Control of Gene Expression in Trypanosomes

LUC VANHAMME AND ETIENNE PAYS*

Department of Molecular Biology, Free University of Brussels, B1640 Rhode Saint Genèse, Belgium

PARASITIC CYCLE OF TRYPANOSOMATIDS ................................................................. 223
SPECIAL FEATURES OF GENE EXPRESSION IN TRYPANOSOMES ............................. 224
Polystronic Transcription ....................................................................................... 224
Coupling between trans Splicing and Polyadenylation: Role of Intergenic Polypyrimidine Tracts ..... 224
Unusual RNA Polymerases .................................................................................. 225
Unusual Transcription Promoters ........................................................................ 226
Is There a Role for Promoters in DNA Replication? ........................................... 226
RNA Editing ......................................................................................................... 227
GENE EXPRESSION SPECIFIC TO THE BLOODSTREAM FORM .............................. 227
VSG and Antigenic Variation ................................................................................ 227
Anatomy of VSG Gene Expression Sites ............................................................. 227
Control of Transcription in the VSG Unit ............................................................ 227
Posttranscriptional Regulation in the VSG Unit ................................................. 228
Mechanisms of Antigenic Variation .................................................................... 229
Distant DNA Rearrangements and the (In)activation of VSG Expression Sites ...... 230
Generation of Diversity in VSG Genes ............................................................... 230
Programming of VSG Gene Expression ............................................................... 231
Expression of Genes for Glycolytic Enzymes ..................................................... 231
GENE EXPRESSION SPECIFIC TO THE PROCYCLIC FORM .................................. 232
Anatomy of the Procyclin Transcription Units ...................................................... 232
Control of Transcription in the Procyclin Units .................................................. 232
Posttranscriptional Regulation in the Procyclin Units ........................................ 233
A Need for Terminators ....................................................................................... 233
Expression of Genes for Mitochondrial Proteins ............................................... 233
INDUCIBLE GENE EXPRESSION ........................................................................... 233
Glucose Transporter Genes .............................................................................. 233
Stress Genes ....................................................................................................... 233
Histone Genes .................................................................................................... 233
TRANSCRIPTION OF RNA GENES ...................................................................... 234
rRNA Genes ....................................................................................................... 234
Spliced Leader Genes ......................................................................................... 234
Small RNA and tRNA Genes ............................................................................ 234
DIFFERENTIATION FROM THE INSECT TO THE BLOODSTREAM STAGE: THE METACYCLIC FORM ................................................................. 234
Metacyclic VSG Expression Sites ....................................................................... 234
DIFFERENTIATION FROM THE BLOODSTREAM TO THE PROCYCLIC FORM .......... 235
Replacement of VSG by Procyclin ...................................................................... 235
CONCLUDING CONSIDERATIONS ...................................................................... 235
ACKNOWLEDGMENTS ......................................................................................... 236
REFERENCES .................................................................................................... 236

PARASITIC CYCLE OF TRYPANOSOMATIDS

Many protozoan species of the genera Leishmania and Trypanosoma are parasites transmitted by specific insect vectors to mammalian hosts, in which they are responsible for several severe diseases. Leishmania species carried by sandflies cause leishmaniasis, Trypanosoma cruzi transmitted by triatomine bugs is the causative organism of Chagas’ disease, and African trypanosomes such as Trypanosoma brucei transmitted by tsetse flies are responsible for sleeping sickness in humans and nagana disease in cattle.

The parasitic cycle of these organisms is characterized by a succession of different forms adapted to the different environments they encounter but also by an alternation of growing stages adapted to infection and nongrowing stages adapted to transmission. In this respect, the cycle of T. brucei is typical (Fig. 1). In the mammalian bloodstream, slender forms of T. brucei are actively dividing cells in which the mitochondrial functions are repressed. These cells cannot engage in oxidative phosphorylation. Instead, they metabolize glucose to pyruvate via a glycolytic pathway which is partially compartmentalized within specialized organelles termed glycosomes. During this stage, the trypanosomes are covered by a uniform coat of a
variant surface glycoprotein (VSG). A continuous variation of this antigen allows the parasite to escape the defenses of the host. As the immune system eliminates the trypanosomes covered by the previous VSG, individual parasites expressing a new VSG increase in numbers, creating the observed pattern of parasitemic waves. As the infection progresses, long slender forms are replaced by nondividing short stumpy forms, preadapted to transmission in the fly. Once taken up by the fly, these cells quickly differentiate into actively dividing procyclic forms. The procyclic forms develop a fully active mitochondrion, respire on proline, and can use the Krebs cycle as a source of energy. At this stage the parasite has lost the VSG and is covered instead by a totally different glycoprotein, termed procyclin or procyclic acidic repetitive protein (PARP). Procyclic forms of the parasite migrate from the midgut to the proventriculus and then to the salivary glands, where they stop dividing. This nondividing form, called the metacyclic form, reacquires a VSG coat in a preadaptation to its future life in the bloodstream of the mammal. The stage-specific morphological and metabolic changes of African trypanosomes are reviewed by Vickerman (195) and Vickerman et al. (196).

This cycle involves transitions between differentiative and proliferative stages, in a pattern similar to that observed for the cellular differentiation of metazoan organisms. In T. brucei, the triggering of differentiation seems to be possible only within a defined window of the cell cycle, which implies an interplay between the life cycle and the cell cycle (116, 117, 143, 206). Genes presumed to encode proteins involved in the control of the cell cycle, namely, a cyclin and several cyclin-dependent kinases, have recently been cloned (2, 27, 121, 123), but nothing is known yet about their function. The expression of markers of differentiation has been the subject of numerous studies. Although knowledge concerning the genetic controls involved in the expression of the mitochondrial and glycosomal proteins is still relatively limited, this is not the case for the genetic controls operating in the expression of the major surface antigens, VSG and procyclin. Consequently, the available information concerning these controls constitutes the core of this review.

SPECIAL FEATURES OF GENE EXPRESSION IN TRYPANOSOMES

The study of gene expression in trypanosomatids has uncovered several novel processes involved in transcription. Some of them, such as trans splicing and RNA editing, have later been found operating in other organisms.

Polycistronic Transcription

In trypanosomes, the protein-coding genes are devoid of class II introns and appear to be generally packed in dense clusters containing both tandem repeats of the same (or very similar) open reading frames and interspersed unrelated sequences. Only short regions separate the successive sequences transcribed into mature mRNAs; these regions will be referred to here as intergenic sequences. The gene arrays are transcribed in long polycistronic units (17, 26, 54, 65, 124, 153, 197, 203; for recent reviews, see references 36 and 139). This particular organization probably explains why only a few trypanosome promoters have been found to date. Despite their common primary transcription, individual genes belonging to the same unit can show markedly different expression patterns (26, 65, 153). Therefore, a major conclusion is inescapable: in trypanosomes, the controls of gene expression operate primarily at posttranscriptional levels.

Coupling between trans Splicing and Polyadenylation: Role of Intergenic Polypyrimidine Tracts

Polycistronic transcription requires a cut-and-paste mechanism to process the primary transcripts into individual mature mRNAs, by the addition of a cap at the 5’ end and a poly(A) tail at the 3’ end. Capping is achieved by trans splicing. A splicing machinery similar to that involved in cis splicing in other eukaryotes—with the notable exceptions of the U1 and U5 small nuclear RNAs (snRNAs) (3, 30, 31, 99, 120, 129, 184, 198, 204)—ensures the addition, at the 5’ end of the cleavage site of the primary transcript, of a common precapped oligoribonucleotide (39 nucleotides in T. brucei). This spliced leader, improperly termed a minixenon, originates from the processing of a minixenon donor RNA (med RNA) (141 nucleotides in T. brucei), which is transcribed on arrays of repeated genes. The spliced leader can function as a U1-independent 5’ exon in heterologous splicing extracts and shows features reminiscent of the U1 snRNA of other organisms (30).

There is no known polyadenylation consensus signal in trypanosomes, and recent observations (103, 118, 169) suggest a constitutive mechanism acting at a conserved distance upstream from the splice site of the next gene (see below). Adenosine residues, in particular when repeated, seem to be preferential targets for polyadenylation (85).

For both splicing and polyadenylation, it was found that instead of a precise site, the processing may occur at several places within a short window (40, 102, 103). In addition, alternative splicing and alternative polyadenylation at clearly separate sites of the same gene environment have been documented (53, 140, 153, 193). As alternative processing generates mRNAs with untranslated regions of different lengths, this variation may be potentially important in the regulation of gene expression (see the example of the expression site-associated gene 6 [ESAG 6] transcripts, below).

The hierarchy of the processing events has been a subject of debate. On one hand, in the hsp70 transcription unit, unspliced but polyadenylated transcripts have been detected (83), pointing to an initial cleavage, linked to polyadenylation, followed by trans splicing of the downstream product. In contrast, 50% of the trans spliced transcripts of the α-tubulin genes were not polyadenylated (187), pointing to a reversed timing of events. However, these observations are not contradictory, since the hierarchy could be different for different genes, or there may be no definite timing. At the very least, these results indicated

FIG. 1. Life cycle of African trypanosomes. This cycle alternates proliferative (+) and quiescent (−) stages. Differentiation intermediates between procyclic and metacyclic forms are not indicated (dotted line).
that the two processing events occur very rapidly after RNA synthesis and are thus most probably cotranscriptional.

A recent series of observations has revealed that these events are actually coupled. The analysis of the processing of transcripts from the DHFR-TS (dihydrofolate reductase-thymidylate synthase) locus of Leishmania major (103), as well as from the tubulin, procyclin, and VSG loci of T. brucei (20, 85, 118, 169, 193), strongly suggested that the choice of the splice site, which is influenced by the presence of polypyrimidine stretches in the intergenic region, determines the site of polyadenylation of the upstream transcript (Fig. 2).

First, the distance between the polyadenylation region and the downstream splice site was found to be generally conserved, around 100 to 150 nucleotides in T. brucei and apparently more in Leishmania species (40, 85, 103, 153). This observation was confirmed by studying deletions in the sequences outside of the polypyrimidine tract of the intergenic region (103). These deletions did not affect the location of the splice acceptor site but shifted the polyadenylation site further upstream, depending on the extent of the deletion. A similar result was obtained following inversion of the polyadenylation region of the VSG gene (20, 147). This inversion created an alternative splice site for a reporter gene placed downstream, and it was found that the generation of this new splice site, located 5' to the previous one, resulted in an upward shift of the polyadenylation site of the upstream gene so that the distance between the two sites was conserved. While this distance appears to be similar in most of the cases analyzed, it is worth mentioning that some exceptions have been observed (118). In particular, alternative splicing does not seem to be obligatorily linked to alternative polyadenylation upstream (140). These cases may be explained by a cryptic processing of short intergenic transcripts, which go undetected because they are very quickly degraded.

