MOLECULAR AND CELLULAR BIOLOGY, Oct. 1995, p. 5598–5606 Vol. 15, No. 10
0270-7306/95/$04.00+0

Specific Binding of Proteins to the Noncoding Strand of a Crucial Element of the Variant Surface Glycoprotein, Procyclin, and Ribosomal Promoters of *Trypanosoma brucei*

LUC VANHAMME, ANNETTE PAYS, PATRICIA TEBABI, SYLVIE ALEXANDRE, AND ETIENNE PAYS*

Department of Molecular Biology, University of Brussels, B-1640 Rhode Saint Genèse, Belgium

Received 16 March 1995/Returned for modification 4 May 1995/Accepted 20 July 1995

The variant surface glycoprotein (VSG) and procyclin promoters of *Trypanosoma brucei* recruit an RNA polymerase sharing characteristics with polymerase I, but there is no sequence homology between them nor between these promoters and ribosomal promoters. We report the detailed characterization of the VSG promoter. The 70-bp region upstream of the transcription start site was sufficient for full promoter activity. Mutational analysis revealed three short critical stretches at positions −61 to −59 (box 1), −38 to −35 (box 2), and −1 to +1 (start site), the spacing of which was essential. These elements were conserved in the promoter for a metacyclic VSG gene. Hybrid sequences containing box 1 of the VSG promoter and box 2 of the ribosomal promoter were active. A specific binding of proteins to the noncoding strand of box 2, but not to double-stranded DNA, occurred. Competition experiments indicated that these proteins also bind to the corresponding region of the metacyclic VSG, procyclin, and ribosomal promoters. Binding of such a protein, of 40 kDa, appeared to be shared by these promoters.

The protozoan parasite *Trypanosoma brucei* shares its life cycle between mammals and an insect vector, the tsetse fly. The bloodstream (mammalian) form and the procyclic (insect) form are uniformly covered by a surface coat, of variant surface glycoprotein (VSG) and procyclin, respectively. The VSG replaces procyclin when procyclic forms differentiate into metacyclic forms in the salivary glands of the fly. This surface antigen persists throughout the development of the parasite in the blood and rapidly disappears when procyclin is reexpressed upon differentiation from the bloodstream to the procyclic form. Therefore, the VSG and procyclin are typical markers for the major developmental stages of the parasite (for recent reviews, see references 4, 10, 24, and 25).

African trypanosomes are responsible for important human and animal plagues. They escape the immune defenses of their hosts by a periodical change of their antigenic VSG coat. The study of the control of VSG expression is thus of major fundamental and economic importance. However, the mechanisms controlling gene expression in trypanosomatids are poorly understood. This is particularly true for transcriptional controls. Because of the general organization of genes into polycistronic transcription units, there is only limited information on transcription promoters and their controls. The only promoters for protein-coding genes that are known so far are those of VSG and procyclin (11, 12, 23, 26, 31, 32). There are two diploid loci for the procyclin transcription units, and it is believed that the promoters of all these units are simultaneously active in the procyclic form (11, 20). The number of VSG transcription units is estimated to be between 6 and 20 (4, 24), but only a single one is expressed at a given time. So far it is not clear if this control operates at the level of promoter activity or transcription attenuation. Interestingly, although the VSG and procyclin units are expressed in a mutually exclusive manner during the trypanosome life cycle, their promoters show constitutive activity. In particular, the VSG promoter is active in procyclic forms, not only as measured in transient assays of episomal reporter constructs (19, 31, 32) but also as measured in situ (23). The down-regulation of the VSG unit during this stage was found to occur primarily at the level of RNA elongation (23, 29).

Detailed analyses of the procyclin promoter suggested similarity with ribosomal promoters (5, 16, 28). This conclusion is interesting in view of the unusual characteristics of the RNA polymerase transcribing the procyclin genes. Because of its resistance to α-amanitin and Sarkosyl, this polymerase was suggested to be Polymerase I (9, 27). The same hypothesis was also put forward for the polymerase for the VSG genes (27, 31). However, this was contradicted by some observations, such as a difference in divalent cation requirement between this polymerase and Polymerase I (15, 24).

