Control of polyadenylation and alternative splicing of transcripts from adjacent genes in a procyclin expression site: a dual role for polypyrimidine tracts in trypanosomes?

Erik Vassella, Richard Braun and Isabel Roditi*

Institut für Allgemeine Mikrobiologie, Baltzerstrasse 4, Universität Bern, CH-3012 Bern, Switzerland

Received February 11, 1994; Revised and Accepted March 18, 1994

ABSTRACT

The procyclin-associated genes (PAGs) of *Trypanosoma brucei* are located downstream of tandemly repeated procyclin genes and belong to the same α -amanitin-resistant polycistronic transcription units. In procyclic form trypanosomes the PAG 1 pre-mRNA is alternatively spliced to give rise to three transcripts of 2.7 kb, 1.8 kb and 1.3 kb. The two larger transcripts contain additional short open reading frames (ORFs) upstream of the major ORF. Trans-splicing to generate these transcripts occurs downstream of three different polypyrimidine tracts. A minor population of procyclin mRNAs is also generated by alternative splicing at a polypyrimidine tract that begins 524 bp upstream of the major splice acceptor site of the procyclin β -gene. The same polypyrimidine tract is also required for accurate polyadenylation of mRNAs from the upstream procyclin α -gene (1). Alternatively polyadenylated forms of PAG 1 mRNAs can also be detected. All polyadenylation sites are found at a similar distance upstream of splice-acceptor sites, in each case with a polypyrimidine tract between them. Our results point to a dual role for polypyrimidine tracts in the maturation of trypanosome mRNAs.

INTRODUCTION

Alternative cis-splicing of precursor mRNA is a common regulatory mechanism of gene expression in higher eukaryotes, generating multiple transcripts from a single gene (for reviews, see refs. 2 and 3). This mechanism can lead to the production of different polypeptides from a single pre-mRNA or alternatively, it can function as an on-off switch for translation of functional polypeptides.

Trypanosomatids lack intervening sequences (introns) in protein coding genes and therefore do not process their pre-mRNA by cis-splicing. In contrast, most genes are transcribed as part of primary polycistronic transcription units, which are subsequently converted into monocistronic mRNAs by trans-splicing (for reviews, see refs. 4 and 5). Trans-splicing involves the cleavage at the 5' boundary of the protein coding exon and the addition

of a capped 39 nt mini-exon (or spliced leader) sequence to the newly generated 5' end. Consensus motifs at the splice acceptor sites for trans-splicing resemble those for cis-splicing in that they are situated at AG dinucleotides downstream of pyrimidine-rich regions (6).

Trypanosoma brucei alternately expresses two types of major surface glycoproteins, the procyclins in procyclic forms (7–9) and the variant surface glycoproteins (VSGs) in bloodstream forms of the parasite (10,11). These surface glycoproteins are expressed stage specifically and only during the differentiation of one form to the other are both types of proteins simultaneously expressed on the surface (12).

Procyclin and VSG genes are transcribed by an α -amanitin-resistant polymerase, which differs from the polymerase which transcribes normal house-keeping genes (13,14,15), and in each case form part of polycistronic transcription units containing other, coordinately regulated genes. Although there are between 10 and 20 potential VSG expression sites, only one is active at a time in bloodstream forms (11). The bloodstream form VSG transcription units are 45 to 60 kb long and contain at least 8 expression site associated genes (ESAGs) upstream of the VSG gene (16–21).

Procyclin genes are organized in tandem repeats of two or three copies on 4 different chromosomes (7,13,15). In contrast to the VSG expression site, all 4 procyclin expression sites appear to be transcribed in procyclic forms (13,22,23). In each locus, the transcription unit extends beyond the procyclin genes to a procyclin associated gene (PAG; ref. 22). The PAGs in the different procyclin expression sites have related 5' ends, but unique 3' ends. The polypeptide encoded by PAG 1 in the Pro A locus is related to ESAG 6 and 7 proteins (22). A gene related to ESAG 2, and accordingly named GRESAG 2, immediately follows the PAG in the Pro C locus (24).