Second, deletions or block substitutions performed in the intergenic polypyrimidine tracts were found to alter the processing of both upstream and downstream genes, leading to aberrant choices of the upstream polyadenylation site and downstream splice acceptor site and/or to reduced levels of both messengers (20, 85, 103, 118, 169, 193). These results indicated that the intergenic polypyrimidine tracts play a key role in both upstream polyadenylation and downstream trans splicing.

Third, in reporter constructs, polypyrimidine tracts involved in the determination of the polyadenylation site were able to also specify the splice acceptor site and vice versa (85).

Fourth, in several cases, the efficiency of splicing was found to depend primarily, if not exclusively, on the richness of pyrimidine in the immediate 5' environment of the gene (46, 82, 93, 212). However, exceptions to this rule have been reported, especially in cases of alternative splicing (20, 153).

Finally, in more general terms, it appears that in many cell types, the polypyrimidine tracts of RNA are preferential targets for interactions with the proteins involved in posttranscriptional modulation of gene expression (119, 130).

Taken together, these results suggest that the RNA-processing machinery binds to the RNA through interaction with the intergenic polypyrimidine tracts, probably as soon as these transcripts are synthesized. The cleavage of the primary transcripts is followed by trans splicing downstream and polyadenylation upstream, the latter process possibly occurring after scanning of the 3' end released by the splicing event (103).

Unusual RNA Polymerases

The three classical RNA polymerases, identified on the basis of their sensitivity to the drug α-amanitin, have been detected in trypanosomatids (41, 42). The major subunit of each of these enzymes was characterized by gene cloning (55, 91, 94, 173). In most species of trypanosomes undergoing antigenic variation, with the exception of T. vivax (172), two slightly different genes for polymerase II (pol II) were found. The largest subunit of the trypanosomal pol II is devoid of the C-terminal heptapeptide repeats characteristic of this enzyme in other eukaryotes. However, this absence does not prevent the C-terminal phosphorylation, which is known to be involved in the regulation of pol II activity in other species (33, 131).

The nature of the RNA polymerase at work on the VSG and procyclin transcription units has been the subject of considerable attention. In both cases, this polymerase was found to be resistant to α-amanitin, a hallmark of the ribosomal polymerase (pol I) (38, 95, 97, 141, 159). The identification of this polymerase as pol I was further supported by its lack of release by the detergent Sarkosyl in run-on transcription assays (163), by the in situ localization of the nascent procyclin transcripts around the nucleolus (34, 161), and by the structural similarities between the VSG, procyclin, and ribosomal promoters (see below). In addition, it has been demonstrated that pol I may be used to synthesize transcripts for protein-coding genes in trypanosomes (161, 208). This is possible because transcription initiation in trypanosomes, unlike in other eukaryotes, is uncoupled from RNA capping. A reason for the use of this particular polymerase to transcribe the genes for the major surface antigens, which represent by far the most abundant
proteins of the cell, may relate to the high efficiency of transcription by pol I.

However, some observations are not in agreement with this interpretation. First, on the basis of its catalytic properties in run-on transcription assays, in particular the requirement for divalent cations (Mg$^{2+}$ and Mn$^{2+}$), the polymerase transcribing the VSG genes behaves as pol II and not pol I (76, 147). Second, the nascent transcripts of the VSG unit are not present in the nucleolus but, rather, in another discrete location of the nucleoplasm (24). Therefore, the polymerase for the VSG and procyclin units appears to share properties with both pol I and pol II. These contrasting properties cannot be easily explained by a mixing of pol I and pol II subunits, since the catalytic domain and the $\alpha$-aminatin-sensitive regions of pol II are present on the same subunit. In addition, these features cannot be ascribed to the few differences that exist between the two copies of the pol II largest-subunit gene. Indeed, the removal of one of these two genes by targeted homologous recombination did not affect the characteristics of $\alpha$-aminatin sensitivity in run-on transcription assays (35). This issue should be resolved by the development of in vitro transcription assays, as well as by the production of antibodies that specifically affect the activity of each of the three polymerases.

**Unusual Transcription Promoters**

Only a few promoters have been characterized in trypanosomes: those of the ribosomal, procyclin, and VSG genes of *T. brucei* (38, 70, 86, 90, 141, 162, 170, 201, 212), and those for some small RNA genes (see below). In addition, a cis-acting regulatory element similar to a promoter (this element is able to determine the strandedness of transcription) has been found upstream of a multidrug resistance gene in *L. enriettii* (202). There is no significant sequence homology between these promoters or with any known eukaryotic ribosomal promoter. In the case of the VSG promoter, the 70-bp region preceding the transcription start site is sufficient to ensure maximal activity in transient expression assays of a reporter gene construct (89, 191, 211). In contrast, in the case of the ribosomal and procyclin promoters, the $-70$ to $+1$ region constitutes a core element whose basal activity is stimulated by an upstream control element located around position $-200$ (28, 86, 170). Apart from this difference, the VSG promoter possesses similar structural features to the core element of the procyclin and ribosomal promoters (Fig. 3). They all contain two stretches, centered approximately at $-60$ and $-35$, that are essential for promoter activity and whose spacing appears to be critical (28, 86, 170, 192). Moreover, despite a lack of homology, one of these stretches appears to bind common factors (192). Interestingly, the binding of specific proteins requires that the target DNA be single-stranded (29, 192), and the binding of proteins common to the three promoters is not observed on double-stranded DNA (87, 192). These results suggest that these promoters should be partially denatured to be functional (29, 192). Proteins able to bind to specific sequences of single-stranded DNA have also been found in *L. major* (199) and *Crithidia fasciculata* (1, 186). However, in contrast to the promoter-binding proteins discussed above, in the last two cases no evidence was found for a role of protein-DNA interaction in gene expression (1, 200).

The characteristics of the VSG and procyclin promoters are reminiscent of properties shared by the eukaryotic ribosomal promoters (126, 174): the absence of sequence conservation between species, a core element containing two important boxes, the importance of the spacing between these elements, and the presence of an upstream control element in the procyclin promoter. These structural similarities were recently strengthened by the demonstration of a functional relationship between the VSG, procyclin, and ribosomal promoters. It was found that the activity of hybrid constructs of these promoters was relatively conserved in transient expression assay of a reporter gene, which implies that their specific protein-binding regions are interchangeable (86, 192).

**Is There a Role for Promoters in DNA Replication?**

Recent work by Cross and coworkers has given the first indication of the mechanism of DNA replication in *T. brucei*. These studies involved the construction of a plasmid that could be used as replication origin trap (135, 136). Restricted genomic DNA was inserted in this plasmid in front of a drug resistance gene placed under the dependence of the procyclin promoter. This gave rise to several plasmids that behaved as stable autonomously replicating episomes. The promoter was then deleted from these constructs, to transform them into transcription promoter traps. These experiments showed that the procyclin promoter was specifically required for DNA replication, as if the presence of this promoter immediately downstream from the entrapped genomic sequence was essential for plasmid maintenance (135). This observation may be linked to the necessity for opening the chromatin structure around the replication origin. This requirement would be ensured by the contiguous presence of a strong transcription promoter.

FIG. 3. Comparison of the core element of the VSG, procyclin, and ribosomal promoters of *T. brucei*. The $-70$ to $+1$ region of the bloodstream VSG promoter (VSG B) is sufficient for maximal activity, while the corresponding region of the ribosomal (RIB) and procyclin (PRO) promoters confers an activity stimulated by the presence of upstream elements. The sequences are aligned on the transcription start sites (arrows). The boxed regions have been defined as crucial in transient-activity assays (RIB [86], PRO [28, 170], and VSG [192]). The bottom sequence (VSG M) is that of the AnTat11.17 metacyclic VSG promoter (192); the sequence elements shared with the crucial regions of the bloodstream VSG promoter are doubly underlined.
RNA Editing

The mitochondrial DNA of trypanosomes, termed kinetoplast DNA, consists of a large network of intercated mini- and maxicircles. Several genes for mitochondrial proteins are encoded by the maxicircles, but these genes are generally incomplete. The generation of functional mRNAs involves the posttranscriptional addition or deletion of U residues, a process known as RNA editing. This process is mediated by small RNAs called guide RNAs encoded on both mini- and maxicircles. Although RNA editing is subjected to stage-specific regulation, this process does not seem to be involved in the control of nuclear DNA-encoded gene expression and will not be discussed here. Several recent and extensive reviews deal with this matter (see reference 18 and references therein).

GENE EXPRESSION SPECIFIC TO THE BLOODSTREAM FORM

Numerous proteins are expressed exclusively or primarily in the bloodstream form. This is the case for the VSG and certain glycolytic enzymes.

VSG and Antigenic Variation

*T. brucei* survives in the bloodstream of its mammalian host as an extracellular parasite subjected to constant attack by the immune system. The main function of the VSG is to ensure a protection against this attack. The parasite surface is covered with a dense coat of 10^7 VSG molecules attached to the plasma membrane by a glycosylphosphatidylinositol (GPI) anchor. Although the genome contains more than 1,000 VSG genes, a single gene is expressed at any given time. Variation of this VSG occurs spontaneously at a rate of up to 10^{-2} per cell and per generation (185). Therefore, each time a population expressing a VSG is eliminated by the immune system, it is replaced by cells expressing another VSG, giving rise to the development of long-lasting, chronic infections.

Anatomy of VSG Gene Expression Sites

While most VSG genes are not telomeric, many are present in telomeres, particularly of minichromosomes (reviewed in reference 139). However, the expressed VSG gene is always located in a telomere (47). The number of telomeric VSG expression sites is probably around 20, with only one being active at a given time. Only a few VSG expression sites have been analyzed in any detail, and they appear to share similar features, although some variations are noticeable (7, 44, 45, 50, 58, 64, 98, 110, 146, 152, 158, 171). In particular, a much shorter VSG unit has recently been described (6).

