MOLECULAR AND CELLULAR BIOLOGY, July 1990, p. 3492–3504
0270-7306/90/073492-13$02.00/0

Vol. 10, No. 7

Procyclic Acidic Repetitive Protein (PARP) Genes Located in an Unusually Small α-Amanitin-Resistant Transcription Unit: PARP Promoter Activity Assayed by Transient DNA Transfection of *Trypanosoma brucei*

GLORIA RUDENKO,[1] SYLVIA LE BLANCQ,[1] JOSEPH SMITH,[1] MARY GWO SHU LEE,[2] ALISON RATTRAY,[1] AND LEX H. T. VAN DER PLOEG[1]*

Department of Genetics and Development, College of Physicians and Surgeons,[1] and Division of Tropical Medicine, School of Public Health,[2] Columbia University, 701 West 168th Street, New York, New York 10032

Received 21 February 1990/Accepted 17 April 1990

At least one of the procyclic acidic repetitive protein (PARP or procyclin) loci of *Trypanosoma brucei* is a small (5- to 6-kilobase) polycistronic transcription unit which is transcribed in an α-amanitin-resistant manner. Its single promoter, as mapped by run-on transcription analysis and UV inactivation of transcription, is located immediately upstream of the first α-PARP gene. Transcription termination occurs in a region approximately 3 kilobases downstream of the β-PARP gene. The location of the promoter was confirmed by its ability to direct transcription of the bacterial chloramphenicol acetyltransferase gene in insect-form (procyclic) *T. brucei*. The putative PARP promoter is located in the region between the 3′ splice acceptor site (nucleotide position 0) and nucleotide position −196 upstream of the α-PARP genes. Regulatory regions influencing the levels of PARP expression may be located further upstream. We conclude that a single promoter, which is located very close to the 3′ splice acceptor site of the α-PARP genes, directs the transcription of a small, polycistronic, and α-amanitin-resistant transcription unit.

The protozoan parasite *Trypanosoma brucei*, causative agent of sleeping sickness in humans, parasitizes the bloodstream of its mammalian host and is transmitted by the tsetse fly. On ingestion of an infected blood meal by the tsetse fly, the bloodstream-form trypanosomes differentiate into insect-adapted (procyclic-form) trypanosomes. The bloodstream form is sheathed by a dense variant cell surface glycoprotein (VSG) coat which is lost upon differentiation into the procyclic form. The procyclic trypanosome is characterized by an abundant cell surface protein, the procyclic acidic repetitive protein (PARP) or procyclin (7, 27–29, 33–35).

The VSG coat is encoded by a single VSG gene, which is transcribed in one of several telomerically located VSG gene expression sites. VSG gene activation can occur through either one of two separate mechanisms. In the first, different types of DNA recombinational events can translocate a VSG gene into an active expression site. In the second, VSG gene switching results from the differential transcriptional control of the expression sites. The expression site transcription units are large, extending up to 60 kilobases (kb) (for reviews, see references 3, 6, 30, 31, and L. H. T. Van der Ploeg, *in* B. D. Hames and D. Glover, ed., *Frontiers in Molecular Biology: Genome Rearrangements and Amplification*, in press).

Transcription of many protein-coding genes in trypanosomes and related species is believed to be polycistronic, generating large precursor RNAs (pre-mRNAs) containing multiple protein-coding genes. Maturation of these pre-mRNAs to generate the mature mRNAs requires addition of a 39-nucleotide (nt) capped miniexon or spliced leader by *trans* splicing and polyadenylation (for reviews, see references 42 and 47). The function of mRNA maturation via *trans* splicing of polycistronic pre-mRNAs and the role of the miniexon are unclear.

A peculiarity of polycistronic VSG gene expression site transcription is its insensitivity to the drug α-amanitin at concentrations up to 1 mg/ml (19). This resistance is characteristic of transcription by RNA polymerase I, which in all other eucaryotes transcribes rRNA genes exclusively. Most other protein-coding genes in trypanosomes (and other eucaryotes) are transcribed by an α-amanitin-sensitive RNA polymerase, presumably RNA polymerase II (8, 11, 45).

Recently, we have shown that transcription of protein-coding genes by an α-amanitin-resistant RNA polymerase is not restricted to the VSG gene expression sites. The PARP protein-coding genes are also transcribed by an α-amanitin-resistant RNA polymerase, suggesting that a subset of protein-coding genes is transcribed by the same type of RNA polymerase (36). To improve our understanding of polycistronic transcription and mRNA maturation by *trans* splicing, we analyzed the transcription of one of the PARP gene arrays. We describe an α-amanitin-resistant transcription unit flanked by defined transcription initiation and termination sites. Polycistronic transcription units can thus be small and encode few, coordinately regulated protein-coding genes.

MATERIALS AND METHODS

Trypanosomes. All experiments with bloodstream-form trypanosomes used *T. brucei* variant 118 clone 1 (23). Procyclic trypanosomes were either established from *T. brucei* 118 clone 1 bloodstream forms adapted to procyclic culture (A. Rattray and L. H. T. Van der Ploeg, unpublished

* Corresponding author.

results) or a *T. brucei* 427 procyclic trypanosome stock originally obtained from R. Brun. We call this strain *T. brucei* 427 RB. The nascent RNA analysis, UV inactivation, and DNA transfection experiments presented here were performed with the latter strain. Nascent RNA analysis and transfection experiments gave the same results with either strain. Bloodstream-form trypanosomes were passaged in rats and harvested by cardiac puncture. Procyclic trypanosomes were maintained in SDM79 medium at 23 to 25°C as described by Brun and Schonenberger (5).

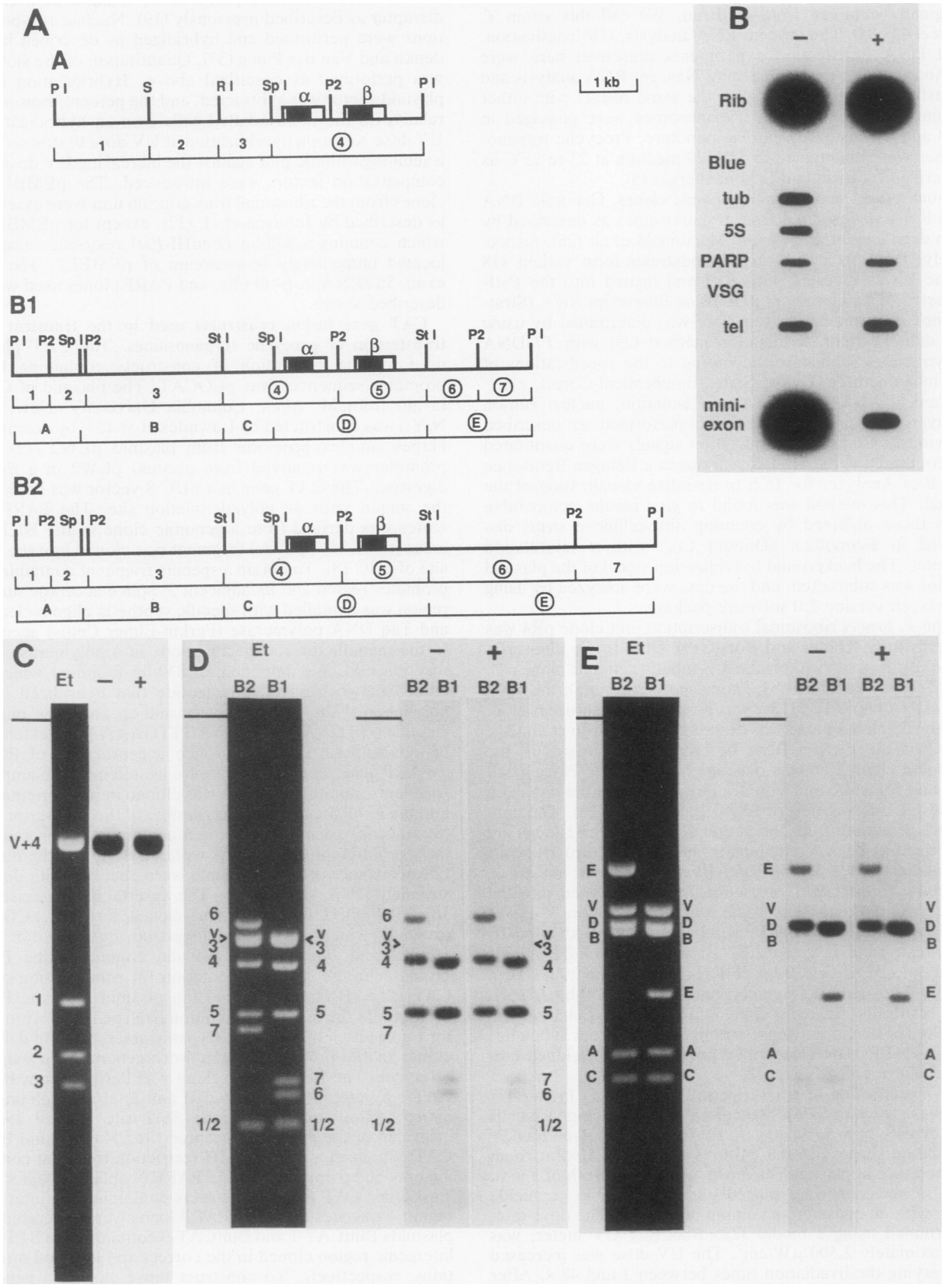
Nucleic acid analysis and genomic clones. Genomic DNA and RNA was isolated from trypanosomes as described by Van der Ploeg et al. (48) and Macdonald et al. (26), respectively. DNA from *T. brucei* bloodstream-form variant 118 clone 1 was digested with *Pst*I and ligated into the *Pst*I-digested plasmid vectors pUC18 or Bluescript SK+ (Stratagene). The nucleotide sequence was determined by using the dideoxy-chain termination method (39) with T7 DNA polymerase (Sequenase) according to the specifications of the manufacturer (United States Biochemical Corp.).