We have recently shown that a 2.7 kb transcript from PAG 1 has a highly unusual structure (22). It consists of a long 5' leader of 1264 bases, containing multiple short open reading frames (ORFs), followed by the major ORF of 1215 nt. This transcript completely lacks a 3' untranslated region (UTR) as the stop codon UAA is also the start of the polyA tail. The GRESAG 2 mRNA is also unusual in that the ORF is interrupted by a stop

*To whom correspondence should be addressed

codon and resumes after a gap of three bases with an initiation codon (24).

In trypanosomal mRNAs, the major ORF normally begins with the first AUG codon after the splice acceptor site. Here we show that PAG 1 pre-mRNA is trans-spliced at three alternative positions, potentially giving rise to polycistronic, bicistronic and monocistronic transcripts. We also present evidence that procyclin pre-mRNAs can be alternatively spliced at a position in the intergenic region. Interestingly, the same polypyrimidine tract that is involved in alternative splicing regulates polyadenylation of transcripts of the upstream gene (1).

MATERIALS AND METHODS

Trypanosomes strains

Procyclic forms of *T.b. brucei* strain 427 (25) were cultivated at 27°C in SDM-79 medium containing 10% fetal bovine serum (26).

Northern blot analysis

Northern blots. Total RNA was isolated as previously described (12). For polysomal RNA preparation, 10⁹ cells from a procyclic culture were lysed (27). The postnuclear supernatant was layered onto a discontinuous sucrose gradient and the polysomes were precipitated as described (7). The RNA was subsequently purified from the polysomes using the same method as for total RNA (12). 8 µg RNA were fractionated on 1.2% agarose gels and transferred to nitrocellulose (13). Radioactive probes were generated by *in vitro* transcription (Stratagene, La Jolla) or random priming (Amersham Inc., Amersham, UK) using the appropriate kits. Hybridization and post-hybridization washing were carried out as described (28).

PCR mapping of 5' and 3' termini

20 µg RNA were used as a template for cDNA synthesis in a 40 µl reaction containing 10 U of AMV reverse transcriptase (Boehringer–Mannheim) and 0.5 µg of the hybrid T₁₇-adaptor primer (5'-GACTCGAGTCGACATCGT₁₇-3'; Ref. 12). The resulting cDNA–RNA hybrid was subsequently treated with 0.1 M NaOH at 65°C to degrade the RNA.

To map the sites of mini-exon addition to PAG 1 transcripts, PAG 1-specific cDNA was amplified using a mini-exon primer (5'-CGCTATTATTAGAACAG-3') and a primer #1 (5'-TGTTTGTAGGTGGCAG-3') corresponding to the complementary sequence of nucleotides 2002 to 2018 of PAG 1 (22). Prior to amplification, PAG 1-specific cDNA was enriched by hybrid-selection (28) with the plasmid cPAG 31 which corresponds to the 3' end of PAG 1 (22). PCR amplification was carried out for 35 cycles of 1 min at 94°C, 2 min at 45°C and 3 min at 74°C in a 100 µl reaction mix containing 1 unit of Supertaq (Stehelin AG, Basel), 200 µmol of each dNTP and 25 pmol of each primer.

The polyadenylation sites of PAG 1 transcripts were mapped after hybrid-selection of PAG 1 cDNA to the plasmid cPAG 21 (22) and amplification using the adaptor primer (5'-GACTCGAGTCGACATCG-3'; Ref. 29) and a primer #2 (5'-TACTTCTGGGGCATTCA-3') from nucleotides 508 to 524 of PAG 1 (22). PCR conditions were similar to those described above for 5' termini mapping, except that the annealing temperature was 50°C.

Alternative splice acceptor sites of procyclin transcripts were mapped by hybrid-selection of procyclin-specific cDNA to the

plasmid pAP2 (13,30) and amplification using the mini-exon primer and a primer #3 (5'-CGGTCGACAAGAATAGAT-3') from the intergenic region between the two tandemly linked procyclin genes in the Pro A locus. The annealing temperature was 37°C for two cycles and 45°C for a further 35 cycles.

