MOLECULAR AND CELLULAR BIOLOGY, Mar. 1991, p. 1473–1479
0270-7306/91/031473-07$02.00/0

Vol. 11, No. 3

A Similar Gene Is Shared by Both the Variant Surface Glycoprotein and Procyclin Gene Transcription Units of *Trypanosoma brucei*

MAGALI BERBEROF, ANNETTE PAYS, AND ÉTIENNE PAYS*

Department of Molecular Biology, University of Brussels, 67, rue des Chevaux, B1640 Rhode Saint Genèse, Belgium

Received 13 September 1990/Accepted 26 December 1990

The genes for the variant surface glycoprotein (VSG) and procyclin are expressed in a mutually exclusive manner during the life cycle of *Trypanosoma brucei* and synthesize the most abundant mRNAs specific to the bloodstream and procyclic stages of the parasite, respectively. Genes belonging to the polycistronic transcription unit of the VSG gene (expression site-associated genes [ESAGs]) are uniquely expressed in the bloodstream form, but some members of ESAG families (genes related to ESAGs [GRESAGs]) are independently transcribed outside the VSG gene expression site. We report here that a gene related to ESAG 2, GRESAG 2.1, is present and expressed in a procyclin gene transcription unit (PARP A locus), which is polycistronic. Members of the ESAG 2 family are thus present in the two major differentially stage-regulated transcription units of this parasite.

The variant surface glycoprotein (VSG) and procyclin are the major surface proteins of the bloodstream and procyclic stages, respectively, of *Trypanosoma brucei*. Except during differentiation of one life cycle stage to another, these two proteins are never found expressed together (29).

There are several hundred different VSG genes in the trypanosome genome, many of which are telomeric (34). Only one VSG gene is expressed at a time, and the active copy is always present at the 3′ end of a large, multigenic, and telomeric transcription unit (13, 16, 27). The genes belonging to this unit have been called expression site-associated genes (ESAGs) (9). All ESAGs are members of large multigene families (16, 27). Although ESAGs are only expressed in bloodstream forms, at least some members of ESAG families, genes related to ESAGs (GRESAGs), are expressed elsewhere, under separate transcriptional control (1, 2, 27).

The procyclin genes exist in four to seven copies per haploid genome, distributed in two or three clusters of two to three tandemly linked genes, depending on the strain (14, 19–21, 28). All these genes appear to be transcribed (14, 21), and evidence has been presented that at least one of the procyclin gene loci (that for procyclic acidic repetitive protein [PARP]B) is expressed as a small transcription unit encompassing the two linked genes (31).

Expression of the VSG ESAGs and procyclin genes occurs in a mutually exclusive manner, the latter solely in procyclic forms and the former in bloodstream forms. However, a link between these transcription units has been found in terms of the properties of the RNA polymerase involved. In both cases the polymerase is resistant to 1 mg of α-amanitin per ml (6, 14, 15, 24, 30), whereas overall synthesis of mRNAs from other protein-coding genes is sensitive to the drug (4, 15). This property is also characteristic of the transcription of telomeric sequences (32) and may be due to a modification of RNA polymerase II (10). Another characteristic which may be common to both transcription units is that their expression seems to be controlled primarily at the posttranscriptional level. We have found that both transcription promoters are active in the two parasite forms and that differential gene expression is achieved through modification of RNA elongation and stability (7, 12, 23, 24). However, this has not been found by others for the procyclin gene promoter (19, 30).

We present here evidence for an additional link between the VSG and procyclin gene transcription units. Each unit appears to contain a member of the ESAG 2 family. The member present in the procyclin unit (GRESAG 2.1) is located 1.2 kb downstream from the second procyclin gene copy in the PARP A locus. Like the procyclin genes, it is only transcribed in procyclin forms, by a polymerase resistant to α-amanitin.

MATERIALS AND METHODS

T. brucei bloodstream forms were from the AnTat 1.3A clone, and procyclic forms were derived from the AnTat 1.1B clone by in vitro transformation (23).

The procedures for DNA and RNA isolation, Southern and Northern (RNA) blot hybridization, and DNA cloning were as described before (1, 25). cDNA libraries were constructed in λgt10 by the method of Gubler and Hoffman (11), with the Amersham cDNA synthesis and cloning kits. The sequences of DNA fragments, subcloned in bacteriophage M13 derivatives, were determined on both strands by the method of Sanger et al. (33), with a modified T7 DNA polymerase (Sequenase; United States Biochemical Corp.).

Run-on transcription assays were conducted as described by Murphy et al. (22). The standard assay mix (1 ml) contained 500 μg of DNA in nuclei, 12.5% (vol/vol) glycerol, 0.8 mg of heparin, 5 mM spermidine, 5 mM $MgCl_2$, 2.5 mM dithiothreitol, 10 mM Tris-HCl (pH 8), 0.5 mM each ATP, UTP, and CTP, and 1 mCi of [α-^{32}P]GTP (2,000 Ci/mmol). The nuclei were usually incubated for 30 min at 30°C, conditions which gave maximal nucleotide incorporation into RNA (7). Both the UV effect and relative efficiency of transcription between different sequences appeared to be

* Corresponding author.

independent of the incubation period up to 60 min (7; data not shown).

UV irradiation was performed under the conditions defined by Johnson et al. (13). Briefly, the blood containing trypanosomes was diluted in Baltz medium at 37°C to reach a parasite concentration of 1×10^7 to 4×10^7/ml. Samples (125 ml) were irradiated at 254 nm (1 J/s/m^2) in square sterile dishes (22 by 22 cm; Bio-Assay, Nunc, Roskilde, Denmark) usually for 1 min with agitation. The irradiated cells were then transferred to culture flasks and kept in the dark, usually for 30 min, until centrifugation.