Nascent RNA analysis. Nuclei isolation, nuclear run-on analysis and hybridization were performed as described previously (20, 37). Hybridization signals were quantitated by counting the nitrocellulose strips in a Betagen Betascope 603 Blot Analyzer for 16 h to optimize visualization of the signal. This method was found to give results comparable with those obtained by counting nitrocellulose strips dissolved in Econofluor (DuPont Co.) with a scintillation counter. The background hybridization signal of the plasmid vector was subtracted, and the data were analyzed by using a Betagen version 2.0 software package.

The *T. brucei* ribosomal transcription unit clone pR4 was described by Kooter and Borst (19); Blue is the Bluescript SK+ (Stratagene) vector; the α-β-tubulin genomic clone pTb α-β T-1 was described by Thomashow et al. (46); the VSG 118 cDNA clone TcV118-2 was described by Bernards et al. (2); the 5S rRNA clone was described by Lenardo et al. (24); PARP is the PARP cDNA described previously (36); the telomere clone TT6 was described by Van der Ploeg et al. (49) and Rudenko and Van der Ploeg (37); and the miniexon clone CL103 was described by Laird et al. (22). The subclones generated from the 3′ end of the PARP B1 locus are described as B1 A, containing a 490-base-pair (bp) *Stu*I-*Sal*I fragment; B1 B, a 560-bp *Nde*I-*Hin*dIII fragment; and B1 C, a 900-bp *Bam*HI-*Pst*I restriction fragment cloned into the Bluescript plasmid vector. The subclones from the 3′ end of the PARP B2 locus were B2 A, containing a 2.9-kb *Hin*dIII-*Pst*I restriction fragment; B2 B, a 960-bp *Hin*dIII-*Hin*cII fragment; B2 C, a 530-bp *Hin*cII-*Xho*I fragment; B2 D, a 650-bp *Xho*I-*Hin*cII fragment; and B2 E, a 750-bp *Hin*cII-*Pst*I restriction fragment cloned into Bluescript. 4-1 and 4-2 are the sense and antisense strands derived from subclones of the PARP cDNA described previously (36), cloned into the phage vector M13 mp18.

UV inactivation of transcription. Procyclic *T. brucei* 427 RB was grown in SDM-79 medium at 23°C. Log-phase cells at a density of 5×10^6 to 1×10^7 per ml were used for UV irradiation. Procyclic-cell cultures were irradiated at room temperature in plastic dishes with a solution depth of 3 mm. Cell suspensions were placed under a portable germicide lamp with an emission maximum at 254 nm. The dose rate, determined using a model J225 Blak-Ray UV meter, was approximately 3,500 μW/cm^2. The UV dose was increased by varying the irradiation times between 3 and 48 s. After irradiation, cells were transferred to 50-ml culture tubes and incubated at 23°C for 1.5 h in the dark. Cells were then spun down, and nuclei were prepared by using a Stansted cell disrupter as described previously (19). Nuclear run-on reactions were performed and hybridized as described by Rudenko and Van der Ploeg (37). Quantitation of the slot blots was performed as described above. Hybridization to the plasmid vector was subtracted, and the percent transcription relative to the nonirradiated cells (run-on hybridization at UV dose X/run-on hybridization at UV dose 0) was set out in a semi-logarithmic plot against the increasing UV doses. No compensation factors were introduced. The pEMBL subclones from the ribosomal transcription unit were essentially as described by Johnson et al. (17), except for pEMBL 2.5, which contains a 850-bp *Hin*dIII-*Pst*I restriction fragment located immediately downstream of pEMBL2. The miniexon, 5S rRNA, α-β-tubulin, and PARP clones used were as described above.

CAT gene fusion constructs used in the transient DNA transfection of procyclic trypanosomes. The CAT plasmid used for the generation of constructs containing PARP promoter sequences was pCOCAT. The plasmid pCOCAT (a gift from M. Allan, Columbia University, New York, N.Y.) was constructed by Lowndes et al. (25) by deleting the herpes simplex promoter from plasmid pLW2 (12). This promoter was removed from plasmid pLW2 in a *Bam*HI digestion. The CAT gene in a pUC 8 vector was flanked by the simian virus 40 polyadenylation site. The PARP promoter was derived from a genomic clone of the B1 PARP locus (the 1,345-bp *Nde*I fragment was cloned into the *Sma*I site of pUC 18). To obtain a specific fragment containing the promoter region and its adjacent 3′ splice acceptor site, the region was amplified with specific synthetic oligonucleotides and Taq DNA polymerase (Perkin-Elmer Cetus) according to the manufacturer's specifications in a polymerase chain reaction (38). We amplified an 830-bp fragment which extended between an oligonucleotide that hybridized in the polylinker of the plasmid vector and an antisense oligonucleotide (-TTGAAGTCTTCAAGTTGAAA-) that extended 7 bp downstream of the 3′ splice acceptor site of the B1 α-PARP gene. *Eco*RI linkers were ligated onto the amplified fragments, and the fragment was cloned in both orientations into the *Eco*RI restriction enzyme site of the Bluescript SK+ (Stratagene) plasmid. Two clones containing the promoter in the two different orientations were chosen, and the nucleotide sequences of the fragments were checked with double-stranded DNA sequencing. The inserts were excised as *Hin*dIII-*Bam*HI fragments and cloned into the pCOCAT plasmid, which had been linearized by digestion with *Hin*dIII and *Bam*HI. The plasmid containing the PARP promoter in the correct orientation for transcription of the CAT gene (BNspCAT-1) and the plasmid with the PARP promoter in the reverse orientation (BNspCAT-0) were used for transfection of procyclic trypanosomes. Several deletion clones of the promoter region were generated to map the promoter. The BΔNspCAT clone was derived from BNspCAT-1 by digesting with *Hin*dIII and *Spe*I and deleting the entire region upstream of the *Spe*I site located 196 bp upstream of the 3′ splice acceptor site. The plasmid BSspCAT contained a *Sca*I-*Bam*HI restriction fragment containing only 80 bp upstream of the PARP 3′ splice acceptor site. Two other CAT fusion constructs containing the intergenic region sequences of the B1 PARP locus were generated: the plasmids BintCAT-1 and BintCAT-0 contained the B1 PARP intergenic region cloned in the correct and inverted orientations, respectively. To construct these clones, a genomic clone containing the intergenic region of the B1 PARP locus (2.2-kb *Nde*I restriction fragment, extending 3′ of the *Nde*I


A
A
P I
S
R I
Sp I
α
P2
β
P I
1 kb
1
2
3
④
B1
P I P2 Sp I P2
St I
Sp I
α
P2
β
St I
P2
P I
1 2 3 ④ ⑤ ⑥ ⑦
A B C Ⓓ Ⓔ
B2
P I P2 Sp I P2
St I
Sp I
α
P2
β
St I
P2
P I
1 2 3 ④ ⑤ ⑥ 7
A B C Ⓓ Ⓔ
B
–
+
Rib
Blue
tub
5S
PARP
VSG
tel
mini-
exon
C
Et – +
V+4
1
2
3
D
Et
B2 B1
– +
B2 B1 B2 B1
6
V
3
4
5
7
1/2
E
Et
B2 B1
– +
B2 B1 B2 B1
E
V
D
B
A
C

restriction enzyme site in the PARP-coding sequence [see Fig. 4]) was annealed with the synthetic oligonucleotides mentioned above to amplify the DNA sequences from 7 bp downstream of the B1 β-PARP 3′ splice acceptor site to an *Nde*I restriction enzyme site located in the 3′ untranslated extension of the α-PARP gene. The 850-bp amplified fragments were cloned upstream of the CAT gene in the *Hin*cII site of pCOCAT (the appropriate restriction enzyme sites are marked in the sequence in Fig. 4).

***T. brucei* transfection.** The protocol used was based on the method described in Bellofatto and Cross (1) for *Leptomonas* species, with modifications. All procedures were performed under sterile tissue culture conditions, and bacterial contamination of the cultures was never detected (controls were essentially as described in reference 1). Procyclic *T. brucei* 427 RB cells were grown to a density of 6×10^6 to 8×10^6 cells per ml, spun down, and washed with phosphate-buffered saline. Cells were suspended at a concentration of 10^8 cells per ml in Zimmerman postfusion medium and incubated at 21°C for approximately 1 h before electroporation. A 0.5-ml portion of the cell suspension (5×10^7 cells) was transferred to 0.4-cm-electrode-gap Bio-Rad Gene Pulser cuvettes and electroporated by using a Bio-Rad Gene Pulser with a capacitance extender and with two charges of 450 V and 125 μF applied 10 s apart. This dose gave approximately 50% cell death (measured immediately after the electroporation). Eighty micrograms of plasmid DNA (160 μg/ml) purified by using pZ523 columns (5′-3′ Inc.) was used for each sample. After all the samples in a series were electroporated, they were transferred to 10 ml of SDM79 medium in 25-cm^2 Falcon flasks. The cells were incubated for 18 to 24 h in SDM79 medium. Cell survival was determined for each sample, and the cells were spun down, washed twice with 5 ml of phosphate-buffered saline and suspended in 150 μl of 250 mM Tris (pH 7.8). The CAT assay was then performed. The cells were freeze-thawed three times, debris from the extracts was removed by 5 min of centrifugation in a microcentrifuge, and the supernatant was transferred to a new tube and incubated for 5 min at 68°C. To each sample, 60 μl of 0.25 M Tris (pH 7.8), 1.2 μl of 50 mM *n*-butyryl coenzyme A (Pharmacia), and 2.4 μl of D-threo-[*dichloroacetyl*-1,2-^{14}C]chloramphenicol (Dupont, NEN Research Products) was added, after which the mixture was incubated for 18 h at 37°C. The incubation mixture was extracted with 400 μl of xylene. The organic phase was back extracted twice with 200 μl of 10 mM Tris and 0.1 mM EDTA. One-tenth of the organic phase was added to 5 ml of scintillation fluid (Betamax; ICN Pharmaceuticals Inc.) and counted. The remaining xylene was dried down and chromatographed on thin-layer silica plates in chloroform-methanol (95:5) to check that the radioactivity was indeed present as the [^{14}C]butyryl-chloramphenicol form.