BS+ -specific cDNA was selected by hybridization to BS+ and amplified with a mini-exon primer (5'-CGCTATTATTAGAACAGTTTCTGTAC-3') and a BS+ primer (5'-GTTTTCCCA-GTCACGACGTTG-3') for 35 cycles with an annealing temperature of 55°C.

The amplified cDNAs were repaired with T4 polymerase and ligated into the EcoRV site of Bluescript+ (Stratagene) or into sites indicated by underlined letters of the corresponding primer sequences. Sequences were determined by the dideoxynucleotide chain termination method (31).

Transient transfections

Transient transfections of procyclic form trypanosomes were carried out with two plasmids that contain chloramphenicol acetyltransferase (CAT) as a reporter gene. The wild type plasmid encompasses the procyclin promoter and the first of three tandemly repeated procyclin genes with a 297 bp 3' UTR and 265 bp of the intergenic region (28). The plasmid ΔPYR was derived from the wild type construct and lacks the polypyrimidine tract from positions 79–112 downstream of the polyadenylation site of the procyclin gene (see Ref. 1). Total RNA was isolated 7 hours after transfection (1,28) and cDNA synthesis was performed as described above.

Amplification of the intergenic sequence

Genomic DNA was extracted from the same stock of *T.brucei* strain 427 as was used for RNA extraction (32). The intergenic region between the posterior procyclin gene and PAG 1 was amplified from 200 ng of genomic DNA (30 cycles) with primer #1, which does not cross-hybridize with the PAGs in the other procyclin expression sites, (22) and primer #4 (5'-ACTCGA-ATTCCAAACGGTTTCTTCTG-3') from the 3' untranslated region of the procyclin gene. An EcoRI site (underlined) was incorporated to facilitate cloning.

RESULTS

Characterization of PAG 1 and procyclin transcripts

To characterize PAG 1-specific mRNAs, Northern blot analysis was carried out with a PAG 1-specific RNA probe. It has been shown previously that there is no cross-hybridisation of this probe to the PAGs from the other procyclin expression sites (22). The Northern blot of total cellular RNA from procyclic forms revealed at least 5 PAG 1-specific transcripts with sizes of 4.4, 2.7, 2.1, 1.8 and 1.3 kb (Figure 1a). The same transcripts are present in similar amounts in RNA isolated from the polyribosomal fraction of procyclic forms of the parasite, indicating that they are mature mRNAs.

As a control for the quality of the RNA, an equivalent blot was also hybridized with a procyclin probe. In this case, prolonged exposure of the blot revealed that in addition to the major transcript of 0.8 kb there is a second transcript of 1.3 kb (Figure 1b).

Mapping of the 5' end of PAG 1 transcripts

The multiple PAG 1 transcripts might be generated from PAG 1 precursor RNA by differential processing. Alternatively, they

Figure 1. Steady state levels of PAG 1 and procyclin transcripts in procyclic RNA. Northern blots with equal amounts of total cellular (1) or polysomal (2) RNA were hybridized with an antisense RNA probe corresponding to positions 1685–2018 of PAG 1 (ref. 22; panel A) or with a procyclin DNA probe synthesized by random priming of the plasmid pAP2 (refs. 13 and 30; panel B). Although not quantified precisely, the hybridization to PAG 1 transcripts was about 500 times lower than to procyclin transcripts.