A refined analysis of the VSG promoter could shed light on the debate. We present a detailed picture of this promoter, using deletional and mutational analyses. Our results indicate that besides the transcription start site, two short stretches are critical for activity. This organization shows overall similarity to the procyclin and rDNA (DNA coding for rRNA) promoters, despite a lack of sequence conservation. The analysis of hybrid promoters and of DNA-binding complexes supports the conclusion that the VSG, procyclin, and ribosomal promoters are related. These results confirm that at least some of the factors required for the activity of these promoters are present at different stages of the trypanosome life cycle. Furthermore, our data suggest that crucial sequence elements bind specific proteins when single stranded.

MATERIALS AND METHODS

Trypanosomes. Procyclic forms were obtained by in vitro cultivation of isolates from the midgut of flies infected with the EATRO 1125 stock of *T. brucei*. They were grown in SDM-79 medium (7) supplemented with 15% heat-inactivated fetal calf serum. Bloodstream forms were passaged in mice or cultivated according to the method of Baltz et al. (2).

Plasmid constructs. The promoter reporter constructs were all derived from pD5 or pD5V, which contain the chloramphenicol acetyltransferase (CAT) gene preceded by the splice site of ESAG 7 and followed by the 3′-end region of a VSG gene in pD5V only (19). The VSG promoter region (0.27-kb *Dra*I-*Sal*I

* Corresponding author. Mailing address: Department of Molecular Biology, University of Brussels, 67, rue des Chevaux, B-1640 Rhode Saint Genèse, Belgium. Phone: 32-2-650 96 21. Fax: 32-2-650 96 25.

fragment, from −188 to +81 with respect to the initiation site) was subcloned into M13 for site-directed mutagenesis, using an Amersham kit according to the manufacturer's instructions. The altered promoter was then cloned back into the original plasmid for CAT activity assays. A complete check of the promoter nucleotide sequence was performed before and after this last cloning. Alternatively, synthetic oligonucleotides covering both strands of the −69 to +6 region of the promoter were annealed and cloned between the *Sph*I and *Stu*I sites of pD5, thus replacing the original promoter region between −188 and +940 with the synthetic promoter. Again, the CAT assays were performed only after a complete verification of the nucleotide sequence of the constructs.

Transfections and CAT assays. The procyclic forms were harvested at mid-log phase and washed once in Zimmerman postfusion medium (3). Cells (10^7) were resuspended in 500 μl of the same solution and mixed with 30 μg of plasmid DNA. Electroporation was performed by two pulses of 1.5 kV at 25 μF with a Bio-Rad gene pulser. Electroporated trypanosomes were kept on ice for 5 min, inoculated into 4 ml of culture medium, and harvested 18 to 24 h after electroporation. Electroporation of bloodstream forms was performed according to the method of Jefferies et al. (19), except that a single pulse of 1.5 kV at 25 μF and 100 Ω was used. CAT assays were also carried out according to the method of Jefferies et al. (19).

Nuclear extracts. Trypanosomes were lysed and nuclei were prepared according to the method of Murphy et al. (22). The nuclei were washed once in 0.25 M sucrose–5 mM $MgCl_2$–10 mM Tris (pH 7.5), resuspended in the same solution containing 0.5% Nonidet P-40 and a cocktail of protease inhibitors (leupeptin, aprotinin, pepstatin A, and phenylmethylsulfonyl fluoride), and then left at 4°C for 10 min. NaCl was then added to 0.3 M. The suspension was left for 10 min at 4°C and then centrifuged for 60 min at 30,000 × *g*. The supernatant was harvested and rendered in 20% glycerol–0.1% Brij 35–1 mM dithiothreitol. It was then dialyzed overnight against 50 mM NaCl–5 mM $MgCl_2$–0.1 mM EDTA–20% glycerol–1 mM dithiothreitol–20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 7.9) with protease inhibitors. It was centrifuged again for 15 min at 13,000 × *g*, and the supernatant was stored at −80°C. One microgram of protein was routinely used for the bandshift assays.