After the manuscript was submitted for publication, procedural modifications to optimize the transfection protocol, affecting the voltage, capacitance, temperature and length of time of the incubations (J. Smith, G. Rudenko, and L. H. T. Van der Ploeg, unpublished), were introduced. Cells were spun down and washed with 4°C phosphate-buffered saline and kept at 4°C throughout the procedure. Cells were suspended at a density of 10^8 cells per ml in Zimmerman postfusion medium supplemented with 0.5% glucose and incubated on ice for 20 min with cesium chloride gradient-purified plasmid DNA before the electroporation. Electroporation was performed as described above, with a single charge at 1,600 V and 25 μF. After electroporation, cells were placed in an ice bath for 10 min and then transferred to 10 ml of SDM 79 medium. The CAT assay was modified and performed for 1.5 h with 80 μl of cell lysate in a volume of 100 μl containing 250 μM *n*-butyryl coenzyme A (Pharmacia), 50 μM chloramphenicol (Boehringer Mannheim Biochemicals), 1.0 μCi of [3,5-^{3}H]chloramphenicol (Dupont, NEN), and 100 mM Tris (pH 8.0). The transfection assays performed by using this modified procedure gave the same relative ratios between the clones (data not shown).

RESULTS

Eight PARP genes are encoded at four separate loci distributed over four chromosomes; there are two A loci, one B1 locus, and one B2 locus (27–29, 36). Each locus contains a PARP α and β gene arranged in a tandem array. The physical maps of PARP loci in different *T. brucei* stocks were shown to differ by several restriction fragment length polymorphisms (27, 28, 34). Since these polymorphisms could complicate our transcription analysis, we isolated genomic clones of the PARP genes from a *T. brucei* bloodstream variant that is also available as a recently generated insect-form cell line (A. Rattray and L. H. T. Van der Ploeg, unpublished results; see Fig. 1A for physical maps). Restriction fragments containing a PARP A locus and the two B loci were obtained from genomic libraries made with *Pst*I-digested genomic DNA. The loci are named A, B1, and B2 (Fig. 1A) (as in Mowatt and Clayton [27]). Differences in restriction enzyme digestion patterns of the PARP genes in bloodstream and procyclic *T. brucei* variant 118 clone 1 were not detected, indicating that genomic rearrangements are unlikely to control PARP gene transcription (data not shown). The PARP genes are not close to a telomere (34, 36; unpublished results).

Nascent RNA analysis. The extent of a transcription unit

FIG. 1. (A) Restriction enzyme maps of the cloned PARP A locus and two alleles of the PARP B locus. PARP α and β coding sequences (■) and 5′ and 3′ untranslated extensions of the mRNA (□) are indicated. The numbers under the A locus map correspond to fragments generated by a *Sal*I-*Eco*RI-*Spe*I restriction enzyme digestion. The fragments labeled with numbers under the two B locus clones were generated by a *Pst*I-*Pvu*II-*Stu*I restriction enzyme digestion, and those labeled with letters were generated by *Pst*I-*Spe*I-*Stu*I restriction enzyme digestion. The circled fragments hybridized with α-amanitin-resistant nascent RNA. Abbreviations: PI, *Pst*I; S, *Sal*I; RI, *Eco*RI; SpI, *Spe*I; P2, *Pvu*II; StI, *Stu*I. (B) Slot blot of *T. brucei* control genes. Abbreviations: Rib, Plasmid containing the *T. brucei* ribosomal RNA transcription unit; tub, α-β-tubulin clone; 5S, 5S rRNA; PARP, PARP cDNA clone; VSG, VSG 118 gene; tel, telomere repeat; miniexon, miniexon repeat clone; Blue, Bluescript plasmid (see Materials and Methods for descriptions of the plasmids). (C) Nascent RNA analysis of the PARP A locus. The genomic *Pst*I clone was digested with *Sal*I-*Eco*RI-*Spe*I. The ethidium-stained agarose gel is shown on the left (Et); blots of duplicate filters that were hybridized to nascent RNA from procyclic *T. brucei* made in the presence (+) of 1 mg of α-amanitin per ml or in the absence (−) of α-amanitin are shown on the right. The largest band (V+4) contains the vector attached to the last A4 fragment. (D) Nascent RNA analysis of B locus genomic clones. Clones containing the B2 and B1 alleles of the PARP B locus digested with *Pst*I-*Pvu*II-*Stu*I. V, Vector band. (E) Nascent RNA analysis of B locus clones. Clones containing the B2 and B1 alleles of the PARP B locus digested with *Pst*I-*Spe*I-*Stu*I. V, Vector band.

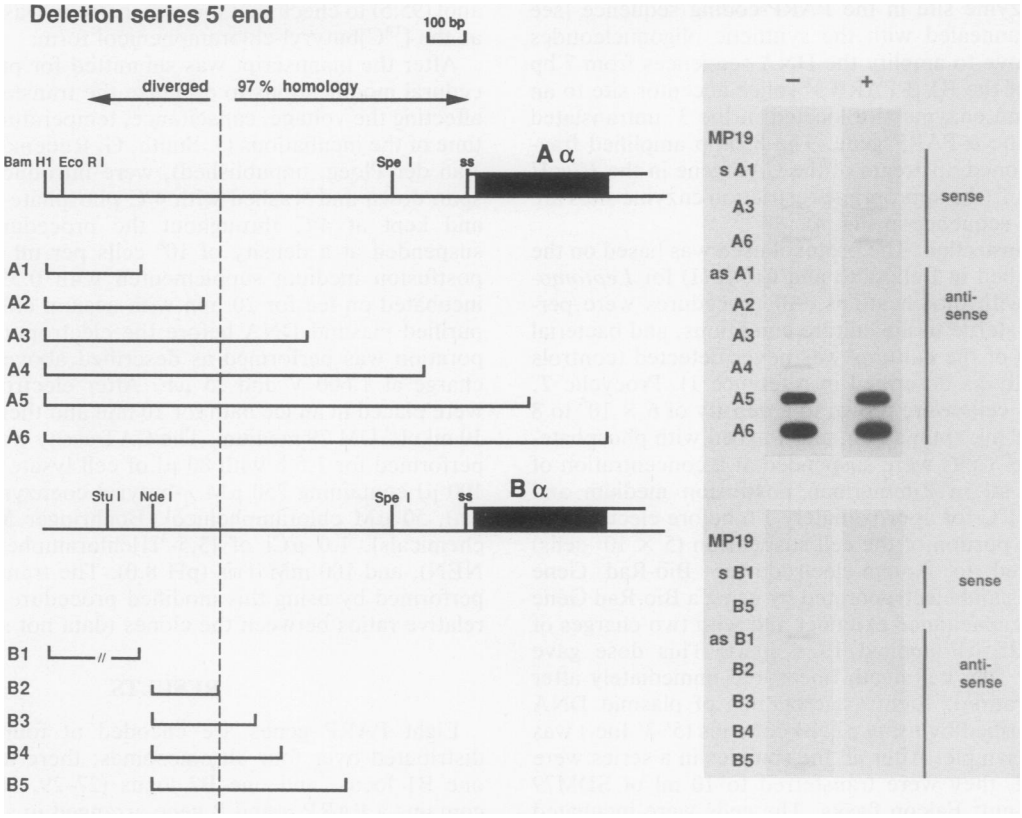


FIG. 2. Nascent RNA run-on analysis of the 5′ end of the PARP transcription units. Nuclear run-on analysis was performed in the presence of 1 mg of α-amanitin per ml (+) or in the absence of α-amanitin (−). *Bal* 31 deletion series of an A locus-derived *Bam*HI-*Pvu*II fragment and a B1 locus-derived *Nde*I restriction enzyme fragment were cloned in two orientations in an M13 mp19 phage vector. The B1 clone contains a 470-bp *Hin*fI restriction fragment. The sizes and locations of the clones are indicated under the map. Single-stranded DNA corresponding to the sense and antisense strands was slot blotted onto nitrocellulose filters and hybridized with nascent RNA from procyclic *T. brucei*. The vertical dotted line demarcates the point in the physical map where the A and B loci diverge; downstream of this point, to the 3′ splice acceptor site, the two loci are 97% homologous. The 3′ splice acceptor site (SS), the coding sequence (■), and the 5′- and 3′-located untranslated extensions of the mRNA (□) are indicated.

cannot be determined by analyzing its steady-state RNA, but it can be analyzed at the nascent RNA level. We thus estimated the size of the PARP gene transcription units by using nuclear run-on assays. These assays show the distribution of RNA polymerases on the DNA template as indicated by ^{32}P labeling of unprocessed RNA by RNA polymerase chain elongation.

Restriction enzyme-digested DNA of the genomic PARP clones was size separated in agarose gels, transferred to filters, and hybridized with ^{32}P-labeled nascent RNA made with procyclic *T. brucei* nuclei in the presence (1 mg/ml; + lanes in Fig. 1B through E) or absence (− lanes in Fig. 1B through E) of the RNA polymerase inhibitor α-amanitin.