Nucleotide sequence accession number. The nucleotide sequence in this article was assigned GenBank accession number M38735.

RESULTS

Cloning of the GRESAG 2.1 cDNA. Data from run-on transcription assays (1, 27) have suggested that sequences related to ESAG 2 were transcribed outside the VSG gene expression site, in both bloodstream and procyclic forms. ESAG 2-specific transcription in bloodstream forms appeared to include both α-amanitin-resistant and -sensitive activities, since the level of inhibition by α-amanitin was reproducibly about 30 to 50% lower under conditions in which transcription of the VSG gene expression site was found to be 100% resistant (1, 27). In procyclic forms, the early data suggested only α-amanitin-sensitive transcription of ESAG 2-related sequences (27), but it was subsequently found that this transcription also included α-amanitin-resistant components (24). Such variability has only been observed in this particular case and is not understood (possible causes of variability include the procedures for cell cultivation and isolation of nuclei). In order to characterize the genes responsible for the ESAG 2-related transcription in procyclic forms, cDNA banks constructed from procyclic mRNA were hybridized with an ESAG 2 probe. From a screening of 5×10^5 recombinants, only 3 specific clones were obtained, all from the same gene. The gene was termed GRESAG 2.1.

Characterization of the GRESAG 2.1 cDNA. The nucleotide sequence of the largest of the three GRESAG 2.1 cDNAs is presented in Fig. 1. This 1.75-kb cDNA appears to be a nearly full-size copy of the mRNA, since it starts with the 5′ spliced leader found ahead of all trypanosome mRNAs (miniexon) and ends with a poly(A) tail. As shown in Fig. 1, the GRESAG 2.1 cDNA is highly homologous to ESAG 2 over approximately 1,400 bp, with some regions (1 to 60, 170 to 290, 340 to 650, 730 to 790, and 840 to 1250) being particularly well conserved. The overall homology is about 70%. The cDNA contains two successive large open reading frames, separated by only three nucleotides.

Predicted GRESAG 2.1 polypeptides. The two open reading frames of the GRESAG 2.1 cDNA encode potential proteins of 199 and 251 amino acids, which can be aligned with the ESAG 2 protein (Fig. 2). The overall amino acid homology between these proteins is about 65%. However, noticeable differences can be found in GRESAG 2.1 polypeptides with respect to both the beginning and the end of the ESAG 2 protein (Fig. 2). The first polypeptide contains a putative signal peptide, with a possible cleavage site according to Von Heijne's program (35), and the second a hydrophobic tail. Both hydrophobic extensions are absent from ESAG 2 (Fig. 3). No significant homology could be found with sequences from data banks.

Transcription of GRESAG 2.1. The GRESAG 2.1 cDNA was used as a hybridization probe on Northern blots of polyadenylated RNA from bloodstream and procyclic forms. In accordance with the length of the cDNA, an mRNA of approximately 1.8 kb could be detected in procyclic RNA (Fig. 4, panel 1). With a probe specific to GRESAG 2.1 (probe d), the 1.8-kb mRNA was only detected in procyclic forms. Other transcripts from the ESAG 2 family include the ESAG 2 mRNA, which is more abundant and is restricted to bloodstream forms (arrowhead), and RNAs common to both forms, suggesting a constitutive expression of other genes from the same family.

In order to measure the level of resistance of GRESAG 2.1 transcription to α-amanitin, the GRESAG 2.1 cDNA was hybridized with run-on transcripts of nuclei from bloodstream and procyclic forms incubated with and without the drug (Fig. 5, panel 1). The actin and procyclin genes were used as controls for α-amanitin-sensitive and -insensitive transcription, respectively (Fig. 5, panel 2). In accordance with the data from Northern blots, the results showed that GRESAG 2.1 and ESAG 2 primarily hybridized with transcripts from procyclic and bloodstream nuclei, respectively. Transcription of GRESAG 2.1 appeared to be fully resistant to 1 mg of α-amanitin per ml, while that of ESAG 2 appeared to be relatively sensitive (panel 1, lanes B and B + am), as reported previously (1, 27).

Cloning GRESAG 2.1. The transcription characteristics of GRESAG 2.1 (specificity for procyclic forms and insensitivity to α-amanitin) were strikingly similar to those of the procyclin genes (Fig. 5, panel 2). This prompted us to evaluate, by Southern blot hybridization, the relative proximity of these different genes in genomic DNA. In some restriction digests, such as with *Bgl*II, probes from the procyclin genes and ESAG 2 and GRESAG 2.1 seemed to recognize the same fragments (arrowhead in Fig. 6) among the different sequences from the ESAG 2 family. Interestingly, these fragments appeared to contain GRESAG 2.1, as they hybridized with a probe specific to GRESAG 2.1 only (Fig. 6, probe d). Such a fragment (10.5 kb *Bgl*II) was cloned into the EMBL4 phage, where it was found to be inserted together with an unrelated 6.5-kb *Bgl*II fragment. The restriction map of this clone is presented in Fig. 7 (panel 1, bottom). From detailed restriction mapping and limited DNA sequencing, the 10.5-kb *Bgl*II fragment turned out to contain both procyclin genes and the genomic GRESAG 2.1. The nucleotide sequence of the latter confirmed the discontinuity in the open reading frame found in the cDNA.