Figure 2. Schematic depiction of alternatively processed procyclin and PAG 1 transcripts. The procyclin genes are shown as black boxes and PAG 1 as a stippled box. The transcripts are shown below the corresponding genes. Black circles symbolize the mini-exon sequence and short stretches of As the poly(A) tails. All ORFs longer than 30 bases are shown for the PAG 1 mRNAs (stippled boxes), while only the major ORFs are shown for the procyclin mRNAs (black boxes).

might be derived from different copies of PAG 1. To analyze if these transcripts were differentially spliced, PAG 1-specific cDNA was amplified with a mini-exon primer and a primer from the 3' end of PAG 1. PAG 1-specific cDNA was enriched by hybrid-selection prior to amplification in order to exclude recombination with cDNAs derived from the PAGs in the other procyclin expression sites. Three fragments of 2.0, 1.2 and 0.7 kb were generated by PCR. The sizes of the three cDNAs are consistent with them being derived from the transcripts of 2.7, 1.8 and 1.3 kb. The corresponding transcripts are designated as PAG 1 α (2.7 kb), PAG 1 β (1.8 kb) and PAG 1 γ (1.3 kb; see Figure 2). The precise sites of mini-exon addition were determined from the nucleotide sequences of the PCR products. Sequence analysis revealed that PAG 1 α contains multiple short ORFs upstream of the major ORF of 1215 bases. It differs from the PAG 1 transcript characterized by Koenig-Martin *et al.* (22) in that it is spliced to the mini-exon 11 bases further upstream. PAG 1 β is potentially bicistronic, with 77 bases of 5' UTR followed by two ORFs of 159 and 1215 bases, respectively. The smaller ORF starts with the first AUG after the mini-exon. A

Figure 3. There are two copies of PAG 1 in the genome of *T. brucei*. The finding that the restriction maps diverge upstream of the Sau3A and AvaII sites (broken line in the second copy) allowed the three cDNAs to be assigned to PAG 1 in the Pro A locus. #1 indicates the primer used for amplification of the cDNAs (see Materials and Methods). The location of the probe used for Northern blots in Fig. 1 is marked. Abbreviations: P, PvuII; B, BamHI; N, NaeI; H, HaeIII; Hi, HindIII; D, DraI; R, RsaI; K, KspI; Au, AluI; S, Sau3A; Al, AvaI; AII, EcoRV.

Figure 4. Characterization of the sequences flanking the three splice acceptor sites of PAG 1. The sequences upstream of the splice acceptor sites of PAG 1 β and PAG 1 γ , respectively, are from the published sequence of PAG 1 (22) and that of PAG 1 α is from a PCR clone generated by amplification of genomic DNA from *T. brucei* strain 427. Black circles denote the splice acceptor sites of the transcripts analyzed and the open circle the splice acceptor site of the previously published sequence (22). The polyadenylation sites of Pro β (upstream of the splice acceptor site of PAG 1 α) and PAG 1 α and PAG 1 β (upstream of the splice acceptor sites of PAG 1 β and PAG 1 γ , respectively) are denoted by black arrows. Polypyrimidine tracts are underlined. Asterisks indicate bases which differ from the sequence from *T. brucei* strain 227 (33) and bold letters the conserved 14-mer in the PAG 1 β and 1 γ splice acceptor sites.

computer search of the GenEmbl database revealed no obvious homology to known sequences. PAG 1 γ is a monocistronic transcript with a short 5' UTR of 13 nucleotides followed by a single ORF of 1152 bases. The splice acceptor site of this transcript is located between the first and the second AUG of the largest ORF of PAG 1 α .

To investigate whether PAG 1 α , 1 β and 1 γ are all derived from the same copy of PAG 1 in the Pro A locus or from other copies, Southern blot analysis of genomic DNA was performed using the same probe as for the Northern blots in Figure 1. This analysis revealed a second copy of PAG 1 (Figure 3). The two

Figure 5. a) Sequences flanking the splice acceptor site of Pro β' . The arrow denotes the polyadenylation site of Pro α and the filled circle the mini-exon addition site of Pro β' . The open circle marks the mini-exon addition site of transcripts synthesised during transient transfection with the construct Δ PYR (see below). The polypyrimidine tract which was deleted from the wild-type construct to generate Δ PYR is underlined and the short pyrimidine tract upstream of the mutant splice acceptor site is indicated by stippled boxes. b) Schematic representation of the wild-type construct and Δ PYR. The numbers above the symbols indicate the numbers of cDNA clones which mapped to a particular position. BS is the Bluescript+ vector backbone. The data for the polyadenylation sites are from Ref. 1.

copies are very similar to each other at their 3' ends, but differ upstream of the *Sau*3A and *Ava*II sites in PAG 1 (Figure 3). The finding that the 2 copies of PAG 1 diverge at this position allowed the 3 cDNA clones to be assigned unequivocally to PAG 1 from the Pro A locus.

