single piece of evidence has suggested promoter regulation. Instead, some evidence has argued against it, pointing to modulation of RNA elongation and processing. First, no difference of chromatin organisation was noted between promoters of active and silent ESs, suggesting similar occupancy by the transcriptional machinery [18]. Second, the silent ESs from both bloodstream and procyclic forms were found to be actively transcribed in their promoter-proximal region [36,37], giving rise to transcripts that are not polyadenylated and are retained within the nucleus [38]. Third, reporter genes inserted at the beginning of silent ESs appeared to be transcribed, but this only led to low production of functional proteins [18]. Fourth, experimental treatments mimicking the environmental conditions that trigger the cellular differentiation of bloodstream into procyclic forms showed striking opposite effects on RNA elongation over the
ESs and procyclin units [39]. Fifth, a recent report indicated that inhibition of DNA synthesis or induction of DNA damage leads to transcriptional stimulation of at least some silent ESs of bloodstream forms, with all signs of incomplete elongation [40]. Finally, the silencing of the active ES that occurs during the transformation of bloodstream forms from proliferative long slender into non-dividing short stumpy forms was shown to involve progressive in situ stalling of RNA polymerase on the ES, also pointing to regulation at the RNA elongation level [41]. Altogether, these observations suggest that environmentally controlled RNA processing on the ESs leads to

**Figure 1**

A model for the mutually exclusive control of VSG and procyclin, the respective major surface antigens of the bloodstream and procyclic (insect) stages of *Trypanosoma brucei*. An estimated 20 VSGs are located in telomeric sites, ESs, which are competent for transcription. Two families of repeats, the telomeric repeats (black circles) and 50 bp repeats (black squares) define the boundaries of the ESs. Only one ES is active at a time, ensuring the homogeneity of the surface coat. The ESs harbour polycistronic transcription units containing several ESAGs. ESAG7 and ESAG6 are the first and most conserved genes of these units, and encode the two subunits of the transferrin receptor. The procyclin genes (Pα and Pβ) are transcribed together with other genes (PAGs, for procyclin-associated genes) in four loci, only one of which is represented here. The exact genomic location of these loci is not known, but they are not telomeric. The figure represents three successive developmental stages of the parasite life cycle, namely (a) the bloodstream form (BF) slender, (b) the BF stumpy and (c) the procyclic form (PF) RNA Pol I (blue beads) seems to be permanently recruited in all units [38]. Full transcription (full green arrow) requires the association of RNA Pol I with an ’RNA factory’ (RNA elongation/processing/export machinery, represented by the red/mauve beads). This would only occur in one locus at a time, either the active ES in slender forms or the procyclin loci in procyclic forms, the other units showing abortive transcription (green dotted arrows). In bloodstream long slender forms, this machinery is concentrated in the ESB [6]. The cellular differentiation into stumpy forms would trigger the decoupling between RNA polymerase and the RNA factory [41]. Following modification induced by environmental signalling, the RNA factory would be redirected to the procyclin loci (perhaps in another body), ensuring the mutually exclusive expression of the stage-specific antigens.
to differential RNA elongation that sensitively controls the access of the VSG by RNA polymerase. Interestingly, transcription of the VSG appears to exclude that of the procyclin genes, and vice-versa. Thus, it is possible that the same transcription machinery inversely controls these transcription units (Figure 1).

At this point, the fact that at least some silent ESs are partially transcribed can be considered as generally accepted. Based on studies involving cells with multiple tagged ESs, a model has postulated the existence, besides the active ES, of a single silent ES in a ‘preactive’ state [42,43]. Therefore, the issue is now to decide if transcripts are synthesized from different silent ESs in each cell, or whether each cell within a cloned population contains a single ‘preactive’ ES that may differ between cells. Single cell PCR should allow the discrimination between these questions.

Studies on the VSG inactivation that takes place during the cellular differentiation of long slender into short stumpy forms suggested a cell cycle arrest-linked decoupling between RNA polymerase and the mRNA factory [41]. This decoupling would be relieved with the cell re-entry into the S-phase upon differentiation in procyclic forms, justifying the abrupt RNA changes that take place during this latter phase [44]. In organisms relying entirely on post-transcriptional regulation to switch genetic expression [45,46], re-coupling transcription with a post-transcriptional machinery that would have undergone signalling-dependent modification (Figure 1) would account for a global change of the pattern of transcripts, ensuring cellular differentiation.

Challenges and mysteries

So far, the control steps and the molecular players involved in both recombination and mono-allelic transcriptional control remain obscure. While it appears that trypanosomal counterparts of yeast homologous recombination factors such as RAD51 [26] are involved, others such as MRE11 [27,28] are not. Does that mean that a specialised recombination pathway is dedicated to antigenic variation? And how does the knockout of a homologous recombination factor impair transcriptional switching? This may suggest a role of cryptic recombination in this process or, alternatively, a secondary function of RAD51 or RAD51-interacting proteins could be involved. If the VSG promoter permanently recruits the basal transcription machinery, the composition of this machinery as well as that of the RNA elongation/processing complex remain to be characterised. The ESB [6] is an obvious candidate for this machinery. But then again, is this body created by the action of transcription (transcription of the active ES has to happen somewhere!) or is it a self-sustained structure, as suggested by its resistance to DNase treatment? Does the ESB have a preferential location in the nucleus and does it change during the parasitic cycle? Is there a bona fide chromatin modification other than J or is the open state of the active ES only related to the presence of RNA Pol I? Does J restrict the access of the RNA factory? Is the ESB the limiting factor that restricts the physical access to a single ES at a time [47]? Does this necessitate crosstalk between ESs and does that support the postulate of a preactive ES ready to take over [42]? Does the apparent crosstalk between ESs and procyclin loci imply that, in procyclic forms, the latter are also organised in a specialised and unique body?

Lastly, questions on the evolutionary advantages of this system still remain. Why are there several ESs? Would it be because this provides the parasite with the capability to construct new VSGs through gene conversion targeted to the active ES, and then to store these new sequences in the genome upon transcriptional inactivation of the ES? Does the presence of multiple ESs confer an extended host range through the variability of the ESG battery? Does the telomeric location of the ESs, a rule for contingency genes in many organisms, relate to better efficiency of genetic recombination and capacity for reversible silencing?

Conclusions

Completion of the T. brucei genome sequencing, expected in August 2004, will clarify several aspects of the antigenic variation system, such as the structure and number of the ESs and the VSG gene arrays. It will probably also allow the identification of new putative factors involved. The workability of RNAi and the recent development of tests enabling the assessment of ES (in)activation [38,40,48] and the measurement of DNA recombination efficiencies [26–29] should complement this effort.

Acknowledgements

The work from our laboratory was supported by the Belgian Fund for Scientific Research (FRSM and Crédit aux Chercheurs) and by the Interuniversity Attraction Poles Programme — Belgian Science Policy. LV is Research Associate at the Belgian Fund for Scientific Research.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- **of outstanding interest


36. Through a combination of in vitro transcription, immunodepletion and RNAi, this paper provides the first definitive evidence that the RNA polymerase transcribing the ES is indeed RNA Pol I.


The transformation of proliferative long slender forms into non-dividing stumpy forms was linked to the stalling of RNA polymerase on the ES. It is suggested that decoupling between RNA polymerase and the RNA factory would occur together with signalling-dependent modification of the latter.