Several VSG expression sites consist of a 45- to 60-kb transcription unit from the promoter to the end of the VSG gene. The length variation appears to depend on the dispensable presence of a 13-kb element at the 5’ extremity (see below) (Fig. 4). The VSG gene is located 5 to 10 kb from the chromosome end and is surrounded by two “barren” regions devoid of restriction sites. While the 3’ barren region essentially consists of telomeric repeats (TTAGGG), the 5’ barren region is
made of a variable array of noncoding 76-bp repeats which seem to behave as recombinational hot spots (111). This unit is polycistronic and also contains several associated genes (ESAGs). In the case of the AnTat 1.3A expression site, which can be taken as a model since it has been almost completely sequenced (7, 110, 146, 152), there are nine ESAGs. All ESAGs are subjected to the same transcriptional control as is the VSG gene, and they are thus expressed exclusively in the bloodstream form.

The function of some of these genes has been recently clarified. ESAGs 6 and 7 are very similar genes encoding the two subunits of a transferrin receptor (108, 165, 167, 177). This heterodimer may be attached to the plasma membrane through a GPI anchor present at the C terminus of pESAG 6. However, the receptor appears to be rapidly released from the surface membrane and accumulates, together with transferrin, in the lumen of the flagellar pocket, an invagination of the plasma membrane at the base of the flagellum (165, 177). The mechanism of transferrin uptake is still unclear, although both transferrin and the pESAG 6/7 complex are internalized via coated pits and vesicles. Interestingly, ESAG 6 and 7 share significant sequence homology with some VSG genes (81), suggesting that the trypanosomal transferrin receptor evolved from a VSG. Many copies of ESAG 6 and 7 are present in the genome. It is not known if all the members of this family are contained within VSG expression sites, but no VSG-independent transcription of these genes has been detected to date (146).

ESAG 4 encodes a receptor-like transmembrane protein with adenylate cyclase activity (134, 146, 158). This gene also belongs to a multigene family, some members of which, termed GRESAG 4.1 to 4.3 (for genes related to ESAG 4), are expressed independently of the VSG (8). The genomic linkage between ESAG 4 and the VSG gene may be related to the functional relationship which exists between adenylate cyclase and the VSG. Indeed, all of the experimental or physiological conditions which induce the release of the VSG also appear to activate the cyclase (156, 157). To date, the mechanism by which the presence of the VSG leads to repression of the cyclase is not understood.

ESAG 8 is present as a single or two slightly divergent copies, depending on the expression site. This gene encodes a soluble protein with a putative nuclear localization signal (152, 171). Evidence for nuclear targeting of pESAG 8 was obtained with the Xenopus oocyte system (109). This protein almost certainly has a regulatory role, since it contains two clearly defined domains: a special class of zinc fingers termed the ring finger, which possibly interacts with a nucleic acid, and a leucine repeat domain, which is probably involved in interactions with a protein. While all attempts to detect interactions between pESAG 8 and DNA or RNA have failed so far, evidence for interaction with a protein has been obtained recently with the yeast “two-hybrid” system (149).

While the role of ESAG 5 is totally obscure, the nucleotide sequences of ESAG 1, 2, and 3 suggest that these genes encode either GPI-anchored or transmembrane surface proteins, similar to other minor invariant surface glycoproteins (7, 45, 132, 205).

The nine ESAGs mentioned above seem to be present in the same order in several other VSG expression sites. However, additional ESAGs have been described. Since these genes are absent from the AnTat 1.3A expression site, their presence is apparently dispensable. Such is the case for ESAG 9, whose product shows some homology with GTP-binding proteins (58), and ESAG 10, which encodes a putative surface protein with multiple transmembrane domains (72). Interestingly, ESAG 10 is contained in the 13-kb transcription unit occasionally found at the beginning of the VSG expression site (70, 72, 210). Like ESAG 4 and 2 (8, 19), ESAG 9 and 10 belong to families some of whose members are expressed independently of the VSG (58, 72).

Recently, a novel type of VSG expression site was identified during a study of the reexpression of a metacyclic VSG in the bloodstream form late in the infection (6). The expressed VSG gene was found to be in a telomere lacking the 5′-barren region and was flanked by ESAG 1 only. In this case, the promoter was located only about 2 kb upstream from the VSG gene, and the ESAG 1 copy was apparently transcribed from another promoter. Thus, this VSG transcription unit was found to be monocistronic. All these characteristics point to a reactivation of a metacyclic expression site in the bloodstream form (see below). At this stage, it is not known if this reactivation is exceptional or how it occurred. It may be used as a salvage pathway when the supply of regular VSGs is exhausted very late in infection.

Control of Transcription in the VSG Unit

In the bloodstream form, all but one VSG expression sites are repressed. So far, it is not known if this repression is controlled at the level of transcription initiation or of premature termination, since the extent of transcription at the beginning of these sites has not been accurately determined. This analysis is rendered difficult by the very high sequence conservation between expression sites in this region.

In the procyclic form, the VSG unit is totally down-regulated. However, run-on transcription assays and cDNA analysis have shown that the VSG promoter is still active at this stage (140, 141, 160). This is consistent with the observation that the VSG promoter is fully active in the procyclic form when present in episomal plasmidic constructs (90, 211, 212) or when inserted in the nontranscribed spacer of the ribosomal DNA (160). However, the extent of activity of the VSG promoter in its genomic environment is difficult to assess. The apparent constitutive transcription at the beginning of the VSG expression site in the procyclic form may actually be due to the addition of a weak transcription from all the VSG promoters at the same time. This view is consistent with the relatively low level of expression exhibited by reporter constructs when targeted to individual expression sites (160). Furthermore, the procyclic transcripts specific to the VSG expression site show sequence heterogeneity, indicating that different VSG transcription units are active at this stage (160). Thus, it appears that transcription initiation in the procyclic form occurs from several, possibly all, VSG promoters. However, it is not known to what extent this initiation is controlled, if at all: as discussed below, the weak transcription of the VSG units may be accounted for by a down-regulation of RNA elongation.

In contrast, it is clear that the genomic environment influences in a specific way the level of transcription driven by the VSG promoter. Indeed, if a ribosomal promoter is targeted to the VSG expression site, the transcription occurring from this promoter in the procyclic form is not repressed (160). A major control for the down-regulation of transcription in the VSG expression site operates at the level of RNA elongation, which is blocked in nuclei from the procyclic form (35, 141, 160, 191). It is possible to detect transcription of regions close to the promoter, but this decreases progressively and cannot be observed at all several kilobases downstream. The arrest of RNA elongation does not seem to be due to a bona fide terminator, although a major attenuation site has been mapped 700 bp downstream from the VSG promoter (160). Interestingly, the
progressive arrest of RNA elongation is also observed in nuclei from bloodstream forms incubated at 20 to 27°C instead of 37°C (7, 98, 141, 191), even for short periods. This indicates that the rate of RNA elongation on the VSG unit is strongly dependent on temperature. Other factors down-regulating RNA elongation in this unit include the components necessary for the differentiation of bloodstream forms into procyclic forms, such as citrate/cis-aconitate (191). The control of transcription by these environmental conditions is very rapid (191), suggesting an early involvement in the differentiation of the parasite. Altogether, the down-regulation of the VSG unit in the procyclic form appears to result from either the loss of an elongation factor or a chromatin repression in this locus. The lack of down-regulation observed when the VSG promoter is not present in its original environment suggests that the factors responsible for repression are not diffusible. In addition, the lack of repression observed when a ribosomal promoter is inserted in this unit suggests that the RNA polymerases recruited by the VSG and ribosomal promoters are controlled by different repressive factors. Interestingly, however, evidence for a temperature- and citrate/cis-aconitate-dependent control of ribosomal DNA transcription, similar to that operating in the VSG unit, has been obtained recently in our laboratory (191).

**Posttranscriptional Regulation in the VSG Unit**

Although the ESAGs and the VSG gene are transcribed together in the bloodstream form by the same polymerase, the steady-state amounts of the ESAG and VSG mRNAs show considerable differences (45, 50, 146). In particular, the VSG mRNA represents by far the most abundant polyadenylated transcript of the cell. A differential rate of RNA maturation and/or degradation must account for this observation. However, the sequences and mechanism responsible for this phenomenon are not characterized. As discussed below for the phosphoglycerate kinase genes (93), the relative richness in pyrimidine of the 5’ environment of each gene may account for a differential rate of trans splicing, which influences the final level of mRNA. In addition, the sequence of the 3’ untranslated region (3’-UTR) may be critical to determining the relative amount of the mRNA: this seems to be particularly true for the VSG mRNA (see below).

The transcripts which are still synthesized in the VSG unit during the procyclic stage appear to be particularly unstable, because their steady-state levels are very low unless RNA degradation is prevented by UV irradiation (39, 40, 140). This phenomenon may be due to a stage-specific regulation of the processing of the transcripts. Two observations support this hypothesis.

First, the ESAG 6 transcripts present in the procyclic form were found to be abnormally spliced (40, 140). In the bloodstream form, alternative splicing generates two ESAG 6 transcripts at approximately similar frequencies as judged by the analysis of RNA stabilized following UV irradiation. However, only the transcripts spliced very close to the gene initiation codon are stable and accumulate in the cytoplasm. In the procyclic form, the gene-proximal splice site is not used, so that only unstable ESAG 6 RNAs with a long 5’ untranslated region (5’-UTR) are produced. The splicing of the ESAG 6 mRNA is thus stage regulated.

Second, it was found that the 3’-terminal region of the VSG mRNA is able both to enhance the steady-state level of a reporter mRNA in the bloodstream form and to decrease this level in the procyclic form (21, 90). This indicates that the nature of the 3’-UTR is of key importance in determining the relative stage-specific abundance of the mRNA. The available data indicate that different mechanisms operate in the bloodstream form and in the procyclic form. In the bloodstream form, the presence of the 3’-terminal region of the VSG mRNA led to an increase in the stability of the transcripts from a reporter gene, and this effect was compatible with the increase in mRNA abundance (21). In contrast, this sequence had no effect on RNA stability in the procyclic form, suggesting that the reduction of mRNA levels at this stage is due to a negative influence of the 3’-terminal sequence on the rate of RNA maturation (21).

**Mechanisms of Antigenic Variation**

The *T. brucei* genome contains more than 1,000 VSG genes, 10 to 20 VSG expression sites, but only a single active VSG transcription unit. Thus, antigenic variation must occur either by gene replacement in the active expression site or by the alternative transcriptional activation of different expression sites. Both mechanisms have been observed (Fig. 5) (for recent
reviews, see references 25, 43, 147, and 189).