GRESAG 2.1 belongs to the transcription unit of the PARP A locus. As determined by hybridization, GRESAG 2.1 was found to be located 1.2 kb downstream from the second procyclin gene of the PARP A locus. Together with the stage specificity and α-amanitin resistance of the RNA polymerase activity, this particular gene linkage strongly suggested that transcription of GRESAG 2.1 was controlled by the procyclin gene promoter. In order to determine whether all these genes were part of the same transcription unit, the effect of UV irradiation on run-on transcription was analyzed. We have shown before (7, 8, 23, 24, 27) that UV irradiation blocks both RNA elongation and RNA decay. The UV effect on RNA decay was demonstrated by measuring the rate of VSG mRNA disappearance from bloodstream forms after inactivating VSG gene transcription by a cold shock and by measuring the decay rate of several mRNAs by pulse-chase RNA labeling experiments. In both cases, a delay in the degradation of RNA was observed following UV irradiation

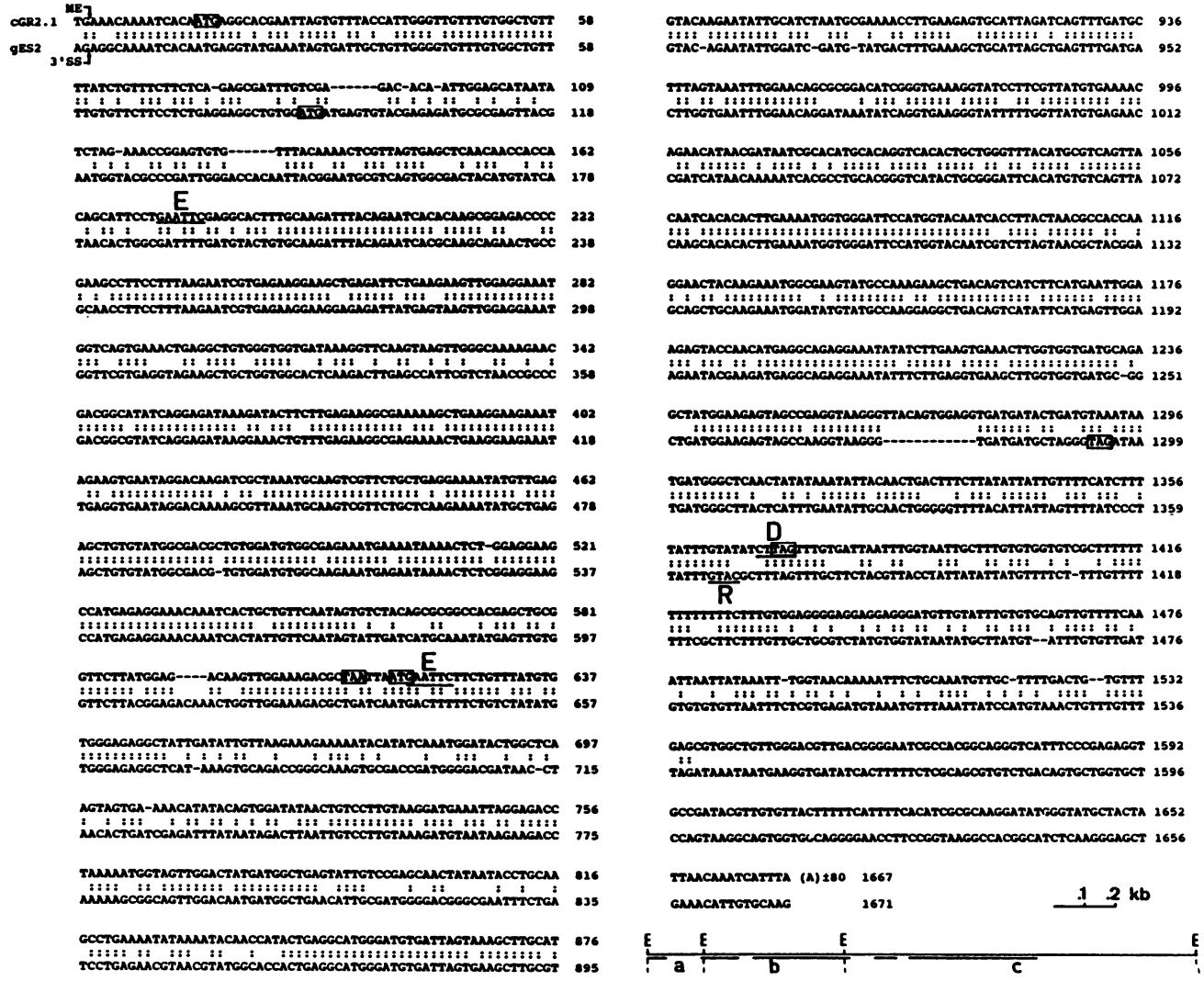

FIG. 1. Nucleotide sequence of the GRESAG 2.1 cDNA, and comparison with ESAG 2. The GRESAG 2.1 cDNA sequence is aligned with that of ESAG 2 in order to achieve maximum homology. Identical nucleotides are indicated by double dots. The 3′ end of the spliced leader (ME, for miniexon) is shown by an arrow, as is the 3′ splice site (3′ SS) of ESAG 2. The initiation and stop codons of the major open reading frames are boxed. A schematic representation of the GRESAG 2.1 cDNA is given below, to indicate the extent of probes a, b, and c, used in Fig. 4 and 6. Regions highly homologous to ESAG 2 are represented by bars under the map. The two gene-internal *Eco*RI sites (E) are underlined in the sequence, as well as *Dde*I (D) and *Rsa*I (R) sites used to generate 3′ gene-specific probes (see Fig. 4); the *Eco*RI sites at the ends of the cDNA were generated by the cloning procedure (11).