Hybridization of nascent RNA to genomic clones of the PARP A locus showed that all the nascent RNA hybridization was confined to the restriction fragments containing the α- and β-PARP-coding sequences (Fig. 1A and C, fragment 4). Three restriction fragments located directly upstream of the α-PARP gene of the A locus did not hybridize with the nascent RNA, indicating that a transcription initiation site may be located immediately upstream of the α-PARP-coding sequence (Fig. 1C, fragments 1, 2, and 3). Similar results were obtained with clones encoding the PARP genes of the B1 and B2 loci (Fig. 1D and E), with hybridization of nascent RNA confined to the PARP-coding region and to sequences located downstream of the PARP genes (Fig. 1E, fragments D and E). The restriction fragments located downstream of the PARP genes in the B loci (Fig. 1D, fragments 6 and 7 in the B1 locus and fragment 6 in the B2 locus, and Fig. 1E, fragment E), showed less hybridization than the restriction fragments containing the PARP-coding sequences. This reduction in hybridization intensities could be due to nucleotide sequence divergence among the loci, which would limit the amount of cross-hybridization. In addition, transcription termination occurs downstream of the β-PARP genes of the B loci (see below). The hybridization patterns were unaffected by the addition of α-amanitin. However, a low level of α-amanitin-sensitive transcription was observed for the upstream fragments in the B loci (Fig. 1E, fragments A, B, and C). We interpret this α-amanitin-sensitive transcription to be the result of readthrough from a different transcription unit located upstream of the α-PARP genes in the B loci (18).

Hybridization of nascent RNA to slot blots containing control *T. brucei* genes (Fig. 1B) showed their predicted α-amanitin sensitivities. Transcription of the rDNA clone (rib) which is transcribed by the α-amanitin-resistant RNA polymerase I is minimally affected by the addition of α-amanitin, as are a PARP cDNA clone and the *T. brucei* telomere repeat clone (36, 37). The hybridization signals of genes transcribed by the α-amanitin-sensitive RNA poly-

merases II and III (α-β-tubulin, miniexon, and 5S rRNA) were reduced to less than 10% of their original levels by the addition of α-amanitin. As expected, the bloodstream-form-specific VSG 118 gene was not expressed in the *T. brucei* insect form.

Location of transcription initiation sites. Restriction fragments from the regions directly upstream of the α-PARP genes did not hybridize with nascent RNA made in the presence of α-amanitin. This result is indicative of the location of transcription initiation sites immediately upstream of the α-PARP-coding sequences. The positions of the 5′ ends of the transcription units were located more precisely by using a series of *Bal* 31 deletion clones in the region upstream of the PARP A loci and the PARP B1 locus (Fig. 2 shows the location of the *Bal* 31 deletion clones). The *Bal* 31 deletion fragments were cloned in both orientations in an M13 mp19 phage vector to determine the strand specificity of transcription. Single-stranded DNA corresponding to the sense and antisense strands was slot blotted onto nitrocellulose filters and hybridized to nascent RNA from procyclic *T. brucei* made in the presence or the absence of α-amanitin (Fig. 2). In all cases, the hybridization to the sense strands of the deletion clones was minimal. There was, however, strong α-amanitin-resistant hybridization with clones containing the antisense fragments A5 and A6. These clones extended 167 and 374 bp downstream of the 3′ splice acceptor site, where the 39-nt miniexon is *trans* spliced to the pre-mRNA. The deletion clone (A4 in Fig. 2) whose 3′ end is 113 bp upstream of the 3′ splice acceptor site did not hybridize to α-amanitin-resistant nascent RNA; neither did the other A locus deletion clones (A1, A2, and A3 in Fig. 2). These data indicate that the transcription initiation site of the α-PARP genes of the A locus is within a few hundred base pairs of the 3′ splice acceptor site.

Our sequence analysis of the genomic clones from the PARP A and B loci had shown that these loci had virtually identical nucleotide sequences extending from the 3′ splice acceptor site to 640 bp upstream. At this point, the nucleotide sequences of the A and B loci diverged (Fig. 2) (see below). We therefore also analyzed single-stranded and B locus-specific deletion clones (B1 and B2 in Fig. 2). As expected, the clones did not hybridize with α-amanitin-resistant nascent RNA.

Quantitation of these experiments showed that the α-amanitin-resistant nascent RNA hybridization to clone A5 was 28-fold higher, and that to clone A6 was 50-fold higher, than the hybridization intensities to fragment A4 or the MP 19 vector control. Since each of the fragments used only recognized the expected PARP loci on Southern blots made with digested *T. brucei* DNA, this transcription is specific for PARP genes, and the initiation site for transcription is therefore likely to be located within the region that extends 200 bp upstream of the 3′ splice acceptor site.

The B PARP locus encodes a small polycistronic transcription unit. The polycistronic transcription units identified in trypanosomes thus far have been large and extended over tens of kilobase pairs. We analyzed the region downstream of the β-PARP genes for possible transcription termination sites which would provide evidence for the presence of well-defined and small polycistronic protein-coding transcription units.

Analysis of transcription termination at PARP loci was performed by using subclones of restriction fragments derived from the region that extended to the 3′ end of the B loci. We had previously determined that transcription at the PARP A locus continued beyond the 3′ end of the PARP A locus genomic clone (results not shown). Slot blots of B-locus plasmid clones were hybridized with nascent RNA from procyclic *T. brucei* made in the presence or in the absence of α-amanitin (Fig. 3). All clones derived from the B1 locus hybridized with nascent RNA. The specificity of this hybridization for the B loci could be determined with clone B1 B, since this fragment was specific for fragments of the B loci and did not detect the A loci (data not shown). Therefore, transcription must continue in at least one of the B loci into regions that extend beyond the point demarcated by this clone. However, at this point the nucleotide sequences of the B loci start to diverge, and the restriction fragments B2 D and B2 E are both specific for the B2 locus (results not shown). Since fragment B2 E does not hybridize at all to the ^{32}P-labeled α-amanitin-resistant nascent RNA, transcription termination of the B2 locus must occur in the region identified by fragments B2 D and B2 E. It is possible that the transcription termination is not abrupt but occurs gradually within the 3′ end of the transcription unit.

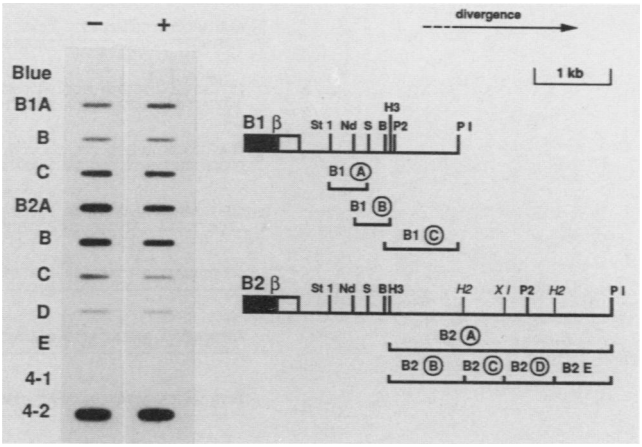


FIG. 3. Run-on transcription analysis of the 3′ end of the PARP B1 and B2 alleles. Slot blots of plasmid subclones were hybridized with nascent RNA from procyclic *T. brucei* made in the presence of 1 mg of α-amanitin per ml (+) or in the absence of α-amanitin (−). The physical map on the right shows the two loci with the relevant subclones underneath. The arrow above the physical map indicates the approximate point of divergence between the B1 and B2 loci. Note that in the map of the B2 locus only the relevant sites are given between the *Hin*dIII and *Pst*I restriction enzyme sites. Fragments that are circled hybridized to nascent RNA. The 490-bp fragment B1 A is generated by digestion with *Stu*I-*Sal*I; B1 B is a 560-bp *Nde*I-*Hin*dIII fragment; B1 C is a 900-bp *Bam*HI-*Pst*I fragment; B2 A is a 2.9-kb *Hin*dIII-*Pst*I fragment; B2 B is a 960-bp *Hin*dIII-*Hin*cII fragment; B2 C is a 530-bp *Hin*cII-*Xho*I fragment; B2 D is a 650-bp *Xho*I-*Hin*cII fragment; and B2 E is a 750-bp *Hin*cII-*Pst*I fragment. Blue is the Bluescript vector; 4-1 and 4-2 are the sense and antisense strands of a PARP cDNA cloned into the M13 single-stranded DNA vector MP18. Abbreviations: B, *Bam*HI; H2, *Hin*cII; H3, *Hin*dIII; P1, *Pst*I; P2, *Pvu*II; Nd, *Nde*I; S, *Sal*I; St1, *Stu*I; X1, *Xho*I.

We conclude that at least one of the PARP transcription units (the B2 locus) is small (5 to 6 kb), with a promoter located within the 200 bp upstream of the α-PARP gene and transcription termination occurring within 3 kb downstream of the β-PARP gene.