Characterization of the three splice acceptor sites

The three splice acceptor sites of PAG 1 (see Figure 4) are situated at the first AG dinucleotide downstream of different polypyrimidine tracts. The splice acceptor site of the previously characterised PAG 1 transcript (22) is situated at an alternative AG dinucleotide 11 nucleotides downstream of that of PAG 1 α . Comparisons of the genomic sequence of the two strains revealed an AC dinucleotide in *T. brucei* strain 227 (22,33) instead of the AG dinucleotide that is used as the splice acceptor site in *T. brucei* strain 427.

The splice acceptor sites used to generate PAG 1 β and PAG 1 γ are both located in a conserved sequence motif, spanning from nucleotide positions -10 to +4 relative to the splice acceptor sites, with only two differences at positions -5 and +1. The last 9 bases of this sequence motif may allow a base pair interaction with a region of the mini-exon donor RNA (data not shown).

Alternatively spliced procyclin transcripts

The existence of a 1.3 kb procyclin-specific transcript raised the possibility that it is an alternatively spliced form of procyclin mRNA. Consistent with this, there is a second polypyrimidine tract beginning 524 bases upstream of the normal splice acceptor site of the procyclin β -gene which could theoretically give rise

Figure 6. Three pathways to produce alternatively spliced transcripts of PAG 1. The shorter transcripts may be generated by independent trans-splicing of the precursor RNA or by sequential trans-splicing of PAG 1 α , respectively, but not by cis-splicing of PAG 1 α . The GU dinucleotide at the mini-exon boundary of PAG 1 α , that has been created by the addition of the mini-exon (G) to the protein-coding exon (U), might theoretically serve as splice donor site for cis-splicing. However, cis-splicing of the intervening sequence to one of the downstream splice acceptor sites would lead to the loss of the G residue at the 3' end of the mini-exon.

to a transcript of this length. Complementary DNA which was synthesized from polysomal RNA was amplified with the mini-exon primer and the primer #3 from the intergenic sequence between the tandemly linked procyclin genes in the Pro A locus. The PCR reaction yielded two DNA fragments of 1.2 kb and 0.4 kb. Sequence analysis revealed that the cDNA of 0.4 kb indeed stems from an alternatively spliced form of the procyclin mRNA of the β -gene (Pro β') containing 471 bases of intergenic sequences at its 5' end (Figure 2). As predicted, the splice acceptor site of Pro β' is situated at the first AG dinucleotide downstream of the polypyrimidine tract in the intergenic region (Figure 5). The second cDNA stems from a bicistronic transcript, designated Pro $\alpha\beta$, which spans both procyclin genes (Figure 2). This RNA must be present in very low amounts as Northern blot analysis did not reveal a transcript of this size (data not shown).

Alternatively polyadenylated transcripts of PAG 1

The polypyrimidine tract immediately upstream of the splice acceptor site of Pro β' starts only 79 bases downstream of the polyadenylation site of the procyclin α -gene. We have recently shown that this polypyrimidine tract is also involved in accurate polyadenylation of mRNA from the α -gene (1). It is striking that the polypyrimidine tract upstream of the splice acceptor site of PAG 1 α begins a similar distance (86 bases) downstream of the polyadenylation site of the procyclin β -gene. If there is a general mechanism by which polypyrimidine tracts regulate both polyadenylation and trans-splicing, it should be possible to find PAG1 transcripts that are alternatively polyadenylated at positions upstream of the splice acceptor sites of PAG 1 β and 1 γ . The hybrid T₁₇-adaptor primer was once again used to prime cDNA synthesis from total polysomal RNA. Amplification of cDNA