(7). The combination of inhibition of both RNA decay and RNA elongation by UV can account for the characteristic accumulation of promoter-proximal transcripts which has been observed for the VSG (8, 23, 24, 27), procyclin (6, 7, 23, 24), rDNA (6, 7, 27), and actin gene (3a) transcription units. Such an apparent stimulation of transcription is thus a marker for the proximity of a promoter, while the degree of inhibition of transcription is indicative of the distance to the promoter.

As shown in Fig. 7 (blots in panel 1, graph in panel 2, thick curve), UV irradiation (60 J/m^2) led to a ca. 1.7-fold apparent stimulation of transcription of the procyclin genes. This result is in accordance with the mapping of a transcription promoter immediately upstream from the procyclin genes (6, 14, 24, 31), in particular in the PARP A locus analyzed here (6). A lower stimulation was observed for GRESAG 2.1 (Fig. 7, panel 2), indicating that this gene is further downstream from its promoter. Taken together with the continuous transcription detected between the second procyclin gene and GRESAG 2.1 (Fig. 7, panel 2), this observation leaves little doubt that all these genes belong to the same transcription unit. The difference in the level of hybridization between the procyclin genes and GRESAG 2.1 (Fig. 7, panel 2, thin curve) is at least partly due to transcription of additional procyclin genes in another locus (14, 21).

Association between procyclin genes and GRESAG 2.1 is not fortuitous. In order to eliminate the possibility that the cotranscription of procyclin genes and GRESAG 2.1 was due to the fortuitous translocation of an ESAG 2 copy to a site adjacent to the procyclin genes only in the *T. brucei* AnTat 1.3A clone, a GRESAG 2.1 probe was hybridized with *Bgl*II digests of trypanosomes from different *T. brucei* stocks and subspecies. As shown in Fig. 6 (panel 2), the same 10.5-kb *Bgl*II fragment hybridized to the probe in *T.*

```
GR2.1 MRHELVFTIGLFVAVLSVSSQSDLSRHNWSIISRNRSVFTKLVSELNNHHSIPEFEALCK 60
ES 2                              MMSVREMRELRMVRPIGTTITECVSGDYMYHNTGDFDVLCK 41

IYRITQAETPKPSFKNREKEAEILKKLEEMVSETEAVGGDKGSSKLGKRTTAYQEIKILL 120
IYRITQAELPQPSFKNREKEGEIMSKLEEMVREVEAAGGTQDLSHSSNRPTAYQEIRKLF 101

EKAKKLKEEIEVNRTRSLNASRSAEENMLRAVYGDAVDVARNENKTLEEAMRGNKSLLFN 180
EKARKLKEEIEVNRTKALNASRSAQENMLRAVYGDVWMWQEMRIKLSEEAMRGNKSLLFN 161

SVYSAATSCGSYGDKLERR*LMN-SSV-YVWERLLILLRKKNTYQMDTGSSSENIYSGYN 237
SIDHANMSCGSYGDKLVGKTLINDFFCLYVGEAHKVQTGQSATDGDDNLTLIE-IYNRLN 220

CPCKDEIRRPKNGSWTMMAEYCPSNYNTCKPENIKYNHTEAWDVISKACMYKNIASNAKT 297
CPCKDVIRRPKSGSWTMMAEHCDGDGRISDPENVTYGTTEAWDVISEACVYKNIGSNVKT 280

LKSALDQFDALVNLEQRGHRVKGILRYVKTEHNDNRTCTGHTAGFTCVSYNHTLENGGIP 357
LKAALAEFDDLVNLEQDKYQVKGIFGYVRTDHNKNHACTGHTAGFTCVSYKHTLENGGIP 340

WYNHLTNATKELQEMAKYAKEADSHLHELEEYQHEAEEIYLEVKLGGDAELWKSSRGKGY 417
WYNRLSNATEQLQEMDMYAKEADSHIHELEEYEDEAEEIFLEVKLGGDAADGRVAKVRVM 400

SGGDDTDVNNDGLNYINITTDFLILLFSSFICIS 451
MLG 403
```

FIG. 2. Comparison of GRESAG 2.1- and ESAG 2-encoded polypeptides. The amino acid sequences of the presumptive polypeptides encoded by GRESAG 2.1 and ESAG 2 are aligned so as to achieve maximum homology. Identical amino acids are indicated by double dots.

brucei brucei, *T. brucei rhodesiense*, and *T. brucei gambiense*, which indicates that the procyclin genes and GRESAG 2.1 did not associate recently in evolution. However, as for other ESAGs (2), GRESAG 2.1 hybridized very poorly to DNA from related kinetoplastids, such as *T. congolense*, *T. vivax*, *T. cruzi*, *Crithidia fasciculata*, and *Leishmania donovani* (data not shown).

Procyclin gene transcription unit extends downstream from GRESAG 2.1. Sequences downstream from GRESAG 2.1 were found to hybridize with run-on transcripts (Fig. 7). As the synthesis of these transcripts resists α-amanitin inhibition and is only slightly stimulated by UV irradiation (Fig. 7, panel 2), it appears likely that these sequences are transcribed from the same promoter as the procyclin genes. Steady-state transcripts from this region (Fig. 4, panel 2) show both bloodstream- and procyclic-form-specific components, as was also observed for ESAG 2 and GRESAG 2. Thus it is possible that other genes, also belonging to multigene families, are present and are transcribed downstream from GRESAG 2.1 in the procyclin gene transcription unit.