Nucleotide sequence analysis of the PARP loci. To study the putative promoter region in more detail, we determined DNA nucleotide sequence of the A and B PARP loci. The DNA nucleotide sequence extending from 1.65 kb upstream of the α-PARP gene of the B1 locus through the intergenic

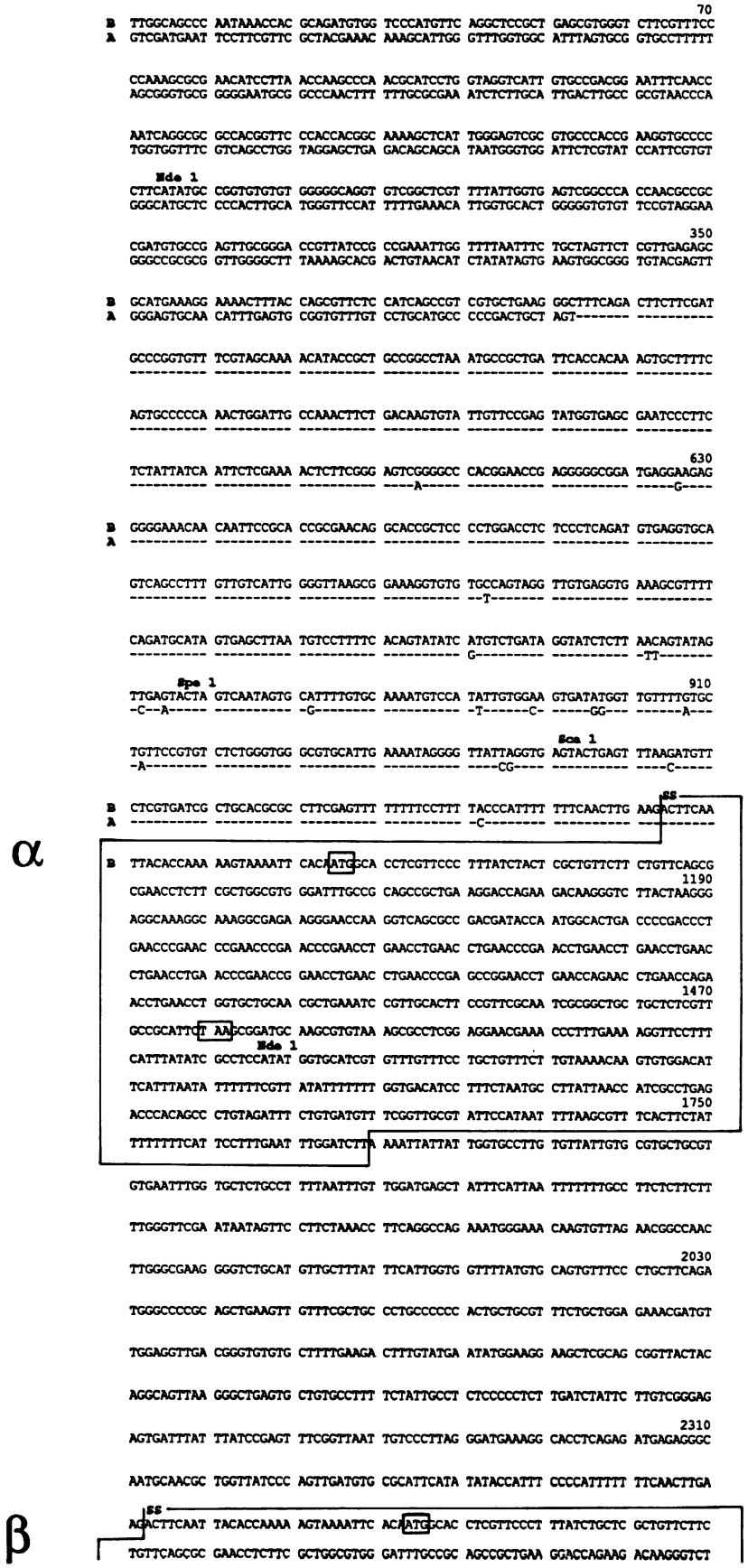


FIG. 4. Sequence analysis of the B1 locus. The region encoding the B1 α and B1 β mRNAs is indicated in large boxes and with single-spaced lines. The sequence of the A locus extending upstream from the α-PARP gene 3′ splice acceptor site is shown underneath the B locus sequence with hyphens representing identical nucleotides. Several relevant restriction enzyme sites are indicated, as are the two 3′ splice acceptor sites of the α- and β-PARP genes (SS). The large box represents the α-PARP gene sequence; the ATG translation initiation codon and the TAA translation termination codon are boxed. The box at the bottom of the figure demarcates the 5′ end of the β-PARP gene.

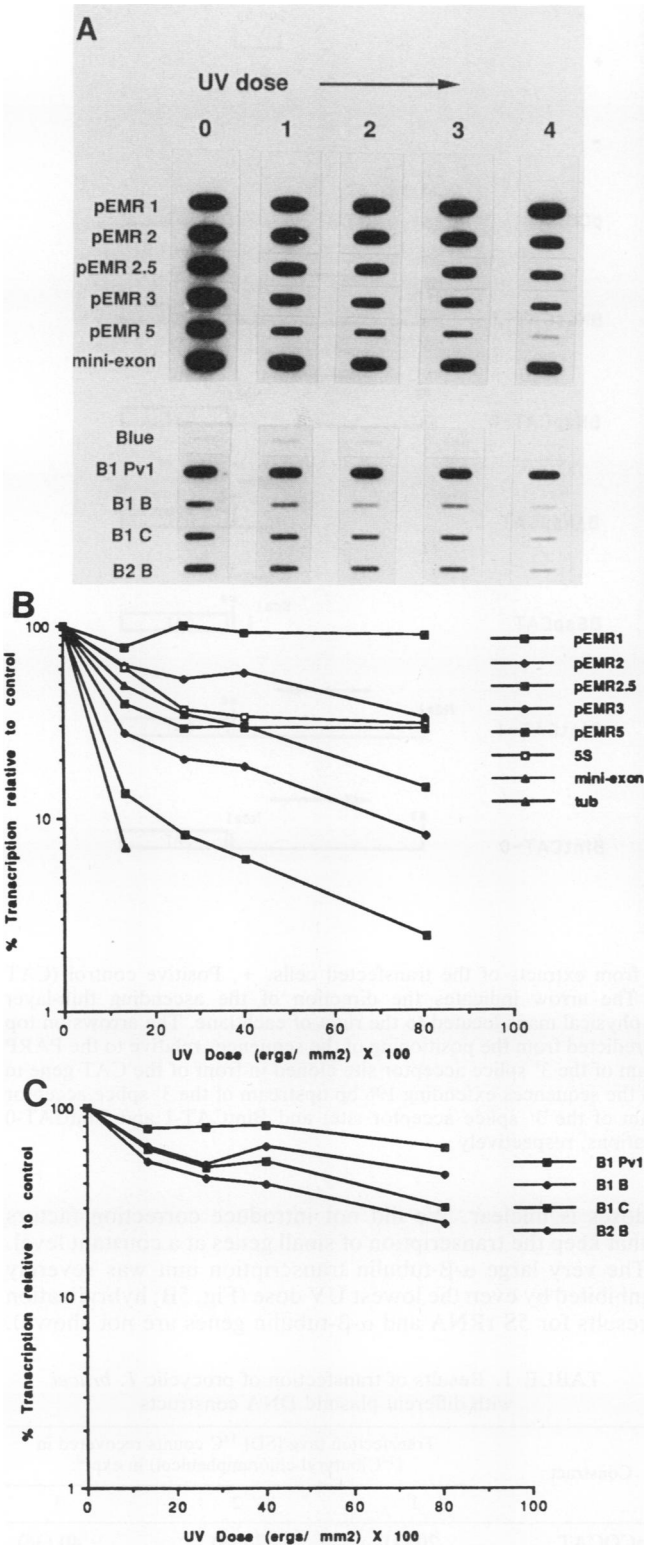


FIG. 5. (A) UV inactivation of transcription in procyclic trypanosomes. Trypanosomes were exposed to increasing doses of UV light, and their nuclei were isolated and used in a nuclear run-on assay to measure the distribution of nascent RNA. The UV doses are given in ergs/mm^2 × 100. The top panel shows the slot blots of subclones from different control genes. The pEMR1-5 clones span the ribosomal DNA transcription unit, and miniexon indicates the miniexon repeat clone. Blue is the plasmid vector Bluescript. B1 Pv 1 is the subclone containing the α-PARP-coding region from the B1 allele (fragments 3 and 4 of the B1 locus in Fig. 1). B1 B and B1 C contain fragments downstream of the B1 β-PARP gene. B2 B is a fragment located downstream of the B2 PARP gene (see map in Fig. 3). The different UV doses were as indicated in Fig. 5B and are here schematically labeled with the numbers 0 through 4 (highest dose). (B) Quantitation of the UV inactivation of transcription data. Semi-logarithmic plot of the relative transcription (signal after UV dose/signal with no UV [as described in Materials and Methods]) set out against the UV dose for subclones from the various control genes. (C) Quantitation of the UV inactivation of the PARP genes. Semi-logarithmic plot of the relative transcription (signal after UV dose/signal with no UV) set out against the UV dose for the PARP subclones. The graphs in panels B and C correspond to the hybridization data shown in panel A. No correction factors were used.

region and into the coding sequence of the β-PARP gene was compared with the A locus sequence (Fig. 4; only 1,043 bp of sequence upstream of the α-PARP genes is shown). Since the B1 and B2 locus physical maps were identical in this region, we did not determine the B2 locus DNA nucleotide sequence. The region extending up to 640 bp upstream from the 3′ splice acceptor site of the α-PARP gene from the B1 locus was 97% homologous to the equivalent region in the A locus sequence. Upstream of this point the nucleotide sequence homology of the A and B loci diverge sharply. As the promoter was estimated to be within 200 bp upstream of the α-PARP gene, the loci may have arisen through gene conversion by a promoter-containing transcription unit.

The DNA nucleotide sequence around the 3′ splice acceptor sites of the B1 α- and β-PARP genes is only identical for 20 bp upstream of the 3′ splice acceptor site. Beyond this point, the nucleotide sequence diverges. Comparison of the putative promoter region and the intergenic region located upstream of the β-PARP gene did not reveal conserved nucleotide sequence elements.

Determination of the location of the transcription initiation site by UV inactivation of transcription. It has been difficult to locate transcription initiation sites in *T. brucei* accurately because of the presence of large polycistronic, perhaps overlapping, transcription units. In addition, presumably all mRNAs in trypanosomes are matured by *trans* splicing. As a result, the 5′ end of the mRNA does not determine the location of the transcription initiation site of the protein-coding gene. We therefore sought to verify the location of the transcription initiation site of the PARP locus, the size of the transcription unit, and the accuracy of the nuclear run-on data by UV inactivation of transcription.

Mapping of transcription initiation sites by UV inactivation of transcription is based on the inability of RNA polymerases to traverse pyrimidine dimers formed on the DNA template by UV irradiation. The level of transcription of a particular gene is thus inversely correlated to the number of pyrimidine dimer cross-links generated in the DNA template and the distance of the fragment from its promoter. At any particular UV dose, DNA sequences located further downstream from a promoter are therefore predicted to be more sensitive to UV irradiation than those located closer to the promoter. Measuring the efficiency of RNA polymerase elongation at different UV doses thus allows mapping of a fragment relative to its transcription initiation site (4, 15–17, 40). This method was used by Borst and co-workers (17) to show that the VSG 221 expression site in bloodstream-form trypanosomes is transcribed as a very large (60-kb) polycistronic unit in bloodstream-form *T. brucei*.