was carried out with the adaptor primer and the primer #2 from the 5' end of PAG1. The poly(A) addition sites of different PCR clones were then determined by nucleotide sequencing. Most of the clones analyzed turned out to be derived from the PAG in the Pro C locus (data not shown). However, we were also able to identify four cDNA clones derived from PAG 1 that were polyadenylated at alternative positions. One clone was found to be polyadenylated 127 nt upstream of the splice acceptor site of PAG 1 β and a further three clones were found to be polyadenylated 169 nt upstream of the splice acceptor site of PAG 1 γ (Figure 4).

Trans-splicing and polyadenylation are controlled by the same polypyrimidine tract

To investigate whether the same polypyrimidine tract that is involved in correct polyadenylation of Pro α mRNAs also directs trans-splicing of the Pro β' mRNA, transient transfections were performed with two different plasmids (Figure 5). Both contain the procyclin promoter and the CAT coding region followed by the procyclin α gene 3' UTR. The wild type plasmid, which contains 265 bases of the downstream intergenic region always gives rise to correctly polyadenylated transcripts (28). In contrast, the majority of transcripts from Δ PYR, a deletion mutant which lacks the polypyrimidine tract in the intergenic region, are polyadenylated at different positions within the 3' UTR of the procyclin gene (1). The splice acceptor sites downstream of the polyadenylation sites were also analyzed by PCR. In the case of the wild type construct, all 4 cDNA clones that were characterized contained the mini-exon at the same position as the Pro β' mRNA. Although this splice acceptor site is also present in Δ PYR, an analysis of 3 cDNA clones revealed that mini-exon addition occurred at a AG dinucleotide 14 bases upstream (Fig. 5). In this case, the splice acceptor site was downstream of a very short pyrimidine tract. These results confirm that the same polypyrimidine tract is involved in both polyadenylation and trans-splicing.

DISCUSSION

In higher eukaryotes, alternative cis-splicing can result from the competition of alternative splice acceptor sites for a single splice donor site or, conversely, of alternative splice donor sites for a single splice acceptor site. In trypanosomes, however, the same splice donor site is used for every splicing event and alternative splicing can only result from splice acceptor site competition.

We have shown that PAG 1 pre-mRNA is alternatively spliced to give rise to three transcripts which differ in the lengths of their 5' leader sequences. These transcripts might be generated independently from the pre-mRNA by trans-splicing (Figure 6). Alternatively, they might be generated by sequential trans-splicing, with PAG 1 α giving rise to PAG 1 β and PAG 1 β to PAG 1 γ . Theoretically, it might also be possible to generate the smaller transcripts from PAG 1 α by cis-splicing. This would involve the deletion of the intervening sequence from the GU dinucleotide at the 3' boundary of the mini-exon sequence to one of the downstream splice acceptor sites. This possibility can be excluded, however, since it would have caused the deletion of a G residue at the 3' end of the mini-exon sequence (see Figure 6).

We have also provided evidence that the procyclin β -gene pre-mRNA gives rise to a major, monocistronic transcript and two minor, alternatively spliced transcripts. The Pro β' mRNA is

clearly detectable by Northern blot analysis. This transcript is spliced at a AG dinucleotide which is 471 bp upstream of the normal splice acceptor site. The existence of this transcript had previously been predicted from primer extension analysis (30) although the exact addition site of the mini-exon was not determined. In contrast, the Pro $\alpha\beta$ mRNA, which spans two procyclin genes, is not detectable on Northern blots and might be an aberrantly processed RNA.