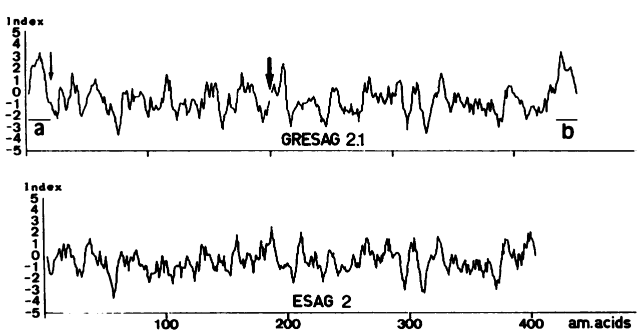

FIG. 3. Hydropathy analysis of the GRESAG 2.1- and ESAG 2-encoded polypeptides. The hydropathy profile of both proteins was determined by the method of Kyte and Doolittle (17), with a window of six amino acids. The relative level of hydrophobicity or hydrophilicity is represented by positive and negative values, respectively. The putative signal peptide of the first GRESAG 2.1 polypeptide and the hydrophobic tail of the second are indicated by the bars marked a and b, respectively. The thin vertical arrow points to the putative cleavage site of the signal peptide, as determined by Von Heijne's method (35). The thick arrow locates the discontinuity in the GRESAG 2.1 open reading frame.

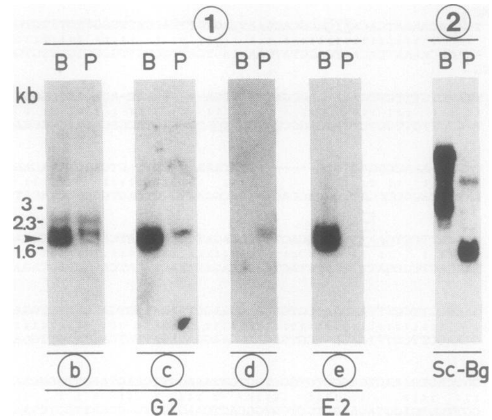

FIG. 4. Steady-state transcripts of the PARP A locus downstream from the procyclin genes. (Panel 1) Transcripts from the ESAG 2 family. A Northern blot of 5 μg of polyadenylated RNA from either bloodstream forms (B) or procyclic forms (P) was hybridized with *Eco*RI fragments from the GRESAG 2.1 cDNA (probes b and c in Fig. 1) and with GRESAG 2.1- and ESAG 2-specific probes (d and e, respectively, corresponding to the 3′-terminal untranslated region of the genes). Probe d is a 400-bp *Dde*I-*Eco*RI fragment from the GRESAG 2.1 cDNA, and probe e is a 500-bp *Rsa*I-*Pst*I fragment from ESAG 2 and 3′-flanking genomic region (the relevant *Dde*I and *Rsa*I sites are underlined in Fig. 1). The ESAG 2 mRNA is designated by an arrowhead. (Panel 2) Transcripts from sequences downstream of GRESAG 2.1. The Northern blot in panel 1 was hybridized with a 1.7-kb *Sca*I-*Bgl*II (Sc-Bg) fragment from the genomic region downstream of GRESAG 2.1 (fragment C7 in Fig. 7).

DISCUSSION

We report here the presence of closely related but different genes in the transcription units coding for the major stage-specific surface proteins of *T. brucei*, the VSG in bloodstream forms and procyclin in procyclic forms.

Although the end products of the VSG and procyclin gene transcription units are expressed in a mutually exclusive manner at different life cycle stages, and despite the fact that the encoded surface proteins are totally different, similarities exist between the two units. The RNA polymerase, which fully resists α-amanitin inhibition, seems to be of the same unusual type in both cases (6, 14, 15, 30). The presence of a related gene in both units strengthens this relationship, although the meaning of this observation is obscure. The following points are worth noting. (i) Only one procyclin gene locus, the PARP A locus, contains GRESAG 2. No evidence could be found for the presence of an ESAG 2-related sequence in PARP B (31). (ii) As suggested by data from both Northern blots and run-on transcription assays, other GRESAGs 2 seem to be transcribed elsewhere in the genome, by a polymerase sensitive to α-amanitin. Similar data have been obtained for the ESAG 4 family (2). (iii)