Gene replacement is achieved by recombinational events changing all or part of the expressed VSG gene in the active site. A few cases of reciprocal recombination between telomeres have been observed, but the most common mechanism for gene replacement is gene conversion, in which the active VSG gene appears to be chased by the copy of another gene. The extent of gene conversion is quite variable: it can range from a few base pairs inside the VSG gene, which leads to the generation of a chimeric gene, to dozens of kilobases, which leads to the conversion of large telomeric segments (137). The length of the conversion domain seems to be dictated by the location of the regions of homology between the donor and target sequences. Thus, these events appear to be a manifestation of the powerful homologous recombination potential revealed by DNA transfection experiments (52, 89, 106, 178, 179). As the different telomeric expression sites appear to share extensive homology, gene conversion events may preferentially occur between expression sites. More generally speaking, the common sequence elements present in all telomeres seem to favor recombination between them, leading to the frequent use of telomeric VSG genes as gene conversion donors early in the parasitic infection (100). The 76-bp repeats of the 5’ barren region are frequently used as 5’ gene conversion endpoints, since they are the most common sequences found between VSG genes. The conserved 3’ end of the VSG gene often represents the other endpoint (137).

The alternative transcriptional activation of expression sites is termed in situ (in)activation and occurs without detectable rearrangement inside or immediately adjacent to the VSG transcription unit. In particular, this phenomenon is not due to a change of the DNA sequence of the transcription promoter, since this sequence is unaltered upon (in)activation (212). Moreover, several complete and potentially functional VSG promoters appear to coexist in the trypanosome genome, probably at least one in each VSG expression site, despite the fact that a single site is expressed at a time (89, 212). Therefore, the genomic environment of each promoter, perhaps influenced by distant rearrangements, seems to be crucially important in achieving the selective activation. Recent observations offer potential answers to this issue and will be discussed in the next two sections.

**Distant DNA Rearrangements and the (In)activation of VSG Expression Sites**

Upon either (in)activation of a VSG expression site or its involvement as a donor in gene conversion, large size variations of the telomeric repeats have been observed (127, 145). These variations probably reflect the existence of recombinations between telomeric repeats. The possible impact of the telomeric rearrangements on the control of the activity of VSG expression sites is worth mentioning. Indeed, it has been recently reported that in human telomeres, short and long arrays of repeats have completely different chromatin structures (181). Moreover, the transcription of telomeric sequences in yeast cells is known to be under unstable control, which is affected by the distance from the chromosome end (166). However, the demonstration that recombination within the telomeric repeats is involved in the (in)activation of the VSG units may prove difficult, given the sequence simplicity and conservation of these regions.

At the other extreme of the VSG unit, DNA rearrangements occurring 13 kb upstream from the promoter were found to be linked to the inactivation of a VSG expression site (70, 210). These rearrangements were due to homologous recombination between two VSG promoter copies located 13 kb apart, resulting in the loss of one of these promoters together with the intervening 13-kb region (71). The latter contained ESAG 10 as well as a transposable element (72, 112). Later reactivation of this expression site was again linked to rearrangements in the same region. Approximately half of the expression sites appear to be preceded by the additional 13-kb unit (71, 112), so that the type of rearrangement described above may occur frequently. Although these observations indicate a possible role of far-upstream control elements in the regulation of transcription, it is worth stressing that the (in)activation of VSG expression sites is also feasible without any detectable rearrangements (210).

**Involvement of a Novel Nucleotide?**

About 10 years ago, the partial resistance of the trypanosome telomeric DNA to digestion by the restriction endonucleases PstI and PvuII was interpreted as being due to a novel type of DNA modification (22, 142). Interestingly, this modification was lost reversibly upon transcriptional activation of the telomere and was specific to the bloodstream form of the parasite. A clue to the nature of this modification was found by Borst and coworkers, who reported the existence of a novel nucleotide, β-D-glucosylhydroxymethyluracil, in the telomeric DNA of the trypanosome bloodstream form only (67–69). This unusual molecule represents only 0.1% of the total nucleotide content. Whether the presence of the modified base is related to VSG gene inactivation is still unclear. First, it is not known if this base is absent from the transcribed telomeres. Second, this base does not easily account for the partial endonuclease digestion, as it is difficult to explain why the presence of a modified version of thymidine would lead to preferential protection of the PstI and PvuII sites (CTGCAG and CAGCTG, respectively), although several other sites such, as HindIII, SalI, or SphI, also show low levels of protection (142). As the extent of modification appears to depend on both the proximity and size of the arrays of telomeric repeats, it was proposed that the phenomenon could be triggered by the binding of a modification machinery to these repeats followed by a scanning toward the chromosome center (22, 142). Current speculation implicates the DNA modification in possible controls of the telomeric chromatin organization and/or interactions with the nuclear matrix, similar to controls involved in the “position effects” which influence transcription in telomeres of yeast cells (for a discussion, see reference 25). However, whether the modification is the cause or the consequence of gene (in)activation is not known. In this respect, it is worth mentioning that the modification is not required for inactivation of the VSG unit in the procyclic form.

**Generation of Diversity in VSG Genes**

By means of an in vitro selection with monoclonal anti-VSG antibodies, it was found that single point mutations within VSG genes may achieve antigenic variation (12). Recent results by Lu et al. have strengthened these experimental observations, by actually demonstrating that point mutations can be generated during gene conversion (114). The gene conversion events examined by these investigators were linked to the generation of VSG sequences for which no identical donor could be detected, even partially. Analysis of three independent duplications of the same VSG gene showed that the point mutations exhibited a strand bias, suggesting that they were not completely random (113). Two explanations may account for these observations: either an error-prone polymerase is at work during the gene conversion process, or the structure of
the donor DNA induces mistakes in copying. Some considerations tend to support the second hypothesis. Indeed, point mutations are not always generated during gene conversion. This was not observed in several events in which nontelomeric VSG genes were used as donors (92, 144, 180). It may therefore be proposed that only the conversion events involving telomeric donors lead to point mutations, although this conclusion should be substantiated by more analyses. If confirmed, this may provide a clue to the resolution of this issue. Indeed, the DNA modification characteristic of the silent telomeric VSG genes could be responsible for a misreading by the polymerase in the gene conversion events that are templated by telomeric VSG genes. This hypothesis has the merit of also providing a possible explanation for the rapid evolution of telomeric VSG genes (59); these genes should preferentially accumulate mutations. In this respect, it is not clear whether the presence of the modified nucleotide can affect the fidelity of DNA replication at telomeres, independently of gene conversion. This question is clearly amenable to experimental verification by comparing the sequences of telomeric and nontelomeric silent VSG genes as a function of cell division.

A mechanism distinct from point mutation also contributes to the evolution of the VSG gene repertoires. Through segmental gene conversion within VSG genes, chimeric genes of a novel type are created by sequence reassortment (92, 144, 180). These new genes may exhibit a complex mosaic structure if numerous donors are involved in the reassortment, which often seems to be the case late in chronic infections (13, 138, 180). This process is able to complement defects in pseudogenes, which allows these sequences to contribute to antigenic variation and thus to extend the size of the antigen repertoires (13, 138). The intragenic recombinational variability appears to be extensive, but it may be limited by some constraints. Indeed, a bias has been observed in the location of the endpoints of gene conversion involving highly homologous genes as donor and target sequences (144). These endpoints were located preferentially within the segment encoding the hinge region between the two domains of the antigen. Therefore, recombination is probably limited by phenotypic constraints. One may imagine that recombinations leading to the exchange of domains between antigens are less detrimental to the generation of successful VSGs than are recombinations within each domain. These observations suggest that the frequency of recombination in the VSG expression site is probably higher than the measured rate of antigenic variation.

**Programming of VSG Gene Expression**

The VSGs are not expressed completely at random during the development of infection. This relative programming is probably due to both genetic and phenotypic constraints.

The frequency of expression of any given VSG gene seems to depend primarily on its probability of recombining with the active expression site, which in turn depends on the degree of homology between the recombination partners (137, 138, 180). This conclusion is particularly illustrated by the fact that genes expressed infrequently or late during infections can be converted into genes expressed early if they are translocated into a telomeric environment showing more homology to the expression site (101). The case of the ArTat 1.1 VSG gene is especially pertinent in this regard. This gene, which is normally expressed late, is telomeric but present in the reverse orientation with respect to the chromosome end. It was found that when this gene was inserted correctly in a telomere, it became predominantly expressed, clearly indicating that the genomic environment, in this case the telomeric repeats, influence the probability of expression (190). When a VSG gene is located in a telomere, the homology of its environment with the telomeric expression site may favor its recombination. This probably explains why telomeric VSG genes are frequently expressed early during infection (180). Similarly, the probability of expression of nontelomeric VSG genes may depend on the presence and number of 76-bp repeats in their 5′ environment. Indeed, these repeats represent the usual homology regions and recombination sites between the VSG donor sequences and the 5′ barren region of the expression site (111).

In addition to the genetic constraints, a phenotypic selection may operate in determining the immediate succession of variants. This selection may occur at the transition between the old and the new VSG during the short period when these antigens are simultaneously present on the cell surface. The viability of the transient double expressors, and thus the selection of the new variant, would depend on the ability of the different VSGs to pack together, because tight packing of the VSG coat is required to protect the parasite against lytic elements from the blood. Mathematical models predict that a variation of this compatibility can lead to VSG programming (5). Although this hypothesis remains equivocal, it is highly likely that not all VSGs can form compatible pairs. The VSGs adopt grossly similar folding patterns (23), but several structural groups have been identified on the basis of their pattern of conserved cysteine residues, and combinations between groups may not be equally successful (32).

**Expression of Genes for Glycolytic Enzymes**

Bloodstream trypanosomes are exclusively glycolytic organisms, whereas procyclic forms respire primarily on carbon derived from amino acids such as proline and threonine. Consequently, certain key enzymes in these two very different metabolic pathways display different levels of activity in both life cycle forms. The genes encoding the glycolytic enzymes are interspersed in multigenic transcription units with either related or unrelated genes, and different levels of control ensuring a developmental regulation of expression have been identified.