There is also evidence for alternative splicing of other transcripts from the procyclin and VSG expression sites. In bloodstream form trypanosomes there are two alternatively spliced forms of ESAG 6 mRNA (34). In procyclic forms, however, only the larger transcript can be detected and then only after stabilisation of the RNA by UV irradiation (35). A cDNA of GRESAG 2 was shown to contain two adjacent ORFs of 597 and 753 bases (24). Northern blots have recently revealed two GRESAG 2-specific transcripts of different lengths (E.Pays, personal communication). We have looked for additional ORFs in published sequences of the VSG and procyclin expression sites and found that a cDNA of ESAG 5 contains two ORFs of 231 and 1251 nt which overlap for 50 bases (20). It remains to be established if there is also an alternatively spliced, monocistronic transcript of ESAG 5.

In addition to alternatively spliced transcripts we have also identified alternatively polyadenylated PAG 1 transcripts. In each case, the polyadenylation site was found upstream of a splice acceptor site with a polypyrimidine tract between them. The same organization also holds true for the procyclin genes, suggesting that polypyrimidine tracts are required for both 5' and 3' processing. Indeed, we have demonstrated that the same polypyrimidine tract that is responsible for accurate polyadenylation of Pro α mRNA (1) also directs trans-splicing of Pro β' pre-mRNA. It has also recently been shown for the DHFR-TS locus of *Leishmania major* that one element regulates both polyadenylation and trans-splicing (36). A connection between trans-splicing and polyadenylation is not unique to kinetoplastids. Both monocistronic and polycistronic transcription units occur in the nematode *C.elegans* (37) and once again, the mRNAs derived from polycistronic transcription units possess 5' spliced leaders. Trans-splicing and polyadenylation also appeared to be coupled in this system since mutation of the polyadenylation signal in the 3' UTR of the *gpd-2* gene reduces the efficiency of trans-splicing of the transcript from the downstream *gpd-3* gene.

Our results suggest that both polyadenylation and trans-splicing can occur at every polypyrimidine tract in the intergenic regions. Transcripts such as PAG 1 α and Pro β' could therefore be regarded as by-products of procyclin mRNA polyadenylation. Similarly, the larger transcript of ESAG 6 might be a by-product from polyadenylation of ESAG 7, which is located immediately upstream.

It may also be the case, however, that alternative splicing constitutes a further level of gene regulation in trypanosomes. This mechanism generates transcripts in which different ORFs start with the first AUG triplet after the mini-exon and may therefore lead to the production of different polypeptides from a single pre-mRNA. The finding that all alternatively spliced PAG 1 transcripts are present in the polysomal fraction lends weight to this hypothesis. Upstream ORFs can also have a profound effect on the translation of downstream ORFs (38,39). Consequently, another function of alternative splicing might be to control expression of the major ORF.

ACKNOWLEDGEMENTS

We thank Adrian Hehl and Nadia Schürch for stimulating discussions. This work was supported by the Swiss National Science Foundation, the Helmut Horten Foundation and the Roche Foundation.