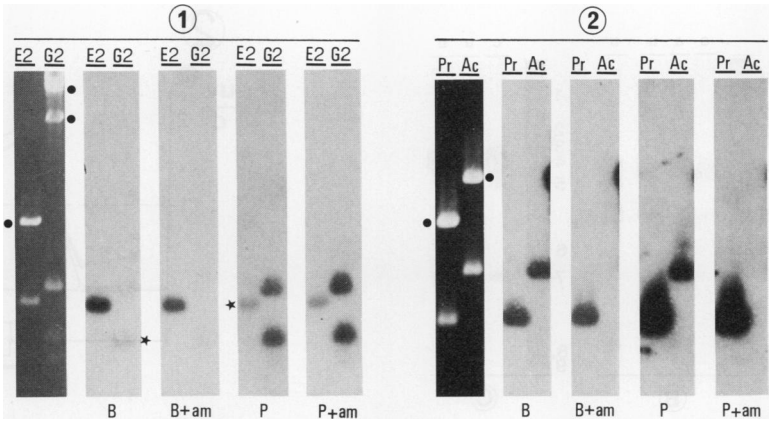

FIG. 5. Procyclic-specific and α-amanitin-insensitive transcription of GRESAG 2.1. In panel 1, run-on transcripts from bloodstream (B) or procyclic (P) nuclei, incubated in the presence of α-amanitin (am, 1 mg/ml) where indicated, were hybridized to Southern blots of either *Hin*dIII-digested plasmid pES200.3, containing ESAG 2 (lanes E2) (1), or *Eco*RI-digested λgt10 DNA containing the GRESAG 2.1 cDNA (lanes G2). The DNA fragments from the plasmid and phage vectors are marked by solid circles, while stars indicate fragments showing possible cross-hybridization with heterologous transcripts from the ESAG 2 family (see text). Panel 2 shows the controls. Run-on transcripts were hybridized to either *Eco*RI digests of a plasmid carrying the procyclin cDNA (lanes Pr) (28) or *Sal*I digests of a plasmid carrying the *T. brucei* actin gene (lanes Ac) (4). Transcription of the actin genes is sensitive to α-amanitin (am) and occurs at the same level in both trypanosome forms, while that of the procyclin genes is resistant to α-amanitin and occurs at a sixfold higher level in procyclic than in bloodstream forms (24).

Among the different gene families of the AnTat 1.3A ESAGs, only members of the ESAG 2 family appear to be shared by the procyclin and VSG transcription units. (iv) The function of ESAG 2 and GRESAG 2.1 is totally unknown. These genes do not have any significant homology with known sequences. Translation of the gene sequences shows only one putative protein for ESAG 2 but two for GRESAG 2.1. Whether only one or both of the GRESAG 2.1 open reading frames can be translated as protein remains unclear. The probable presence of a signal peptide in the first GRESAG 2.1 protein and of a hydrophobic tail in the second suggests interaction with a membrane, which is not the case for the ESAG 2 protein. Heterologous expression of either of the ESAG 2 or GRESAG 2.1 proteins in the alternative life cycle stage, by means of the recently developed DNA transfection systems for kinetoplastid protozoa (3, 6, 12, 18, 31), should help to settle this question.

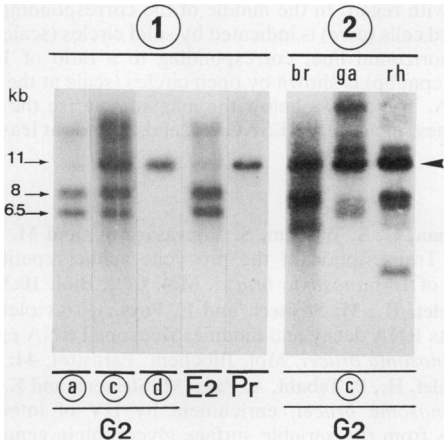

FIG. 6. Genomic hybridization patterns of GRESAG 2.1. (Panel 1) Comparison with ESAG 2 and procyclin probes. Southern blots of *Bgl*II-digested AnTat 1.3A genomic DNA were hybridized with three probes from the GRESAG 2.1 cDNA(G2 a, c, and d; the extent of a and c is indicated in Fig. 1, and probe d is from the *Dde*I [D] site to the end of the GRESAG 2.1 cDNA; see Fig. 1), and with probes specific to the ESAG 2 (E2) and procyclin (Pr) genes. (Panel 2) Comparison of the hybridization pattern in different *T. brucei* subspecies. Probe c was hybridized to *Bgl*II digests of *T. brucei brucei* FEO(s) (br), *T. brucei gambiense* LiTat 1.6 (ga), and *T. brucei rhodesiense* AnTat 12.2 (rh). Posthybridization washes were carried out at high stringency (1.5 mM NaCl, 65°C). A fragment seemingly encompassing both the GRESAG 2.1 and procyclin genes is shown by an arrowhead.

Our data suggest that the PARP A procyclin gene transcription unit includes different genes and may thus be polycistronic. The length of this unit cannot be determined as yet, but our Northern blot and run-on transcription data suggest that it may encompass additional genes. This organization differs from that of the PARP B locus, in which the transcription unit appears to contain only the two tandemly linked procyclin genes (31). One may speculate that the PARP B unit is a truncated copy of the full procyclin gene expression site (PARP A), which, like that of the VSG gene, carries a battery of associated genes. That the procyclin gene loci have a tendency to duplicate or translocate is indicated by the fact that some strains of trypanosomes seem to show an additional procyclin gene locus (14). All these loci seem to be transcribed, reflecting duplication of the transcription promoter which is present immediately upstream from the first gene of each locus (6, 24, 31). However, the variation in the number of loci between strains would imply that the activity of at least some loci is not essential.

ACKNOWLEDGMENTS

We thank D. Jefferies for comments on the manuscript, S. Rolin and D. Jefferies for providing procyclic trypanosomes, P. Tebabi for help in DNA sequencing, and I. Roditi (Bern) for the procyclin cDNA probe.

This work was supported by the Belgian Fonds de la Recherche Scientifique Médicale and research contracts between the University of Brussels and the "Communauté Française de Belgique"