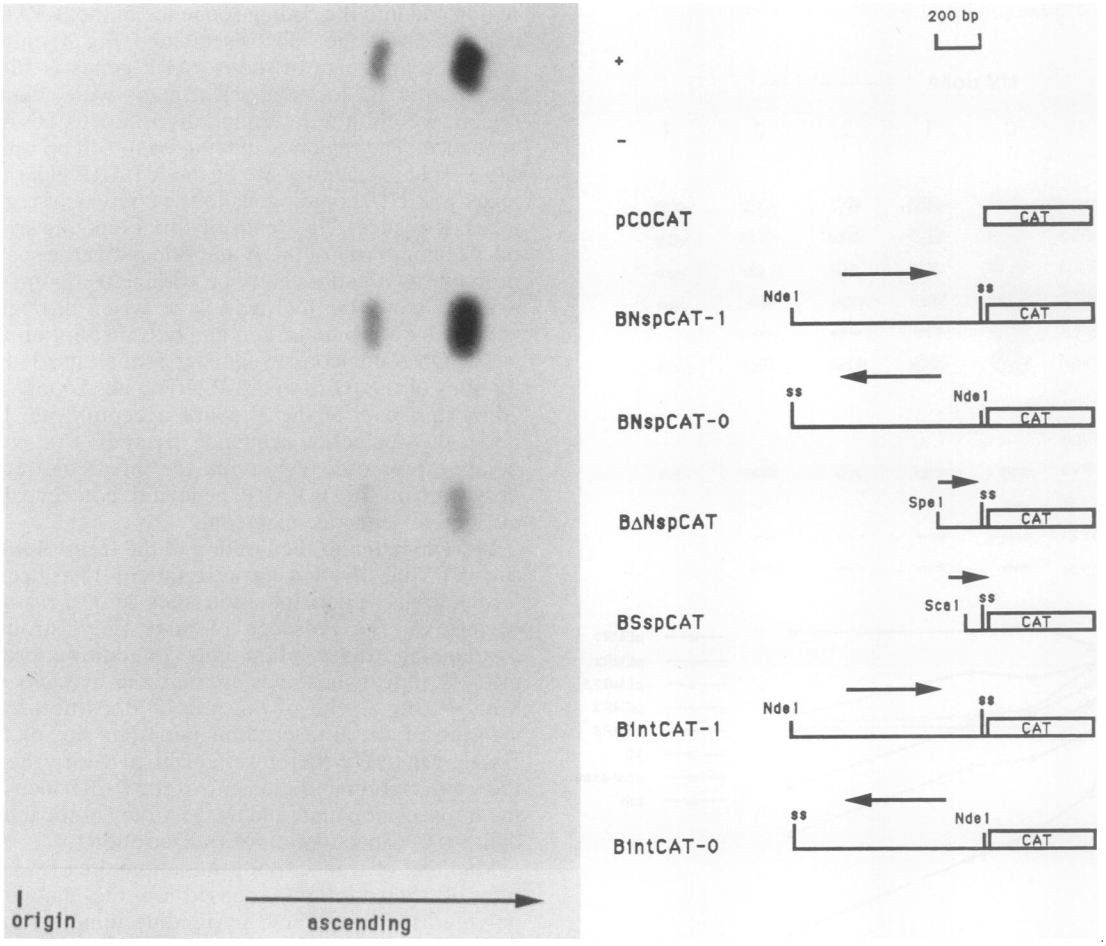


FIG. 6. Thin-layer chromatography of [^{14}C]butyryl-chloramphenicol from extracts of the transfected cells. +, Positive control (CAT enzyme added to a trypanosome cell lysate); −, negative control. The arrow indicates the direction of the ascending thin-layer chromatography. Plasmid constructs are schematically outlined with the physical maps located to the right of each lane. The arrows on top of the physical maps indicate the expected directions of transcription as predicted from the positioning of the sequences relative to the PARP genes. BNspCAT-1 and BNspCAT-0 contained the 830-bp region upstream of the 3′ splice acceptor site cloned in front of the CAT gene in the correct and reverse orientations, respectively. BΔNspCAT contained the sequences extending 196 bp upstream of the 3′ splice acceptor site; the BSspCAT construct contained sequences up to 80 bp upstream of the 3′ splice acceptor site; and BintCAT-1 and BintCAT-0 contained the intergenic region cloned in the correct and inverted orientations, respectively.

Slot-blots of *T. brucei* rDNA clones containing fragments spanning the transcription unit (from 5′ to 3′, pEMR1 to pEMR5) are shown in the top panel of Fig. 5A. Slot blots containing these rDNA clones, the miniexon gene, and α-β-tubulin clones (as described in Materials and Methods) were hybridized with nascent RNA from procyclic *T. brucei* that had been exposed to various doses of UV light before the nuclei isolation. The efficiency of transcription at each UV dose relative to the zero point was set out in a semilogarithmic plot against the UV dose (Fig. 5B). As can be seen from the curves, at increasing UV doses hybridization to the most-downstream rDNA fragment (pEMR5) decreased more rapidly than hybridization to the most-upstream promoter-containing fragment (pEMR1). The hybridization intensity was inversely correlated to the distances of the fragments from the promoter. The 5S rRNA and miniexon subclone containing the promoter and coding sequence of a very small (140-bp) transcription unit were only affected to a limited extent by the increasing UV doses. (The reason for the initial drop in the efficiency of transcription of the 5S rRNA and miniexon genes and the absolute resistance to increasing UV doses is unclear. We did not introduce correction factors that keep the transcription of small genes at a constant level. The very large α-β-tubulin transcription unit was severely inhibited by even the lowest UV dose (Fig. 5B; hybridization results for 5S rRNA and α-β-tubulin genes are not shown).

TABLE 1. Results of transfection of procyclic *T. brucei* with different plasmid DNA constructs

Construct	Transfection (avg [SD] ^{14}C counts recovered in [^{14}C]butyryl-chloramphenicol) in expt[a]:		
	1	2	3
pCOCAT	20 (1)	25 (9)	40 (36)
BNspCAT1	8,941 (985)	5,467 (1269)	3,337 (292)
BNspCAT0	23 (5)	31 (15)	40 (26)
BΔNspCAT	4,892 (352)	1,184 (485)	457 (72)
BSspCAT	45 (2)	26 (5)	33 (23)
BIntCAT1	ND[b]	28 (7)	60 (20)
BIntCAT0	ND	26 (6)	27 (21)

[a] Most experiments were performed in triplicate.
[b] ND, Not done.

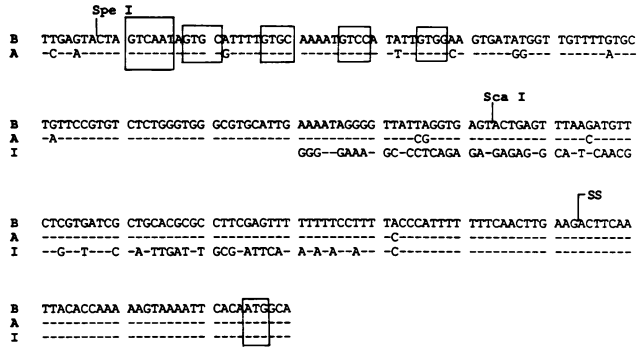


FIG. 7. Nucleotide sequence of the PARP promoter regions. DNA nucleotide sequence comparison of the sequences upstream of the α-PARP genes at the A locus, the B locus, and the B locus-derived intergenic region sequence (I). Hyphens indicate identity in the DNA sequences, and base substitutions are shown. The intergenic region sequence is only shown in the region where sequence conservation exists. SS, 3′ splice acceptor site. *Spe*I and *Sca*I indicate these restriction enzyme digestion sites. Boxed sequences are discussed in the text.

The ribosomal transcription unit was used to calibrate the system, and an estimated size for the transcription unit of 10 kb was obtained (as previously shown by Johnson et al. [17]).

Hybridization of the PARP subclones with nascent RNA showed a similar decrease in the nascent RNA signal, dependent on increasing doses of UV light (hybridization results are shown in the bottom panel of Fig. 5A and the graph shown in Fig. 5C; hybridization to the plasmid control [Blue] was negligible). As expected, hybridization to the subclone containing the putative PARP promoter plus coding region (labeled B1Pv1) was the most resistant to increasing UV doses. Hybridization to two fragments located downstream of the β genes in the two different B loci (fragment B1 C and B2 B in Fig. 3, located at the 3′ end of the B2 PARP locus) were the most sensitive to the UV-induced inactivation of transcription. A fragment located just upstream of B2 B (fragment B1 B) hybridized as expected with a slightly greater intensity. From this data, we can conclude that all the fragments behaved as if they were downstream from a single promoter, located within hundreds of base pairs of the α-PARP-coding sequence. The technique of mapping initiation sites with UV inactivation of transcription is not sensitive enough to determine whether a second promoter exists in the intergenic region upstream of the β-PARP gene.

PARP promoter activity in transient transfection of procyclic *T. brucei*. Dissection of the PARP promoter was performed by analysis of in vivo transient transcription assays. Recently, methods to express the bacterial CAT gene in the trypanosomatid *Leptomonas seymouri* was developed by Bellofatto and Cross (1), and a similar method for *Leishmania enrietii* was developed by Laban and Wirth (21). The Bellofatto and Cross procedure relied on the electroporation of an *L. seymouri* expression vector containing the CAT marker gene with an upstream *L. seymouri*-derived miniexon and miniexon promoter region. Downstream of the CAT gene, these authors placed the 3′ end and downstream region of the *L. seymouri* α-tubulin gene. By using a more sensitive modification (41) of the standard CAT enzyme assay (14), CAT activity was detectable 3 h after transfection.