In the case of fructose bisphosphate aldolase, the steady-state mRNA levels in bloodstream and procyclic forms parallel the levels of enzyme activity (37). As the transcription rate is very similar at both stages, the observed sixfold difference in mRNA level must be related to posttranscriptional controls. The 3′-UTR of the mRNA, but not the splice acceptor site, has been found to influence the level of expression of a reporter gene in transient-activity assays (84). It is not known if this effect occurs through a modulation of the amount of RNA.

Similar observations have been made for the phosphoglycerate kinase genes (65). The locus containing the three genes (PGK A, PGK B, and PGK C) is transcribed at the same rate in both developmental stages, but the genes show a markedly differential expression: the PGK A mRNA level is low in both bloodstream and procyclic forms, whereas the PGK B and PGK C mRNAs are highly expressed but in the procyclic form and the bloodstream form, respectively. Although the poor efficiency of the splice acceptor site of the PGK A gene probably accounts for its low expression in both forms of the parasite (93), it is not known what determines the differential expression of the PGK B and PGK C genes. As this regulation must operate during RNA processing, intergenic sequences are likely to play an important role.
GENE EXPRESSION SPECIFIC TO THE PROCYCLIC FORM

The procyclic form, multiplying in the insect digestive tract, is characterized by the derepression of the mitochondrion and the replacement of the VSG by procyclin. Like the VSG in the bloodstream form, procyclin and its mRNA represent the most abundant components of their developmental form (38, 154, 155). The procyclin units are subjected to several regulatory controls (reviewed in reference 79).

Anatomy of the Procyclin Transcription Units

Procyclin genes are arranged in two diploid, nontelomeric, and unlinked loci (PARP A and PARP B), each containing two genes in direct repeat (Fig. 6). In some strains, a third gene is present in one of the loci. Slight sequence differences between the open reading frames and major differences between their UTRs allowed the demonstration that both loci are expressed (38, 95). As mentioned above, the transcription of these units is mediated by a polymerase resistant to α-amanitin. The procyclin transcription units are polycistronic, and their length is relatively short: probably around 10 kb for PARP B (38, 162) and 8 kb for PARP A, in which the termination region has been identified (19a). Apart from the procyclin genes, the PARPB locus contains an additional gene termed PAG1 (for procyclin-associated gene 1) (96). This gene shares sequence homology with ESAG 7 and 6, PAG 2 and PAG 3 share a long 5′-UTR with the longest transcript of PAG 1, and GRESAG 2.1 is related to ESAG 2. The size of the PARP A locus is around 8 kb, while that of PARP B is still undefined but probably does not exceed 10 kb. α am-s and α am-r, α-amanitin-sensitive and -resistant transcription, respectively.

Control of Transcription in the Procyclin Units

Although the procyclin units are expressed only in the procyclic form, nuclear run-on transcription assays have indicated that the procyclin promoter is still active in the bloodstream form (141). The minimal level of this activity appears to be between 5 and 15% of that measured in procyclic forms, but this value is biased by the strong transcription attenuation which occurs in this unit at that stage (see below). The observation of a constitutive activity of the procyclin promoter is consistent with the results of transient-activity assays of a reporter gene placed under the dependence of this promoter. In the bloodstream form, this activity appears to be equivalent to that of the VSG promoter (88). Through a homologous recombination-mediated targeting of a reporter construct in the procyclin promoter, it was confirmed that the primary transcription at the beginning of this unit is about 10-fold lower in the bloodstream than in the procyclic form, in the absence of any selection to force this activity (191). As for the VSG unit in the procyclic form, evidence was obtained that RNA elongation on the procyclin unit is progressively blocked in the bloodstream form (191). Remarkably, the factors responsible for this effect were the same as those controlling RNA elongation in the VSG unit: these factors include temperature and components known to trigger differentiation, such as citrate/cis-aconitate (191). Thus, opposite but related controls of RNA elongation operate on the procyclin and VSG units at each stage of the parasite life cycle.
Posttranscriptional Regulation in the Procyclin Units

In addition to RNA elongation control, posttranscriptional regulation probably accounts for the large difference in the level of mature procyclin mRNA between the bloodstream and procyclic forms. The 3′-UTR of the procyclin mRNA contains elements able to play such a role. As a whole, the presence of this 3′-UTR leads to an increase in the amount of a reporter mRNA in the procyclic form (21). Furthermore, a 16-mer sequence conserved among the different 3′-UTRs of the procyclin mRNAs, as well as in the 3′-UTR of an unrelated mRNA for a major surface antigen of the procyclic form of *T. congolense* (16), was found to be required for the efficient expression of a reporter gene in the procyclic form (80). This sequence is involved in the formation of a stem-loop structure whose conservation appears to be essential. As the deletion of the 16-mer does not affect the mRNA level, its polyadenylation, or its nucleocytoplasmic distribution, it may be required for efficient translation of the mRNA. However, the involvement of this element in the control of procyclin expression remains to be demonstrated.

A Need for Terminators

The available data suggest that the procyclin loci are pol I transcription units located between pol II units. Therefore, specific transcription terminators must be present to prevent interference between polymerases of neighboring transcription units. Evidence for such an interference was obtained through the insertion of a procyclin promoter in the tubulin locus. In bloodstream forms of these transgenic trypanosomes, the promoter was not used by the α-amanitin-resistant RNA polymerase but was bypassed by the α-amanitin-sensitive polymerase from the tubulin locus (21). In contrast, in procyclic forms, it was able to stop upstream transcription and to recruit its normal RNA polymerase (21). These observations suggest that the procyclin promoter is associated with a pol II terminator, whose strength may vary between developmental stages. In addition, a terminator present at the other end of the PARP A unit has recently been characterized in our laboratory (19a).

Expression of Genes for Mitochondrial Proteins

In bloodstream slender forms, the mitochondrion lacks cristae and detectable cytochrome-mediated respiration. The mitochondrial activation starts with the arrest of cell proliferation and the differentiation into stumpy forms. When the stumpy forms are ingested by the fly and transform into procyclic forms, the mitochondrion expands, the cristae become discoid, and the oxidative metabolism becomes functional (195). These adaptations in energy metabolism are accompanied by a regulation in the expression of the mitochondrial proteins. These proteins are encoded by both nuclear and mitochondrial genes. The few data available suggest multiple levels of control for both sets of genes. Only data concerning nuclear genes will be discussed here.

Cytochrome c reductase subunit 4 and cytochrome c are not detectable in bloodstream long slender forms, while their mRNAs are present at approximately 20 and 30%, respectively, of the levels found in procyclic forms (150, 183). This observation suggests a double control at both transcriptional and posttranscriptional levels. In the case of cytochrome c, regulation appears to operate primarily at the level of the stability of the protein, since the mRNA is equally translatable at both stages but the half-life of the protein is considerably reduced in the bloodstream form (182). This observation probably relates to a defect in the subcellular targeting of the protein during this stage, since it is likely that mitochondrial proteins cannot assemble properly in the repressed organelle.

**INDUCIBLE GENE EXPRESSION**

While it is clear that the majority of the genes studied to date in trypanosomatids show developmental variation of the steady-state amount of their mRNAs, only a few examples are known of genes whose expression can be physiologically modulated in a given stage.

**Glucose Transporter Genes**

The glucose transporters of *T. brucei* are encoded by two gene families (*THT1* and *THT2*) (26). The THT locus contains six direct repeats of *THT1* followed by five direct repeats of *THT2*. All genes are probably transcribed as a polycistronic unit. *THT1* genes are highly expressed in bloodstream forms but are repressed in procyclic forms. *THT2* genes are moderately expressed in bloodstream forms but appear to be regulated by external signals at the procyclic stage. Glucose starvation triggers a 60-fold increase of the *THT2* mRNA level (26). The mechanisms underlying this regulation are still unknown, although posttranscriptional controls are likely to be important.

**Stress Genes**

Heat shock proteins (HSPs) must play important functions in parasitic organisms such as trypanosomes, whose life cycle imposes successive shifts from low temperature in insects to 37°C (or more during high parasitemia) in the bloodstream. In different trypanosomatids, the *hsp* genes are arranged in arrays of tandem repeats; for example, the *hsp70* and *hsp83* genes are present in 6 and 10 to 12 copies, respectively, in *T. brucei* (66, 124). Although run-on assays point at polycistronic transcription of the *hsp70* locus (105), elements conserved in the intergenic regions of the *hsp70* and *hsp83* loci have been detected that possess sequence homology to the eukaryotic heat shock promoter element (66, 124). Moreover, the *hsp70* intergenic region has been reported to manifest a weak promoter activity in transient-activity assays, although no site of transcription initiation could be mapped (66). The steady-state levels of both *hsp70* and *hsp83* mRNAs are lower in procyclic than in bloodstream forms, but this difference may arise during in vitro cultivation of the procyclic forms (121, 188). The levels of *hsp70* and *hsp83* mRNAs are both elevated upon heat shock at 42°C, with the notable exception of the transcripts from the first gene of the *hsp70* locus, which is located 6 kb upstream from the tandem array of the other genes (104, 105). It was recently shown that this increase does not involve a change in the transcription rate in *Leishmania* species (10). Again, the role of the mRNA UTRs (5′ and 3′) appears to be important (9). It is not clear at which level this effect is produced.

Interestingly, it was found that the *trans* splicing of the *hsp70* transcripts is resistant to the heat-induced disruption which affects the processing of the other trypanosomal RNAs (125). Whether this reflects the presence of a specific category of splicing components has not been determined.

**Histone Genes**

The structure of chromatin in trypanosomatids is poorly understood, although it is clear that all classes of histones are present and that the DNA is packed into nucleosomes (78). The interactions between histones and the DNA appear to be much weaker than in other eukaryotes. In addition, structural
and functional differences exist between chromatin of the various stages of the parasitic life cycle (168). The genes for histones H1, H2A, and H2B of several trypanosomatid species have been cloned (11, 57, 61, 63, 151, 175). The mRNAs of these genes are polyadenylated. In T. cruzi and T. brucei, the level of the H2B mRNA is transiently reduced upon inhibition of DNA synthesis by aphidicolin or hydroxyurea (61, 62). Conflicting results were obtained for Leishmania enrietti (63), but the discrepancy may relate to the different time scales of the experiments. In addition, non-proliferative stages such as the stumpy form of T. brucei or trypanomastigotes of T. cruzi show lower H2B mRNA levels than do proliferative stages (61, 62). These observations suggest that the production of the histone mRNAs is coupled to DNA replication. The mechanism involved in this control is not understood, but it seems to depend on de novo protein synthesis (61).