Bandshift assays and UV cross-linking. Synthetic oligonucleotides (Eurogentec, Liège, Belgium) covering bases −80 to −51 (referred to as box 1) or −50 to −21 (referred to as box 2) of the different promoters, or their mutated versions (see Fig. 1 and 2) were used as probes or competitors. The oligonucleotides were end labelled with T4 polynucleotide kinase, gel purified, and eluted. Probe (0.1 to 0.5 ng; 20,000 to 50,000 cpm) was routinely mixed with nuclear extracts, 200 ng of sheared salmon sperm DNA, and cold competitor DNA if desired (in increasing excesses of 5-, 50-, 250-, and 2,500-fold) in a final volume of 50 μl of binding buffer (5% glycerol, 60 mM KCl, 7.5 mM $MgCl_2$, 1 mM dithiothreitol, 0.5 mM EDTA, 10 mM HEPES, 4 mM Tris [pH 7.9]). The binding reaction was allowed to take place for 15 min at room temperature, and then the mixture was loaded on a 5% polyacrylamide gel and electrophoresed at 150 V in 0.5× Tris-borate-EDTA for 90 min at room temperature. The gels were then vacuum dried and exposed with intensifying screens. For cross-linking, after electrophoresis the wet gels were placed on ice at a distance of 5 cm under a UV lamp and irradiated for 25 min at 254 nm. After overnight exposure, the complexes of interest were cut out, boiled in sample buffer, and subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

RESULTS

Anatomy of the VSG promoter. Previous results demonstrated that a 188-bp region upstream of the transcription start site of the AnTat 1.3A VSG gene expression site of *T. brucei* was able to drive transcription of a reporter gene in both bloodstream and procyclic forms (pD5 and pD5V constructs) (19). The promoter activity was conserved after deletion of the sequence upstream from position −88, whereas it was abolished after deletion at position −57. The addition of 4 kb upstream from the transcription start site did not affect the CAT activity (18). Taken together with other reports (31, 32), these results suggested that the last 88 bp are sufficient for full basal promoter activity. We undertook a detailed analysis of this region.

First, a set of gross deletions was performed (delA to delG; Fig. 1). The resulting constructs were transiently transfected in procyclic trypanosomes, and their ability to drive expression of the CAT gene was monitored. All the deletions reduced the activity of the promoter to less than 10% of that of the wild type, except delG, for which 18% of wild-type activity was observed (Fig. 1). Thus, any gross alteration of the promoter affected its activity, suggesting that its overall configuration is important.

This region was subjected to a more refined analysis by performance of a series of point, double, or triple mutations, concentrating on regions conserved between different VSG promoters (AL, ES, 221, and 118; Fig. 1 [top]). As shown in Fig. 1, three short sequences appeared to be very important. A double mutation in nucleotides −61 and −62 (mutD and mutE) largely inactivated the promoter, while mutations in flanking nucleotides at −59 and −60 (mutG) had less of an effect (29% of activity remaining). The important nucleotide was the T at −61 and not the C at −62, as determined by the analysis of single nucleotide mutants (mutDD and mutEE). Mutation of the triplet from −37 to −35 (mutP) and the doublet at −38 and −37 (mutO) resulted in only 12 and 30% of promoter activity, respectively. In this region, the G at −36 was particularly important (mutPP). Finally, only 6 and 16% of the activity remained when the doublets at positions −1 and +1 and −2 and −1 were mutated (mutY and mutX, respectively). The combination of data obtained with single nucleotide mutants (mutXX and mutZ) revealed that the C and A at −1 and +1, respectively, were the most important. In summary, the nucleotides TAT at −61 to −59 (box 1), CAGG at −38 to −35 (box 2), and CA at −1 to +1 (start site) were found to be crucial for VSG promoter activity. Single mutations at positions −61, −36, and +1 were able to inhibit this activity by 70 to 90%.