REFERENCES

- Schürch, N., Hehl, A., Vassella, E., Braun R. & Roditi, I. (1994). *Mol. Cell. Biol.*, in press.
- Maniatis, T. (1991) *Science* **251**, 33–34.
- McKeown, M. (1992) *Annu. Rev. Cell Biol.* **8**, 133–55.
- Agabian, N. (1990) *Cell* **61**, 1157–60.
- Borst, P. (1986) *Annu. Rev. Biochem.* **55**, 701–732.
- Huang, J. & Van der Ploeg, L. H. T. (1991) *EMBO J.* **10**, 3877–3885.
- Mowatt, M. R. & Clayton, C. E. (1987) *Mol. Cell. Biol.* **7**, 2833–2844.
- Richardson, J. P., Beecroft, R. P., Tolson, D. L., Liu, M. K. & Pearson, T. W. (1988) *Mol. Biochem. Parasitol.* **31**, 203–216.
- Roditi, I., Carrington, M. & Turner M. (1987) *Nature* **325**, 272–274.
- Cross, G. A. M. (1990) *Annu. Rev. Immunol.* **8**, 83–110.
- Pays, E. and Steinert, M. (1988) *Annu. Rev. Genet.* **22**, 107–126.
- Roditi, I., Schwarz, H., Pearson, T. W., Beecroft, R. P., Liu, M. K., Richardson, J. P., Buehring, H.-J., Pleiss, J., Buelow, R., Williams, R. O. & Overath, P. (1989) *J. Cell Biol.* **108**, 737–746.
- Koenig, E., Delius, H., Carrington, M., Williams, R.O. and Roditi, I. (1989) *Nucleic Acids Res.* **17**, 8727–8739.
- Kooter, J. M. & Borst, P. (1984) *Nucl. Acids Res.* **12**, 9457–9472.
- Rudenko, G., Le Blanc, S., Smith, J., Lee, M. G. S., Rattray, A. & Van der Ploeg, L. H. T. (1990) *Mol. Cell. Biol.* **10**, 3492–3504.
- Alexandre, S., Guyaux, M., Murphy, N.B., Coquelet, H., Pays, A., Steinert, M. and Pays, E. (1988) *Mol. Cell. Biol.* **8**, 2367–2378.
- Cully, D.F., Ip, H.S. and Cross, G.A.M. (1985) *Cell* **42**, 173–182.
- Florent, I., Raibaud, A. and Eisen, H. (1991) *Mol. Cell. Biol.* **11**, 2180–2188.
- Hobbs, M.R. and Boothroyd, J.C. (1990) *Mol. Biochem. Parasitol.* **43**, 1–16.
- Pays, E., Tebabi, P., Pays, A., Coquelet, H., Revelard, P., Salmon, D. & Steinert, M. (1989) *Cell* **57**, 835–845.
- Revelard, P., Lips, S. and Pays, E. (1990) *Nucl. Acids Res.* **18**, 7299–7304.
- Koenig-Martin, E., Yamage, M. & Roditi, I. (1992) *Mol. Cell. Biol.* **55**, 135–146.
- Mowatt, M.R., Wisdom, G.S. and Clayton, C.E. (1988) *Mol. Cell. Biol.* **9**, 1332–1335.
- Berberof, M., Pays, A. and Pays, E. (1991) *Mol. Cell. Biol.* **11**, 1437–1479.
- Cross, G. A. M. & Manning, J. C. (1973) *Parasitology* **67**, 315–331.
- Brun, R. & Schoenenberger, M. (1979) *Acta Tropica* **36**, 286–292.
- Aline, R. F. J., Scholler, J. K. & Stuart, K. (1988) *Mol. Biochem. Parasitol.* **32**, 169–178.
- Hehl, A., Vassella, E., Braun, R & Roditi, I. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 370–374
- Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
- Pays, E., Coquelet, H., Tebabi, P., Pays, A., Jeffries, D., Steinert, M., Koenig, E., Williams, R.O. & Roditi, I. (1990) *EMBO J.* **9**, 3145–3151.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc Natl. Acad. Sci. USA* **74**, 5463–5467.
- Young, J.R., Donelson, J.E., Majiwa, P.A., Shapiro, S.Z. & Williams, R.O. (1982) *Nucl. Acids Res.* **10**, 803–819.
- Miller, N. & Turner, M. J. (1981) *Parasitology* **82**, 63–80.
- Coquelet, H., Tebabi, P., Pays, A., Steinert, M. & Pays, E. (1989) *Mol. Cell. Biol.* **9**, 4022–4025.
- Pays, E., Coquelet, H., Pays, A., Tebabi, P. & Steinert, M. (1989) *Mol. Cell. Biol.* **9**, 4018–4021.
- LeBowitz, H.J., Smith, H.Q., Rusche, L. and Beverley, S.M. (1993) *Genes & Development* **7**, 996–1007.
- Spieth, J., Brooke, G., Kuersten, S., Lea, K. and Blumenthal, T. (1993) *Cell* **73**, 521–532.
- Fütterer, J. and Hohn, T. (1992). *Nucleic Acids Res.* **20**, 3851–3857.
- Mueller, P.P. and Hinnebusch, A.G. (1986) *Cell* **45**, 201–207.