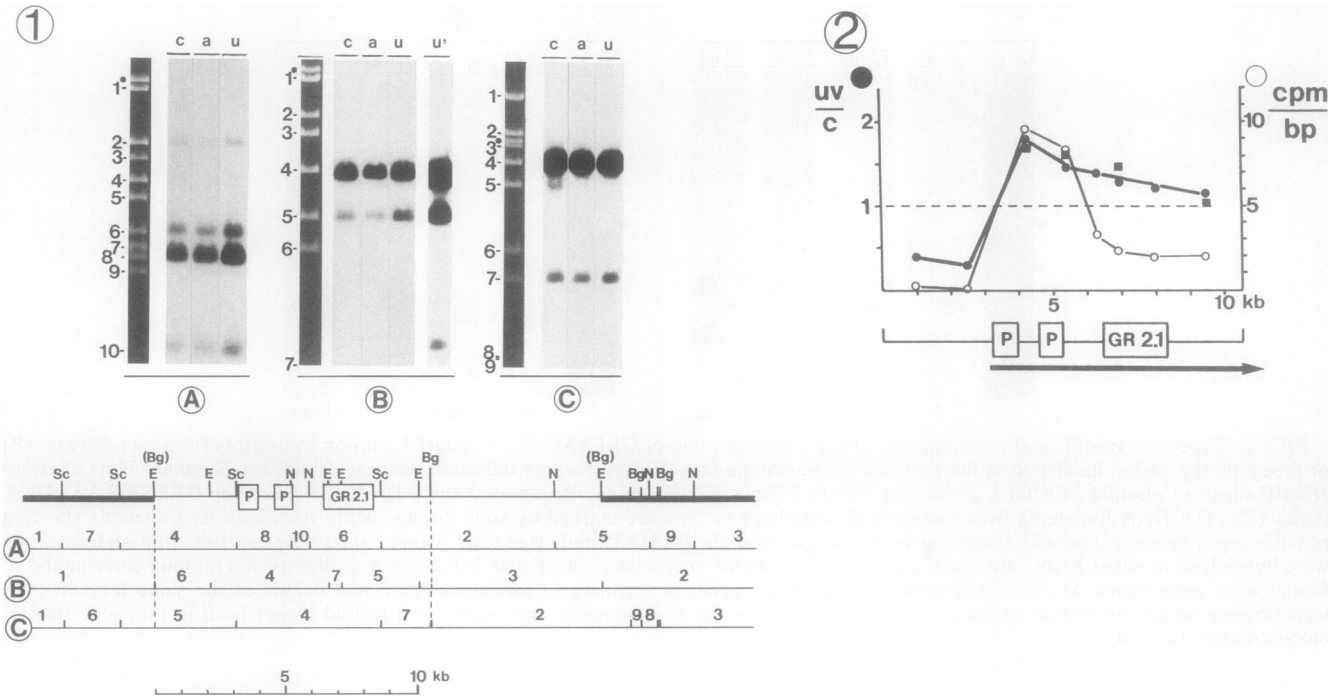

FIG. 7. Procyclin genes and GRESAG 2.1 belong to the same transcription unit. In panel 1, run-on transcripts from procyclic nuclei were hybridized with restriction digests of the EMBL4 recombinant containing GRESAG 2.1. For each digest (A, B, and C), the first lane shows ethidium bromide staining of the DNA fragments, whose numbering corresponds to that in the maps below (fragments marked with asterisks are phage arms annealed through their cohesive termini). The transcripts were from nuclei isolated from either control (c) or UV-irradiated (u) trypanosomes, incubated in the presence of α-amanitin (1 mg/ml) where indicated (a). The lane marked u′ in B shows a fivefold-longer exposure of lane u to reveal hybridization to fragment 7. In the maps, GRESAG 2.1 (GR2.1) and the procyclin genes (P) are boxed, while the bacteriophage arms are indicated by thick lines. Abbreviations for restriction endonuclease sites: Bg, *Bgl*II; E, *Eco*RI; N, *Nde*I; Sc, *Sca*I. The two *Bgl*II sites in brackets were destroyed by the cloning procedure, and the two *Eco*RI sites present instead are from the bacteriophage. Only the region between the dotted vertical lines is specific to the procyclin gene transcription unit and its environment. In the genome, the 6.5-kb region to the right of the central *Bgl*II site is not contiguous with the 10.5-kb *Bgl*II fragment (data not shown), the two having been artifactually cloned together. Panel 2 shows a quantitative plot of the data in panel 1. The nitrocellulose areas containing the hybridized DNA fragments were cut out and counted by liquid scintillation. Each value is plotted with regard to the middle of the corresponding restriction fragment. The ratio of hybridization between samples from UV-treated and untreated cells (uv/c) is indicated by solid circles (scale at the left). Solid squares are values obtained in an independent experiment. The dotted horizontal line, corresponding to a ratio of 1, separates stimulation and inhibition by UV. The level of radioactivity per base pair of DNA (cpm/bp) is shown by open circles (scale at the right). This reflects the relative level of hybridization of the run-on transcripts along the DNA. The arrows below the map summarize the results; the transcription unit of the procyclin genes starts immediately upstream from these genes, includes GRESAG 2.1, and extends at least to the end of the 10.5-kb *Bgl*II fragment.

(Convention ARC no. 89/94-134), Solvay & Cie (Brussels), and the Commission of the European Communities (TDS-M-023B).

REFERENCES

1. **Alexandre, S., M. Guyaux, N. B. Murphy, H. Coquelet, A. Pays, M. Steinert, and E. Pays.** 1988. Putative genes of a variant-specific antigen gene transcription unit in *Trypanosoma brucei*. Mol. Cell. Biol. **8**:2367–2378.
2. **Alexandre, S., P. Paindavoine, P. Tebabi, A. Pays, S. Halleux, M. Steinert, and E. Pays.** 1990. Differential expression of a family of putative adenylate/guanylate cyclase genes in *Trypanosoma brucei*. Mol. Biochem. Parasitol. **43**:279–288.
3. **Bellofatto, V., and G. A. M. Cross.** 1989. Expression of a bacterial gene in a trypanosomatid protozoan. Science **244**: 1167–1169.

3a. **Ben Amar, F., and E. Pays.** Unpublished data.