We used a modification of the same procedure to electroporate CAT constructs containing the putative PARP promoter, using a different vector, into procyclic *T. brucei* (see Materials and Methods for constructs used and protocol). Table 1 shows the results from three representative experiments using PARP CAT gene constructs, and Fig. 6 shows a relevant example of the [^{14}C]butyryl-chloramphenicol as analyzed by thin-layer chromatography. The PARP promoter construct (BNspCAT-1) contained the sequence extending from 7 bp downstream of the α-PARP 3′ splice acceptor site up to 833 bp upstream of the B1 α-PARP gene (see Materials and Methods for details). The CAT vector without a promoter (pCOCAT) did not show CAT activity, exhibiting background levels only, while the putative α-PARP promoter construct containing the promoter in the correct orientation (BNspCAT-1) always showed significant levels of CAT activity (250- to 450-fold above background; transfections were usually performed in triplicate). The same fragment cloned in the inverted orientation (BNspCAT-0) never showed activity above background. A deletion construct (BΔNspCAT) containing only 200 bp upstream of the B1 α-PARP 3′ splice acceptor site gave a CAT activity of 50- to 250-fold above background. Since all *T. brucei* mRNAs are matured by *trans* splicing, the possibility remained that promoter activity was in the plasmid vector and that CAT enzyme activity was found only with constructs containing the 3′ splice acceptor site in the proper orientation. We therefore cloned an 80-bp fragment containing the 3′ splice acceptor site and its adjacent pyrimidine-rich region from the B1 α-PARP gene in front of the CAT gene (construct BSspCAT). This construct gave either no CAT activity or a very slight CAT activity (no more than twofold above background). The experiments presented reflect the range of relative ^{14}C counts obtained with the different constructs. Data for experiments two and three are the most representative of the results obtained from multiple experiments (Table 1 and data not shown). We assume that the variation in the transfection efficiencies resulted from differences in the batches of cells used for each experiment.

We next determined whether a promoter was present in front of the β genes of the polycistronic PARP array. Since sequence conservation could not be found in the region extending beyond 20 bp upstream of the B1 α- and β-PARP genes, these promoters would have to be drastically different in their primary DNA sequences (Fig. 7). A fragment measuring 780 bp (containing, from 5′ to 3′, 170 bp from the 3′-untranslated extension of the B1 locus-derived α-PARP gene, the entire intergenic region, and 7 bp downstream of the 3′ splice acceptor site of the B1-derived β-PARP gene) was cloned in front of the CAT gene in the correct (BintCAT-1) or inverted (BintCAT-0) orientation. Neither construct gave CAT activity above background (Table 1). The transfection results thus indicated that the PARP transcription unit is controlled from a single promoter located upstream of the α-PARP genes.

All our efforts to map the exact location of the transcription initiation site by using procyclic RNA have failed thus far. We have used primer extension, nuclease S1 protection mapping of total RNA, and RNase T1 protection of ^{32}P-labeled nascent RNA. Even though the latter technique has allowed us to map α-β-tubulin and heat shock 70-kilodalton nascent RNA precursor transcripts (J. Huang and L. H. T. Van der Ploeg, unpublished results), PARP nascent RNA transcripts that determine the location of the transcription initiation site could not be detected. We assume that the PARP nascent RNA transcripts are processed more rapidly or that they are extremely unstable. We consider it unlikely

that this instability results from *trans* splicing of nascent RNA, since we have thus far not been able to detect PARP nascent RNA molecules with miniexon sequences (J. Huang and L. H. T. Van der Ploeg, unpublished results).

Description of the α-amanitin-resistant PARP promoter. The putative promoter regions of the A and B loci aligned above the sequence of the intergenic region extending directly upstream of the 3′ splice acceptor site of the B1 β-PARP gene are shown in Fig. 7. Analysis of nascent RNA and UV inactivation of transcription provided evidence for the location of the PARP promoter immediately upstream of the α-PARP gene. Transient transfection of procyclic *T. brucei* pinpointed the putative PARP promoter to the restriction enzyme fragment extending from the *Spe*I restriction enzyme site to the 3′ splice acceptor site (SS). The smaller fragment containing the 3′ splice acceptor site only (construct BSspCAT, extending up to the *Sca*I restriction enzyme site) did not exhibit promoter activity. This data narrows the promoter region to an area of roughly 200 bp upstream of the 3′ splice acceptor site. A canonical CAAT consensus sequence (large box in Fig. 7) is located just downstream of the *Spe*I restriction enzyme site. There were no obvious TATA boxes. A characteristic pattern of four boxes with a GTGC consensus sequence (small boxes in Fig. 7) interspersed with five A's or T's is found downstream of the *Spe*I restriction enzyme site. A 7-bp stem hairpin could be formed in the region just downstream of the second GTGC box. Determination of the significance of these DNA sequence elements for the initiation of transcription awaits analysis of promoter mutant constructs in a CAT transfection assay.

DISCUSSION

At least one of the α-amanitin-resistant PARP transcription units is unusually small (between 5 and 6 kbp), as judged from nuclear run-on hybridization analysis. This is in sharp contrast to the VSG gene expression sites, whose α-amanitin-resistant transcription units range in size from 45 to 60 kb (17, 20, 32, 43, 44). The α-amanitin sensitively transcribed α-β-tubulin and phosphoglycerate kinase transcription units are also large, measuring tens of kilobase pairs in size (13; M. G. S. Lee et al., unpublished results).

The PARP genes contain a promoter located immediately upstream, presumably within 196 bp, of the 3′ splice acceptor site of the α-PARP gene. Since the PARP A and B loci (29) diverge from each other 450 to 550 bp upstream of the putative promoter, all proximal regulatory elements for transcription initiation are probably located within this 550-bp region. Transcription termination occurs within 3 kb downstream of the β gene of the B2 locus. These small polycistronic transcription units generate mature mRNAs with the capped 39-nt miniexon via *trans* splicing of the pre-mRNA. Thus, the function of *trans* splicing is not simply limited to defining the 5′ ends of mature mRNAs in large polycistronically transcribed domains. We do not have evidence for the existence of promoters in the intergenic regions of the PARP transcription unit. The PARP arrays are therefore unlikely to consist of overlapping transcription units and are truly polycistronic. This conclusion is mainly based on the observation that the intergenic region failed to function as a promoter in transient transfection assays. Indirect evidence is provided by the absence of significant sequence homology between the intergenic region and the region upstream of the α-PARP genes. Our results indicate that the polycistronic transcription does not reflect a general absence of transcription terminators at protein-coding genes in trypanosomes, since we identified a region with a transcription terminator in the B2 locus.

The physical maps of the PARP loci were indistinguishable in insect- and bloodstream-form trypanosomes; hence, DNA recombinational events are unlikely to control PARP gene transcription. We could not detect obvious similarities between the PARP promoter region and the putative promoter of the ribosomal RNA transcription unit that is transcribed by RNA polymerase I (51). Similarly, obvious homologies with the region upstream of the α-amanitin sensitively transcribed miniexon genes could not be detected (10). These comparisons are, however, hampered by our inability to determine the exact location of the transcription initiation sites. As discussed above, we assume that the very 5′ end of the PARP precursor RNA is very unstable.

The data presented here and the previously published reports on transfection of *Leptomonas* (1) and *Leishmania* (21) species indicate that it is likely that many different kinetoplastid species can be transformed by electroporation. We have not analyzed the CAT mRNA in the transfected cells and therefore do not know whether a miniexon is added by *trans* splicing. We assume that the addition of the spliced leader is necessary for proper translation (9, 50). Further optimization of the transfection system for *T. brucei* may be required before RNA detection becomes feasible. The characterization of the PARP promoter will aid in the identification of the polymerase that transcribes the PARP and VSG genes, and a comparison of VSG gene promoters and the PARP promoter may allow the identification of characteristic common regulatory domains.

Finally, the analysis of the PARP promoter showed that transcription initiation sites and the 3′ splice acceptor site can be located very close to each other. This juxtaposition implies that the sequence requirements for *trans* splicing are limited to a region that is smaller than the 196-bp domain which contains the PARP promoter and the 3′ splice acceptor site.

ACKNOWLEDGMENTS

We thank David Bishop for help with the DNA sequencing, Jin Huang for help with the analysis of the PARP precursor RNA, Christine Clayton for communicating unpublished results, Piet Borst for critical reading of the manuscript, Piet Borst and colleagues for the gift of ribosomal clones used for the UV inactivation experiments, and Maggi Allen for a gift of plasmid pCOCAT. We also thank Steven Brown, Hui-min Chung, Keith Gottesdiener, Jin Huang, Reggie Kase, Demetrios Vassilatis, and Michael Weiden for critical reading of the manuscript.

This work was supported by Public Health Service grant AI 21784 from the National Institutes of Health and by a grant from the John D. and Catherine T. MacArthur foundation to L.H.T.V.D.P. L.H.T.V.D.P. is a Burroughs Wellcome Scholar in Molecular Parasitology. A.R. was supported by a National Institutes of Health Training Grant fellowship (5T32CA09503-05).