**TRANSCRIPTION OF RNA GENES**

---

**rRNA Genes**

In trypanosomes, the rRNAs are synthesized in multiple transcription units clustered on several large chromosomes, although ribosomal promoters were occasionally found in minichromosomes (209). These units encode the canonical eukaryotic 18S, 5.8S, and 28S rRNAs, the 28S rRNA being split into two separate 28Sa and 28Sb rRNAs. In addition, they encode four small RNAs of 220, 180, 140, and 70 nucleotides (77, 201). Run-on assays suggested that these RNAs are processed in several steps from a single primary transcript (201). The promoter has been mapped to a short distance upstream of the first rRNA gene (1.2 kb in T. brucei, and 1.55 and 1.76 kb in T. cruzi) (48, 115, 201). As discussed above, in T. brucei this promoter possesses typical features of other eukaryotic rDNA promoters and shows structural similarities with the VSG and procyclin promoters. A stage-specific regulation of RNA elongation on these units has recently been observed in our laboratory (191). The combination of factors which triggers the differentiation of bloodstream into procyclic forms was found to reduce RNA elongation in both the ribosomal and VSG units. The significance of the rDNA transcription control is unclear, but it may relate to a higher stability of the ribosomal transcripts in the procyclic form.

**Spliced Leader Genes**

In T. brucei, the minicron donor RNA is encoded by clusters of 1.35-kb units (around 200 copies per genome) carried on large chromosomes. Run-on assays indicated that each unit is transcribed individually. In different Leishmania species, elements required for the synthesis of the med RNA were recently located very close to the transcription start site (at −67 to −58 and at −1 to +19) (4, 164). The −67 to −58 region was found to be critical for promoter activity, while the −1 to +19 region was important to confer stability to the RNA (164). These analyses revealed structural similarities with vertebrate promoters of one of the small U RNAs involved in splicing, the U2 snRNA. Moreover, the essential −67 to −58 region showed homology with the proximal sequence element of the promoters for snRNA genes of other eukaryotes (164). These findings are consistent with the hypothesis that the spliced leader is the analog of one of the U snRNAs, U1 (30). In vertebrates, some of the snRNA genes are transcribed by pol II and others are transcribed by pol III. Whether the med RNA is synthesized by pol II or pol III in the trypanosomes is still being debated (76, 164).

---

The med RNA is extensively polyadenylated in short stumpy forms but only barely so in long slender forms and not at all in procyclic forms (148). The significance of this observation is obscure but indicates that the process of polyadenylation can be stage regulated.

**Small RNA and tRNA Genes**

In trypanosomatids, small RNA genes such as those for the U snRNAs involved in splicing, are closely linked with different, divergently oriented tRNA genes (120, 122, 128). The use of the polymerase inhibitors α-amanitin and tagetitoxin indicated that these genes are transcribed by pol III (56). Analyses performed on several snRNA genes revealed three regions critical for promoter function, located in the immediate vicinity of the coding region (56, 128). Two upstream elements are centered respectively at positions −160 and −110, and the first 20 bases of the coding region are also important. The two upstream elements were identified as being the two intragenic promoter elements (boxes A and B) of the flanking tRNA genes (128). Thus, these boxes appear to play a dual role, as intragenic promoter elements for the respective tRNA genes and as extragenic regulatory elements for the associated small RNA genes.

**DIFFERENTIATION FROM THE INSECT TO THE BLOODSTREAM STAGE: THE METACYCLIC FORM**

In the salivary glands of the tsetse fly, the metacyclic trypanosomes manifest several preadaptations to the future blood environment. In particular, they resume the synthesis of VSG. At this stage, the expressed VSGs represent a limited and relatively conserved subset of the total VSG repertoire, termed the metacyclic VSG repertoire. The study of metacyclic VSG gene expression has been hampered by major problems. First, the metacyclic forms do not multiply and exist in only very small numbers. Second, for T. brucei, no reproducible system of in vitro differentiation into metacyclic forms is yet available. Third, although each cell synthesizes only a single type of VSG, the metacyclic population is heterogeneous, the whole metacyclic repertoire being expressed at the same time. Limited data were obtained from the study of moderately stable clones derived from metacyclic forms injected in the blood, where the expression of metacyclic VSGs continues for a few days.

**Metacyclic VSG Expression Sites**

The few data available, as well as theoretical considerations based on the characteristics of the phenomenon, strongly suggest that the activation of the metacyclic VSG genes occurs in situ, through stimulation by specific factors from the fly (14, 75, 107). This activation would require the continuous presence of these factors and disappears progressively when the metacyclic trypanosomes are diluted in the bloodstream.

All the metacyclic VSG genes cloned to date were found to be located in telomeres of large chromosomes. Their 5' environment differs from that of other telomeric VSG genes, in particular those present in bloodstream VSG expression sites (Fig. 4). Indeed, this environment is almost completely devoid of the 5' barren region, whose 76-bp repeats have been found to play an important role in the recombinations underlying antigenic variation in the bloodstream (74, 107). The absence of the 76-bp repeats probably accounts for the genetic isolation of the metacyclic VSG genes, by preventing the recombinations between the metacyclic VSG genes and the bloodstream VSG expression sites; hence the relative conservation of the
metacyclic repertoires (74). Another characteristic is the total or partial absence of ESAGs (73, 74).

The promoter of the metacyclic VSG genes was roughly mapped by nuclear run-on transcription assays and Northern (RNA) blot analysis (74). It was found to be approximately 2 kb upstream from the VSG gene and thus to be downstream from ESAG 1 when this gene is present. Recently, a breakthrough regarding the characterization of this promoter was obtained, following the observation of the expression of what appears to be a metacyclic VSG unit in bloodstream forms (6). The promoter of this monocistronic unit was present approximately 2 kb upstream from the VSG gene. Its 70-bp core element showed about 60% identity with that of promoters from bloodstream VSG expression sites (Fig. 3). Exactly the same promoter, located at a similar distance from the VSG gene, was found ahead of the gene for the major metacyclic VSG of T. gambiense, AnTat 11.17 (192). When inserted in plasmidic constructs, these promoters were able to drive the expression of a reporter gene in procyclic forms (6, 192).

These observations raise many questions. Although there is little doubt that the metacyclic VSG promoter has now been identified, it is still unclear how it does not function in procyclic forms when present in its genomic environment. Moreover, before the recent report of the apparent reexpression of a metacyclic unit in the bloodstream form (6), it was generally considered that metacyclic promoters were inactive during this stage. Is the case reported by Alarcon et al. (6) an exception? A reactivation of metacyclic expression sites late in the infection may allow the trypanosomes to extend their potential for antigenic variation when the generation of new bloodstream variants becomes problematic.

Finally, the expression of the ESAGs in metacyclic forms is very puzzling. The ESAGs are obviously absent from the metacyclic units, although some of them are present in close proximity. However, at least some ESAGs appear to be expressed at this stage, apparently by a polymerase resistant to α-amanitin (73). These observations suggest that some ESAG-containing transcription units driven by promoters similar to those of the VSG units are active in metacyclic forms. Clearly, this question awaits further experimental analysis.

DIFFERENTIATION FROM THE BLOODSTREAM TO THE PROCYCLIC FORM

As the parasitemia develops, long slender forms proliferating in the bloodstream are progressively replaced by nondividing short stumpy forms. These cells die unless they are ingested by the tsetse fly, where they rapidly differentiate into procyclic forms. This process can be mimicked in vitro by lowering the temperature from 37 to 27°C and by supplementing the cells with citrate/cis-aconitate (133, 206). When subjected to these conditions, the stumpy forms differentiate rapidly, synchronously, and completely, which is not the case for long slender forms (15, 155, 156, 206). These observations strongly suggest that the stumpy forms are blocked in a phase of the cell cycle when they are competent for differentiation (116, 117, 143).

Two factors presumably involved in these processes were recently uncovered. First, the expression and activity of a serine-threonine kinase termed Nrk were found to be highly increased as bloodstream forms differentiated from long slender to short stumpy forms (60). This increase was due to a translational control (60). Second, a transient increase of adenylate cyclase activity was detected when short stumpy forms were triggered to differentiate into procyclic forms (156). The role of these two enzymatic activities in differentiation remains totally obscure, although it was observed that cyclic AMP may increase the synchrony of procyclin appearance at the transition between bloodstream and procyclic forms (156).

Replacement of VSG by Procyclin

As soon as the stumpy forms are placed under appropriate conditions for differentiation into procyclic forms, the expression of the VSG is terminated. It is probable that the very first event is a cold-shock-dependent blocking of RNA elongation on the VSG expression site, in view of the rapidity of this phenomenon, even in long slender forms (7, 98, 141, 191). The VSG mRNA becomes preferentially degraded (51) and disappears almost totally within 2 h (143). The VSG is massively released from the cell surface between 4 and 6 h after initiation of differentiation, through the activation of a protease cleaving the antigen at its base (207).

Transcripts of the procyclic genes start to accumulate within the first 2 h and reach high levels around 4 to 6 h, at which point the amount of procyclin RNA is 20-fold higher than in established procyclic forms (143). This huge overproduction is probably needed to allow the parasite to renew its antigen coat as rapidly as possible. It is generally believed that procyclin protects the parasite against the defenses of the fly. As discussed above, the stimulation of procyclin expression is probably achieved through both transcriptional and posttranscriptional processes.

Strikingly, inhibitors of protein synthesis mimic the effects of differentiation: they superinduce the production of procyclin mRNA, trigger a 20-fold decrease in the rate of VSG mRNA synthesis, and cause a 6-fold decrease in its half-life (49, 51). A transient inhibition of protein synthesis has actually been detected in vivo at differentiation (15), although it remains to be seen if the timing of this inhibition is in accordance with a possible effect on the RNA changes.

Taken together, these observations point at the following model: upon cold shock, a temperature-dependent inhibition of protein synthesis causes the rapid loss of labile proteins involved in transcription elongation on the VSG expression site as well as in repression of procyclin mRNA synthesis. In addition to these controls, sequence elements present in the 3'-UTR of the transcripts differentially regulate the respective RNA amounts. All the factors involved in this highly speculative scenario have yet to be identified.