4. **Ben Amar, M. F., A. Pays, P. Tebabi, B. Dero, T. Seebeck, M. Steinert, and E. Pays.** 1988. Structure and transcription of the actin genes of *Trypanosoma brucei*. Mol. Cell. Biol. **8**:2166–2176.
5. **Borst, P.** 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annu. Rev. Biochem. **55**:701–732.
6. **Clayton, C. E., J. P. Fueri, J. E. Itzakhi, V. Bellofatto, D. R. Sherman, G. S. Wisdom, S. Vijayasarathy, and M. R. Mowatt.** 1990. Transcription of the procyclic acidic repetitive protein genes of *Trypanosoma brucei*. Mol. Cell. Biol. **10**:3036–3047.
7. **Coquelet, H., M. Steinert, and E. Pays.** Ultraviolet irradiation inhibits RNA decay and modifies ribosomal RNA processing in *Trypanosoma brucei*. Mol. Biochem. Parasitol. **44**:33–42.
8. **Coquelet, H., P. Tebabi, A. Pays, M. Steinert, and E. Pays.** 1989. *Trypanosoma brucei*: enrichment by UV of intergenic transcripts from the variable surface glycoprotein gene expression site. Mol. Cell. Biol. **9**:4022–4025.
9. **Cully, D. F., H. S. Ip, and G. A. M. Cross.** 1985. Coordinate transcription of variant surface glycoprotein genes and an expression site-associated gene family in *Trypanosoma brucei*. Cell **42**:173–182.
10. **Grondal, E. J. M., R. Evers, K. Kosubek, and A. W. C. A. Cornelissen.** 1989. Characterization of the RNA polymerases of *Trypanosoma brucei*: trypanosomal mRNAs are composed of transcripts derived from both RNA polymerase II and III. EMBO J. **8**:3383–3389.
11. **Gubler, U., and B. J. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene **25**:263–269.
12. **Jefferies, D., P. Tebabi, and E. Pays.** 1991. Transient activity assays of the *Trypanosoma brucei* VSG gene promoter: control

of gene expression at the posttranscriptional level. Mol. Cell. Biol. **11**:338–343.

13. **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.
14. **Koenig, 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.
15. **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.
16. **Kooter, J. M., H. J. Van der Spek, R. Wagter, C. E. d'Oliveira, F. Van der Hoeven, 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.
17. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
18. **Laban, A., and D. F. Wirth.** 1989. Transfection of *Leishmania enrietii* and expression of chloramphenicol acetyltransferase gene. Proc. Natl. Acad. Sci. USA **86**:9119–9123.
19. **Mowatt, M. R., and C. E. Clayton.** 1987. Developmental regulation of a novel repetitive protein of *Trypanosoma brucei*. Mol. Cell. Biol. **7**:2838–2844.
20. **Mowatt, M. R., and C. E. Clayton.** 1988. Polymorphism in the procyclic acidic repetitive protein (PARP) gene family of *Trypanosoma brucei*. Mol. Cell. Biol. **8**:4055–4062.
21. **Mowatt, M. R., G. S. Wisdom, 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.
22. **Murphy, N. B., A. Pays, P. Tebabi, H. Coquelet, M. Guyaux, M. Steinert, and E. Pays.** 1987. *Trypanosoma brucei* repeated element with unusual structural and transcriptional properties. J. Mol. Biol. **195**:855–871.
23. **Pays, E., H. Coquelet, A. Pays, P. Tebabi, and M. Steinert.** 1989. *Trypanosoma brucei*: posttranscriptional control of the variable surface glycoprotein gene expression site. Mol. Cell. Biol. **9**:4018–21.
24. **Pays, E., H. Coquelet, P. Tebabi, A. Pays, D. Jefferies, M. Steinert, E. Koenig, R. O. Williams, and I. Roditi.** 1990. *Trypanosoma brucei*: constitutive activity of the VSG and procyclin gene promoters. EMBO J. **9**:3145–3151.
25. **Pays, E., M. Delronche, M. Lheureux, T. Vervoort, J. Bloch, F. Gannon, and M. Steinert.** 1980. Cloning and characterization of DNA sequences complementary to messenger ribonucleic acids coding for the synthesis of two surface antigens of *Trypanosoma brucei*. Nucleic Acids Res. **8**:5965–5981.
26. **Pays, E., and M. Steinert.** 1988. Control of antigen gene expression in African trypanosomes. Annu. Rev. Genet. **22**:107–126.
27. **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.
28. **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.
29. **Roditi, I., H. Schwartz, T. W. Pearson, R. P. Beecroft, M. K. Liu, J. P. Richardson, H. Büring, J. Pleiss, R. Bülow, 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.
30. **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. **8**:4259–4263.
31. **Rudenko, G., S. Le Blancq, J. Smith, M. G. S. Lee, A. Rattray, and L. H. T. Van der Ploeg.** 1990. 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*. Mol. Cell. Biol. **10**:3492–3504.
32. **Rudenko, G., and L. H. T. Van der Ploeg.** 1989. Transcription of telomere repeats in protozoa. EMBO J. **8**:2633–2638.
33. **Sanger, F., S. Nicklen, and A. R. Coulsen.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
34. **Van der Ploeg, L. H. T., D. Valerio, T. De Lange, A. Bernards, P. Borst, and F. G. Grosveld.** 1982. An analysis of cosmid clones of nuclear DNA from *Trypanosoma brucei* shows that the genes for variant surface glycoproteins are clustered in the genome. Nucleic Acids Res. **10**:5905–5923.
35. **Von Heijne, G.** 1983. Patterns of amino acids near signal sequence cleavage sites. Eur. J. Biochem. **133**:17–21.