To ensure that these important regions of the promoter are the same at the two main developmental stages, we investigated the activity of some constructs in the bloodstream form. Transient activity assays in this form are more difficult than in the procyclic form, because of a 100-fold-lower level of transfection efficiency. The wild-type promoter as well as mutated versions with full (mutU), mildly affected (mutJ and mutM), or strongly reduced (mutD, mutP, and mutY) activity were inserted in a plasmid, derived from pD5V, where the CAT gene is followed by the 3′-untranslated region of the VSG gene to allow high expression in the bloodstream form (19). This gave rise to pD5V constructs mutU′, mutJ′, mutM′, mutD′, mutP′, and mutY′. In accordance with the results for the procyclic forms, mutU, mutJ, and mutM did not abolish promoter activity while mutD, mutP, and mutY strongly reduced the activity (Fig. 1, right).

The effect of small deletions or insertions was also assessed. The promoter was largely inactivated by a 4-bp deletion in the region of box 1 (delH), as expected from the mutational analysis. The observation of a similar effect by a 4-bp deletion between boxes 1 and 2 (delI) was less expected, as the promoter activity was not strongly affected by mutations in this region (mutK). This result suggested that the spacing between the boxes was critical. In order to confirm this hypothesis, we performed two insertions between box 1 and box 2, respectively, of 4 bp (roughly half a helix turn; Fig. 1, insA) and 10 bp (roughly a helix turn; Fig. 1, insB). As shown in Fig. 1, both insertions strongly reduced promoter activity. Note that delG also down-regulated the promoter, suggesting an additional spacing effect between box 2 and the transcription start site.

Finally, we investigated the minimal sequence length required for promoter function by assaying the activity of reporter constructs containing synthetic oligonucleotides of the promoter region. As shown in Fig. 2 (VSG), it was found that a fragment of 75 bp, from −69 to +6, was necessary and sufficient to confer a high activity. Shorter constructs showed reduced or no activity (data not shown), although we did not evaluate the effect of deletions between +2 and +6. In agreement with the results obtained before, a six-nucleotide sequence alteration in the region encompassing box 2 (Fig. 2, VSG1) strongly reduced the promoter activity, while a change