LITERATURE CITED

1. **Bellofatto, V., and G. A. M. Cross.** 1989. Expression of a bacterial gene in a trypanosomatid protozoan. Science **244:** 1167–1169.
2. **Bernards, A., L. H. T. Van der Ploeg, A. C. C. Frasch, P. Borst, J. C. Boothroyd, S. Coleman, and G. A. M. Cross.** 1981. Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3′ end. Cell **27:**163–170.
3. **Borst, P.** 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annu. Rev. Biochem. **55:**701–732.
4. **Brautigam, A. R., and W. Sauerbier.** 1973. Transcription unit

mapping in bacteriophage T7. J. Virol. **12**:882–886.
5. **Brun, R., and M. Schonenberger.** 1979. Cultivation and in vitro cloning of procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Acta Trop. **36**:289–292.
6. **Clayton, C. E.** 1988. The molecular biology of kinetoplastidae. Genet. Eng. News **7**:1–56.
7. **Clayton, C. E., and M. Mowatt.** 1989. The procyclic acidic repetitive proteins of *Trypanosoma brucei*. J. Biol. Chem. **264**:15088–15093.
8. **Cornelissen, A. W. C. A., R. Evers, and J. Kock.** 1988. Structure and sequence of genes encoding subunits for eukaryotic RNA polymerases. Oxf. Surv. Eukaryotic Genes **5**:91–131.
9. **Cornelissen, A. W. C. A., M. P. Verspieren, J.-J. Toulmé, B. W. Swinkels, and P. Borst.** 1986. The common 5′ terminal sequence on trypanosome mRNAs: a target for anti-messenger oligodeoxynucleotides. Nucleic Acids Res. **14**:5605–5614.
10. **De Lange, T., T. M. Berkvens, H. J. G. Veerman, A. C. C. Frasch, J. D. Barry, and P. Borst.** 1984. Comparison of the genes coding for the common 5′ terminal sequence of messenger RNAs in three trypanosome species. Nucleic Acids Res. **12**: 4431–4443.
11. **Evers, R., A. Hammer, J. Kock, W. Jess, P. Borst, S. Memet, and A. Cornelissen.** 1989. *Trypanosoma brucei* contains two RNA polymerase II largest subunit genes with an altered C-terminal domain. Cell **56**:585–597.
12. **Gaffney, D. F., J. McLauchlan, J. L. Whitton, and J. B. Clements.** 1985. A modular system for the assay of transcription regulatory signals: the sequence TAATGARAT is required for herpes simplex virus immediate early gene activation. Nucleic Acids Res. **13**:7847–7863.
13. **Gibson, W., B. W. Swinkels, and P. Borst.** 1988. Post-transcriptional control of the differential expression of phosphoglycerate kinase genes in *Trypanosoma brucei*. J. Mol. Biol. **201**:315–325.
14. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
15. **Hackett, P. B., and W. Sauerbier.** 1974. Radiological mapping of the ribosomal RNA transcription unit in E. coli. Nature (London) **251**:639–641.
16. **Hackett, P. B., and W. Sauerbier.** 1975. The transcriptional organization of the ribosomal RNA genes in mouse L cells. J. Mol. Biol. **91**:235–256.
17. **Johnson, P. J., J. M. Kooter, and P. Borst.** 1987. Inactivation of transcription by UV irradiation of *T. brucei* provides evidence for a multicistronic transcription unit including a VSG gene. Cell **51**:273–281.
18. **Konig, E., H. Delius, M. Carrington, R. O. Williams, and I. Roditi.** 1989. Duplication and transcription of procyclin genes in *Trypanosoma brucei*. Nucleic Acids Res. **17**:8727–8739.
19. **Kooter, J. M., and P. Borst.** 1984. Alpha-amanitin insensitive transcription of variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. Nucleic Acids Res. **12**:9457–9472.
20. **Kooter, J. M., H. J. Van der Spek, R. Wagter, C. E. d'Oliveira, F. Van der Hoeven, P. Johnson, and P. Borst.** 1987. The anatomy and transcription of a telomeric expression site for variant-specific surface antigens in *T. brucei*. Cell **51**:261–272.
21. **Laban, A., and D. F. Wirth.** 1989. Transfection of *Leishmania enrietii* and expression of chloramphenicol acetyl transferase gene. Proc. Natl. Acad. Sci. USA **86**:9119–9123.
22. **Laird, P. W., J. M. Kooter, N. Loosbroek, and P. Borst.** 1985. Mature mRNAs of *Trypansoma brucei* possess a 5′ cap acquired by discontinuous RNA synthesis. Nucleic Acids Res. **13**:4253–4266.
23. **Lee, M. G.-S., and L. H. T. Van der Ploeg.** 1987. Frequent independent duplicative transpositions activate a single VSG gene. Mol. Cell. Biol. **7**:357–364.
24. **Lenardo, M. J., D. M. Dorfman, L. V. Reddy, and J. E. Donelson.** 1985. Characterization of the *Trypanosoma brucei* 5S ribosomal RNA gene and transcript: the 5S rRNA is a spliced leader-independent species. Gene **35**:131–141.
25. **Lowndes, N. F., J. Paul, J. Wu, and M. Allan.** 1989. C-Ha-ras gene bidirectional promoter expressed in vitro: location and regulation. Mol. Cell. Biol. **9**:3758–3770.
26. **Macdonald, R. J., G. H. Swift, A. E. Przybla, and J. M. Chirgwin.** 1987. Isolation of RNA using guanidinium salts. Methods Enzymol. **152**:219–227.
27. **Mowatt, M. R., and C. E. Clayton.** 1987. Developmental regulation of a novel repetitive protein of *Trypanosoma brucei*. Mol. Cell. Biol. **7**:2838–2844.
28. **Mowatt, M. R., and C. E. Clayton.** 1988. Polymorphism in the procyclic acidic repetitive protein gene family of *Trypanosoma brucei*. Mol. Cell. Biol. **8**:4055–4062.
29. **Mowatt, M. R., and C. E. Clayton.** 1989. Variation of tandem repeats in the developmentally regulated procyclic acidic repetitive proteins of *Trypanosoma brucei*. Mol. Cell. Biol. **9**:1332–1335.
30. **Parsons, M., R. G. Nelson, and N. Agabian.** 1984. Antigenic variation in African trypanosomes: DNA rearrangements program immune evasion. Immunol. Today **5**:43–50.
31. **Pays, E.** 1989. Pseudogenes, chimaeric genes and the timing of antigen variation in trypanosomes. Trends Genet. **5**:389–391.
32. **Pays, E., P. Tebabi, A. Pays, H. Coquelet, P. Revelard, D. Salmon, and M. Steinert.** 1989. The genes and transcripts of an antigen gene expression site from *T. brucei*. Cell **57**:835–845.
33. **Richardson, J. P., R. P. Beecroft, D. L. Tolson, M. K. Liu, and T. W. Pearson.** 1988. Procyclin: an unusual immunodominant glycoprotein surface antigen from the procyclic stage of African trypanosomes. Mol. Biochem. Parasitol. **31**:203–216.
34. **Roditi, I., M. Carrington, and M. Turner.** 1987. Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of *Trypanosoma brucei*. Nature (London) **325**:272–274.
35. **Roditi, I., H. Schwartz, T. W. Pearson, R. P. Beecroft, M. K. Liu, J. P. Richardson, H. J. Buhring, J. Pleiss, R. Bulow, R. O. Williams, and P. Overath.** 1989. Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. J. Cell Biol. **108**:737–746.
36. **Rudenko, G., D. Bishop, K. Gottesdiener, and L. H. T. Van der Ploeg.** 1989. Alpha-amanitin resistant transcription of protein coding genes in insect and bloodstream form *Trypanosoma brucei*. EMBO J. **13**:4259–4263.
37. **Rudenko, G., and L. H. T. Van der Ploeg.** 1989. Transcription of telomere repeats in protozoa. EMBO J. **8**:2633–2638.
38. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.
39. **Sanger, F., S. Nicklen, and A. R. Coulsen.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
40. **Sauerbier, W., R. L. Millette, and P. B. Hackett.** 1970. The effects of ultraviolet irradiation on the transcription of T4 DNA. Biochem. Biophys. Acta **209**:368–386.
41. **Seed, B., and J.-Y. Sheen.** 1988. A simple phase-extraction for chloramphenicol acytransferase activity. Gene **67**:271–277.
42. **Sharp, P.** 1987. Trans splicing: variation on a familiar theme. Cell **50**:147–148.
43. **Shea, C., M. G.-S. Lee, and L. H. T. Van der Ploeg.** 1987. VSG gene 118 is transcribed from a cotransposed pol I-like promoter. Cell **50**:603–612.
44. **Shea, C., and L. H. T. Van der Ploeg.** 1988. Stable variant specific transcripts for the VSG gene 1.8 expression site. Mol. Cell. Biol. **8**:854–859.
45. **Smith, J. L., J. R. Levin, C. J. Ingles, and N. Agabian.** 1989. In trypanosomes the homolog of the largest subunit of RNA polymerase II is encoded by two genes and has a highly unusual C-terminal domain structure. Cell **56**:815–827.
46. **Thomashow, L. S., M. Milhausen, W. J. Rutter, and N. Agabian.** 1983. Tubulin genes are tandemly linked and clustered in the genome of *Trypanosoma brucei*. Cell **32**:35–43.
47. **Van der Ploeg, L. H. T.** 1986. Discontinuous transcription and splicing in trypanosomes. Cell **47**:479–480.
48. **Van der Ploeg, L. H. T., A. Bernards, F. A. M. Rijswijk, and P. Borst.** 1982. Characterization of the DNA duplication-transpo-

sition that controls the expression of two genes for variant surface glycoproteins in *Trypanosoma brucei*. Nucleic Acids Res. **10**:593–609.

49. **Van der Ploeg, L. H. T., A. Y. C. Liu, and P. Borst.** 1984. Structure of the growing telomeres of trypanosomes. Cell **36**: 459–468.
50. **Walder, J. A., P. S. Eder, D. M. Engman, S. T. Brentano, R. Y. Walder, D. S. Knutzon, D. M. Dorfman, and J. E. Donelson.** 1986. The 35-nucleotide spliced leader sequence is common to all trypanosome messenger RNA's. Science **233**:569–571.
51. **White, T. C., G. Rudenko, and P. Borst.** 1986. Three small RNAs within the 10 kb trypanosome rRNA transcription unit are analogous to domain VII of other eukaryotic 28S rRNAs. Nucleic Acids Res. **14**:9471–9489.