CONCLUDING CONSIDERATIONS

The genome organization in trypanosomatids is reminiscent of that in yeasts, with arrays of densely packed intronless genes, subject to a powerful homologous recombination machinery and interspersed with numerous retroposons (139). However, in trypanosomes, the characteristics of gene expression are totally different from those of yeasts and actually appear to be unique among eukaryotes, with the notable exception of a category of genes in worms (129, 176). So far, examples of canonical eukaryotic transcription units, with single genes flanked by individually controllable promoters, are the exception in trypanosomes. The genes seem to be generally contained within long polycistronic transcription units, under the influence of elusive and far-distant promoters. Only two promoters for protein-encoding genes are known, those for the major and stage-specific surface antigens of T. brucei, the VSG and procyclin. Numerous arguments tend to indicate that these promoters are of the ribosomal type. The originality of a ribosomal polymerase transcribing protein-encoding genes is possible in trypanosomes because of transcription in polycistronic units, which uncouples transcription initiation and RNA cap-
Polypyrimidine stretches located in the intergenic regions of these latter processes has been considerably improved in the primary polycistronic transcripts. Our understanding of differentiation of the parasite can be achieved at will under havethegenesandtheirpromotersbeenidentifiedbutalsothe form, represents an experimental model: in this case, not only of the bloodstream form and procyclin of the insect procyclic expression of the major surface antigens of during the embryology of metazoans. Therefore, concerted ological transformations comparable to cell differentiation ment: the cyclical transmissions between the insect vector and numerous cell adaptations to a drastically changing environ- ment. It is the rule that genes belonging to the parasitic lifecycle. Itistherulethatspeciesbelongingtothe genes are reversibly (in)activated by the regulation of gene expression. The continuous differential targetsofthehomologousrecombinationmachinery,meres. Indeed, the trypanosome telomeres appear to control the transcription of the telomeric VSG expression sites, which also contributes to antigenic variation. Lastly, the presence in telomeric DNA of a novel nucleotide, β-glucosylhydroxymethyluracil, may be responsible for the generation of point mutations in VSG genes, a newly identified source of antigenic variation.

As is clear from these considerations, many of the basic components and mechanisms involved in the modulation of trypanosome gene expression are still to be characterized. In particular, virtually nothing is known about factors able to bind specifically to DNA or RNA, and this also includes enzymes involved in the recombination of DNA. The future avenues of research are obvious.

ACKNOWLEDGMENTS

We thank D. Nolan for comments on the manuscript. The work in our laboratory was supported by the Belgian FRSM and FIRCIM, by research contract ARC 89/94-134 with the Communauté Française de Belgique, and by the Agreement for Collaborative Research between ILRAD (Nairobi) and Belgian Research Centres. L. V. is a senior research assistant of the Belgian National Fund for Scientific Research.

REFERENCES

1. Abielovich, H., T. Yehuda, and J. Shlomai. 1993. A trypanosomal CCHC-type zinc finger protein which binds the conserved universal sequence of kinetoplast DNA minicircles: isolation and analysis of the complete cDNA from Crithidia fasciculata. Mol. Cell. Biol. 13:7766–7773.
2. A Franchino, J. L., S. A. Gonzalez, and E. Pays. 1993. Isolation of a mitotic-like cyclin homologue from the protozoan T. brucei. Gene 132:75–82.
3. Agabian, N. 1990. Trans-splicing of nuclear pre-mRNAs. Cell 61:1157–1160.
4. Agami, R., R. A. Ly, S. Halman, and M. Shapira. 1994. Functional analysis of cis-acting DNA elements required for expression of the SL RNA gene in the parasitic protozoan Leishmania amazonensis. Nucleic Acids Res. 22: 1959–1965.
5. Agur, Z., D. Abiri, and L. H. T. Van der Ploeg. 1989. Ordered appearance of antigenic variants of African trypanosomes explained in a mathematical model based on a stochastic switch process and immune-selection against putative switch intermediates. Proc. Natl. Acad. Sci. USA 86:9626–9630.
6. Allarcon, C. M., H. Jin Sun, T. Hall, and J. E. Donelson. 1994. A monocis- tronic transcript for a trypanosome variant surface glycoprotein. Mol. Cell. Biol. 14:579–591.
7. Alexandre, S., P. Paindavoine, P. Toubabi, A. Pays, S. Hallens, M. Steinert, and E. Pays. 1990. Differential expression of a family of putative adenylate/ guanylate cyclase genes in T. brucei. Mol. Biochem. Parasitol. 43:279–288.
8. Aly, R., M. Argaman, S. Halman, and M. Shapira. 1994. A regulatory role for the 5’ and 3’ untranslated regions in differential expression of hsp83 in Leishmania. Nucleic Acids Res. 22:2922–2929.
9. Argaman, M., R. A. Aly, and M. Shapira. 1994. Expression of heat shock protein 83 in Leishmania is regulated posttranscriptionally. Mol. Biochem. Parasitol. 64:55–110.
10. Aslund, L., L. Carlsson, J. Henriksson, M. Rydaker, G. C. Toro, N. Galanti, and U. Pettersson. 1994. A gene family encoding heterogeneous histone H1 proteins in Trypanosoma cruzi. Mol. Biochem. Parasitol. 65:317–330.
11. Baltz, T., C. Giroud, F. Bringaud, H. Eisen, C. Jacquemot, and C. W. Roth. 1991. Exposed epitopes in a Trypanosoma equiperdum variant surface glycoprotein altered by point mutations. EMBO J. 10:1653–1659.
12. Barbet, A. F., and S. M. Kamper. 1993. The importance of mosaic genes to trypanosome survival. Parasitol. Today 9:63–66.
13. Barry, J. D., S. V. Graham, K. R. Matthews, P. G. Shiels, and O. A. Shonekan. 1990. Stage-specific mechanisms for activation and expression of variant surface glycoprotein genes in Trypanosoma brucei. Biochem. Soc. Trans. 18:708–710.
14. Bass, K. E., and C. C. Wang. 1992. Transient inhibition of protein synthesis accompanies differentiation of Trypanosoma brucei from bloodstream to procyclic forms. Mol. Biochem. Parasitol. 56:129–140.
15. Bayne, R. A., E. A. Kilbride, F. A. Lainson, L. Tetley, and J. D. Barry, 1993.
A major surface antigen of procyclic stage \textit{Trypanosoma congoense}. Mol. Biochem. Parasitol. 61:295–310.

17. Ben Amar, M. F., A. Pays, P. Tebabi, B. Dero, and T. Seebeck. 1988. Structure and transcription of the actin gene of \textit{Trypanosoma brucei}. Mol. Cell. Biochem. 121:216–217.

18. Benne, R. 1994. RNA editing in trypanosomes. Eur. J. Biochem. 219:5–23.

19. Berberof, M., A. Pays, and E. Pays. 1991. A similar gene is shared by both the VSG and procyclin gene transcription units of \textit{Trypanosoma brucei}. Mol. Cell. Biol. 11:1473–1479.

20. Berberof, M., and E. Pays. Unpublished data.

21. Berberof, M., L. Vanhamme, and E. Pays. \textit{Trypanosoma brucei}: a preferential splicing at the inverted polyadenylation site of the VSG mRNA provides further evidence for coupling between trans-splicing and polyadenylation. Exp. Parasitol., in press.

22. Berberof, M., L. Vanhamme, P. Tebabi, A. Pays, S. Welburn, D. Jefferies, and E. Pays. The 3'-terminal region of the mRNAs for VSG and procyclin can confer stage-specificity to gene expression in \textit{Trypanosoma brucei}. EMBO J., in press.

23. Bernardi, A., N. Van Harten-Loosbroek, and P. Borst. 1984. Modification of telomeric DNA in \textit{Trypanosoma brucei}, a role in antigenic variation? Nucleic Acids Res. 12:4153–4167.

24. Blum, M. L., J. A. Down, A. M. Garnett, M. Carrington, M. J. Turner, and D. C. Wiley. 1993. A structural motif in the variant surface glycoproteins of \textit{Trypanosoma brucei}. Nature (London) 362:603–609.

25. Borst, P. Personal communication.

26. Brinza, D., L. V. Harten-Loosbroek, P. T. J. Christiaens, A. R. J. Van der Veken, J. J. M. A. Van Thillo, B. S. Swinkels, and P. Borst. 1991. The \textit{3′}-terminal polyadenylated RNA of the \textit{Stage-specific 17.5 kDa protein (SSP-1)} gene family in \textit{Trypanosoma brucei} can confer stage-specificity to gene expression in \textit{Trypanosoma brucei}. Mol. Cell. Biol. 11:3489–3496.

27. Brown, S. D., and L. H. T. Van der Ploeg. 1994. Single-stranded DNA-binding protein in the procyclic acidic repetitive protein (PARP) promoter of \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 65:109–122.

28. Bruzik, J. P., and J. A. Steitz. 1990. Spliced leader RNA sequences substitute for the essential 5' end of U1 RNA during splicing in a mammalian in vitro system. Cell 62:809–899.

29. Carrington, M., N. Miller, M. Blum, I. Roditi, D. Wiley, and M. Turner. 1991. Variant specific glycoprotein of \textit{Trypanosoma brucei} consists of two domains each having an independently conserved pattern of cysteine residues. J. Mol. Biol. 221:823–835.

30. Chapman, A. C., and N. Agabian. 1994. \textit{Trypanosoma brucei} RNA polymerase II is phosphorylated in the absence of carboxy-terminal domain heptapeptide repeats. J. Biol. Chem. 269:4754–4760.

31. Chung, H. M., M. G. S. Lee, and L. H. T. Van der Ploeg. 1992. Characterization of the procyclic acidic repetitive protein (PARP) gene of \textit{Trypanosoma brucei}. Mol. Cell. Biochem. 121:216–222.

32. Clarry, A. W. C. A., R. Evers, E. Grondal, A. H. Hammer, W. Jess, and J. Kock. 1990. Transcription analysis in \textit{Trypanosoma brucei}. Biochem. Soc. Trans. 18:710–714.

33. Cornelissen, A. W. C. A., R. Evers, E. Grondal, A. Hammer, W. Jess, and J. Kock. 1991. Transcription and RNA polymerases in \textit{Trypanosoma brucei}. Nova Acta Leopold. 279:215–216.

34. Cross, G. A. M. 1990. Cellular and genetic aspects of antigenic variation in trypanosomes. Annu. Rev. Immunol. 8:83–110.

35. Czolij, F., S. H. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei} in Trypanosoma brucei. Mol. Biochem. Parasitol. 21:12–1.

36. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

37. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

38. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

39. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

40. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

41. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

42. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

43. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

44. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

45. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

46. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.

47. Cuzin, D., F. S. Ip, and G. A. M. Cross. 1985. Coordinate transcription of variable surface glycoprotein genes and an expression site associated gene in \textit{Trypanosoma brucei}. Mol. Biochem. Parasitol. 21:12–1.
and promoter-associated DNA rearrangement events. Mol. Cell. Biol. 11: 2467–2480.

71. Gottesdiener, K., L. Goriparthi, J. Masucci, and L. H. T. Van der Ploeg. 1992. A proposed mechanism for promoter-associated DNA rearrangement events at a variant surface glycoprotein gene expression site. Mol. Cell. Biol. 12:2478–4790.

72. Gottesdiener, K., M. J. Turner, and J. S. Cordingley. 1977. The transcription of the procyclin gene in Trypanosoma brucei: analysis of splicing and poly(A) addition at nascent RNA transcripts from the hsp70 locus. Mol. Cell. Biol. 11:3180–3190.

73. Hasen, Q. M. J., T. E. Rovers, K. Kosubek, and A. W. C. A. Cornelissen. 1989. Characterization of the RNA polymerases of Trypanosoma brucei: trypansomarial mRNAs are composed of transcripts derived from both RNA polymerase II and III. EMBO J. 8:3383–3389.

74. Hecker, H., B. Betschart, K. Bender, M. Burri, and W. Schlimme. 1984. Distinct functional requirements of RNA polymerase II and III. EMBO J. 3:3383–3389.

75. Hobbs, M., H. R. Hotz, C. Hartmann, and C. Clayton. 1993. Differential response to RNA trans-splicing in trypanosomes: implications for splice site selection. Mol. Cell. Biol. 13:2565–2573.

76. Lips, S., M. Gentens, F. Paturiaux-Hanocq, J. Hanocq-Quertier, and E. Pays. Unpublished data.

77. Lu, Y., T. Hall, L. S. Gay, and J. E. Donelson. 1993. Point mutations are associated with a gene duplication leading to the bloodstream reexpression of a trypanosome metacyclic VSG. Cell 72:397–406.

78. Martinez-Cavillo, S., and R. Hernandez. 1994. Trypanosoma cruzi ribosomal DNA: mapping of a putative distal promoter. Gene 142:221–232.

79. McKeown, R. H., and C. Clayton. 1989. Cycles within cycles: the interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes. J. Cell Biol. 125:1147–1156.

80. Mattheyssen, G., and M. White. 1983. Poly-pyrimidine tracts and their binding proteins: regulatory sites for posttranscriptional modulation of gene expression. Biochemistry 22:2321–2327.

81. Mattheyssen, G., C. K. Perry, P. M. Lizardi, R. Luhmann, and N. Agabian. 1989. Isolation and sequence of four small nuclear U RNA genes of Trypanosoma brucei subsp. brucei: identification of the U2, U4, and U6 RNA analogs. Mol. Cell. Biol. 9:1212–1223.

82. McKeown, R. H., and C. Clayton. 1989. Cdc2-related protein kinases and cell cycle control in trypanosomatids. Parasitol. Today 5:253–257.

83. McKeown, R. H., S. D. Bell, R. G. Nelson, and J. D. Barry. 1991. RNAs of Trypanosoma cruzi. Parasitol. Today 5:253–257.
174. Sollner-Webb, B., and J. Tower. 1986. Transcription of cloned eukaryotic ribosomal RNA genes. Annu. Rev. Biochem. 55:801–830.

175. Soto, M., J. M. Requena, L. C. Gomez, I. Navarrete, and C. Alonso. 1992. Molecular characterization of a Leishmania donovani infantum antigen identified as histone H2A. Eur. J. Biochem. 205:211–216.

176. Spieth, J., G. Brooke, S. Kuersten, K. Lea, and T. Blumenthal. 1993. Operons in C. elegans: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. Cell 73:521–532.

177. Steverding, D., Y. D. Stierhof, M. Chaudhri, M. Littenberg, D. Schell, A. G. Beck-Schinning, and P. Overath. 1994. ESAG 6 and 7 products of Trypanosoma brucei form a transferrin binding protein complex. Eur. J Cell Biol. 64:78–87.

178. ten Asbroek, A. L. M. A., C. A. A. Mol, R. Kieft, and P. Borst. 1993. Stable transformation of Trypanosoma brucei. Mol. Biochem. Parasitol. 59:133–142.

179. ten Asbroek, A. L. M. A., M. Ouellette, and P. Borst. 1990. Targeted insertion of the neomycin phosphotransferase gene into the tubulin gene cluster of Trypanosoma brucei. Nature (London) 348:174–175.

180. Thon, G., T. Baltz, C. Giroud, and H. Eisen. 1990. Trypanosome variable surface glycoproteins: composite genes and order of expression. Genes Dev. 4:1374–1383.

181. Tommerup, H., A. Dousmanis, and T. de Lange. 1994. Unusual chromatin in human telomeress. Mol. Cell. Biol. 14:5777–5785.

182. Torri, A. F., K. I. Bertrand, and S. L. Hajduk. 1993. Protein stability regulates the expression of cytochrome C during the developmental cycle of Trypanosoma brucei. Mol. Biochem. Parasitol. 57:305–316.

183. Torri, A. F., and S. L. Hajduk. 1988. Posttranscriptional regulation of cytochrome C expression during the developmental cycle of Trypanosoma brucei. Mol. Cell. Biol. 8:4625–4633.

184. Tschudi, C., and E. Ullu. 1989. Destruction of U2, U4, or U6 small nuclear RNA spliced leader RNA in Trypanosoma brucei. Mol. Biochem. Parasitol. 30:45–52.

185. Tzfati, Y., H. Abeliovich, I. Kapeller, and J. Shlomai. 1993. Organization of two invariant surfactant protein genes in Homo sapiens. J. Biol. Chem. 268:713–716.

186. VanderPloeg, L. H. T., K. Gottesdiener, and M. G. S. Lee. 1986. Three small RNAs within the genomes of P. falciparum that interact with the 5′-splice site of the Trypanosoma brucei spliced leader RNA in vivo. Cell 41:171–182.

187. Webb, J. R., and W. R. McMaster. 1993. Molecular cloning and expression of a Leishmania major gene encoding a single-stranded DNA-binding protein containing nine “CCHC” zinc finger motifs. J. Biol. Chem. 268:13994–14002.

188. Webb, J. R., and W. R. McMaster. 1994. Leishmania major HEXBP deletion mutants generated by double targeted gene replacement. Mol. Biochem. Parasitol. 63:231–242.

189. White, T. C., G. Rendenko, and P. Borst. 1986. Three small RNAs within the 10 kb Trypanosoma brucei RNA transcription unit are analogous to domain VII of other eukaryotic 28S rRNAs. Nucleic Acids Res. 14:9471–9480.

190. Vickerman, K. C., M. A. Curotto de Lafaille, and D. F. Wirth. 1994. Identification of a small RNA that interacts with the 5′ splice site of the Trypanosoma brucei spliced leader RNA in vivo. Mol. Biochem. Parasitol. 63:231–242.

191. Ziegelbauer, K., B. Stahl, M. Karas, Y. D. Stierhof, and P. Overath. 1993. Proteolytic release of cell surface proteins during differentiation of Trypanosoma brucei from bloodstream to procyclic forms in vitro. Eur. J. Biochem. 192:373–378.

192. Ziegelbauer, K., B. Stahl, M. Karas, Y. D. Stierhof, and P. Overath. 1993. Proteolytic release of cell surface proteins during differentiation of Trypanosoma brucei. Mol. Cell. Biol. 13:207–216.

193. Xu, G. L., B. Wieland, and A. Bindereif. 1994. trans-splicing of U6 RNA in trypanosomes: a mechanism conserved in some other eukaryotes. EMBO J. 13:521–532.

194. Vanhamme, L., M. Berberof, D. Le Ray, and E. Pays. 1990. A ribosomal gene promoter at the telomere of a mini-chromosome in Trypanosoma brucei. Nucleic Acids Res. 18:2725–2733.

195. Vanhamme, L., A. Pays, P. Tébaï, S. Alexandre, and E. Pays. 1991. Alpha-amanitin resistant transcription units in trypanosomes: a comparison of promoter sequences for variant-specific surface glycoprotein genes with novel features. Nucleic Acids Res. 19:1359–1368.

196. Ziegelbauer, J. C. B. M., R. Kieft, and P. Borst. 1993. Efficient production of functional mRNA mediated by RNA polymerase I in Trypanosoma brucei. Nature (London) 363:772–775.

197. Zomerdijk, J. C. B. M., R. Kieft, and P. Borst. 1993. Proteolytic release of cell surface proteins during differentiation of Trypanosoma brucei: a telomeric expression site for variant-specific surface glycoprotein genes with novel features. Nucleic Acids Res. 19:1359–1368.

198. Zomerdijk, J. C. B. M., R. Kieft, P. G. Shiels, and P. Borst. 1991. Antigenic variation in Trypanosoma brucei: a telomeric expression site for variant-specific surface glycoprotein genes with novel features. Nucleic Acids Res. 19:1359–1368.

199. Zomerdijk, J. C. B. M., R. Kieft, P. G. Shiels, and P. Borst. 1991. Antigenic variation in Trypanosoma brucei: a telomeric expression site for variant-specific surface glycoprotein genes with novel features. Nucleic Acids Res. 19:1359–1368.

200. Zomerdijk, J. C. B. M., R. Kieft, P. G. Shiels, and P. Borst. 1991. Antigenic variation in Trypanosoma brucei: a telomeric expression site for variant-specific surface glycoprotein genes with novel features. Nucleic Acids Res. 19:1359–1368.

201. Zomerdijk, J. C. B. M., R. Kieft, P. G. Shiels, and P. Borst. 1991. Antigenic variation in Trypanosoma brucei: a telomeric expression site for variant-specific surface glycoprotein genes with novel features. Nucleic Acids Res. 19:1359–1368.

202. Zomerdijk, J. C. B. M., R. Kieft, P. G. Shiels, and P. Borst. 1991. Antigenic variation in Trypanosoma brucei: a telomeric expression site for variant-specific surface glycoprotein genes with novel features. Nucleic Acids Res. 19:1359–1368.