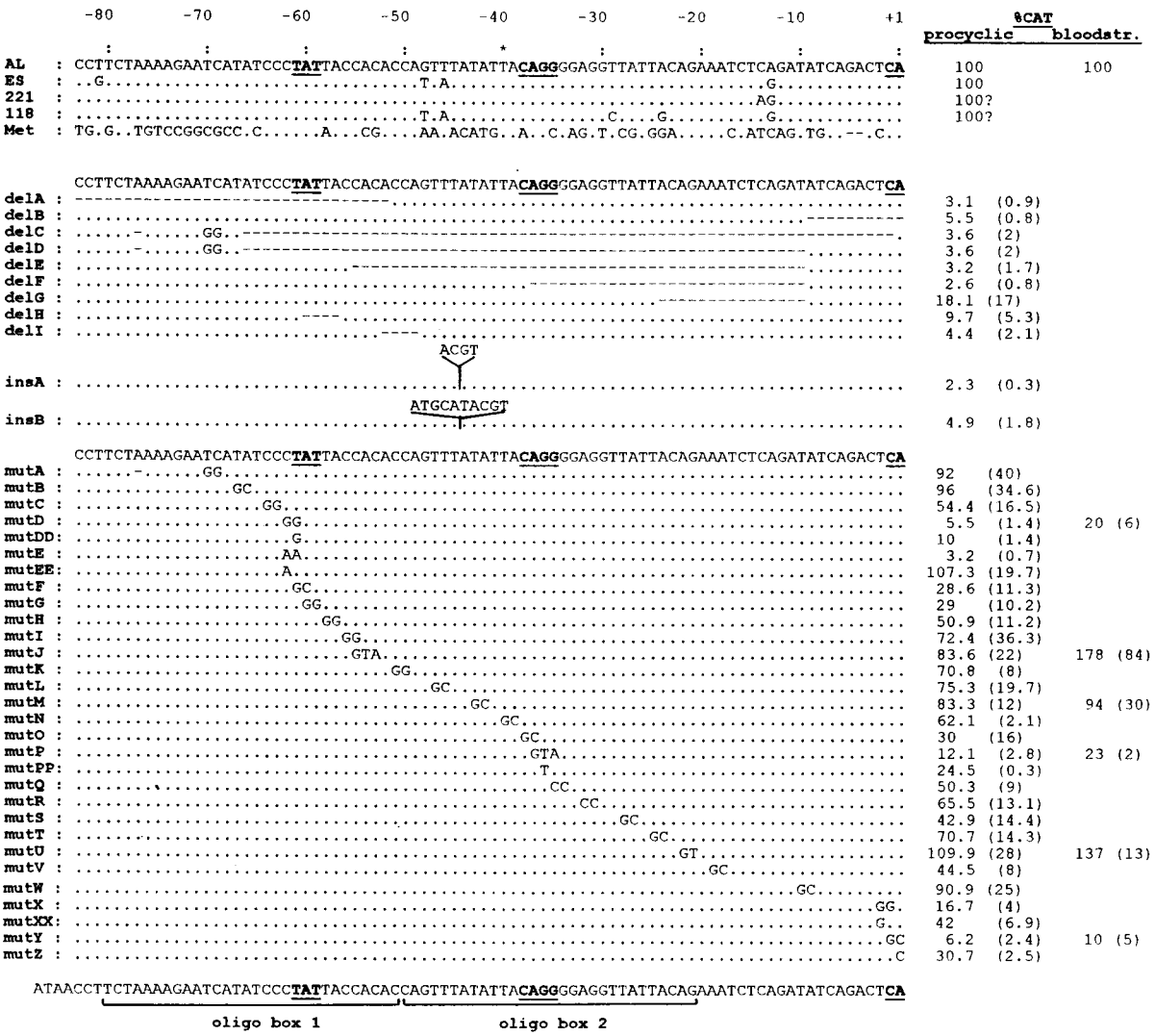

FIG. 1. Mutagenesis of the VSG promoter. At the top, the sequences of the core regions of four promoters from different bloodstream VSG expression sites are compared, with single dots standing for identical nucleotides (AL and ES [23]; 221 [32]; 118 [12]). In addition, the sequence of a metacyclic (Met) VSG promoter (AnTat 11.17 [30]) is aligned for maximal homology, with dashes for spacing. The sequence of AL was altered by site-directed mutagenesis in the regions showing sequence conservation between bloodstream VSG promoters (del, deletion; ins, insertion; mut, mutagenesis). Transient CAT activity assays were performed in procyclic forms transfected with pD5 constructs (19) containing the different versions of the promoter. Some mutants were constructed in pD5V (19) and tested in bloodstream forms. The results are expressed in percentages relative to the activity of AL. The numbers in parentheses represent the standard deviations measured from at least three independent assays. The crucial elements are underlined and in boldface type. The asterisk at −40 refers to a putative alternative start site (32). The extent of the oligonucleotide probes used for the protein binding experiments is shown at the bottom.

of the same extent next to box 2 (Fig. 2, VSG2) did not affect the activity.

In conclusion, the results shown in Fig. 1 and 2 indicated that the 70-bp region upstream from the start site contains the basal promoter of the VSG transcription unit. They show that inside this region, three short stretches are crucial, and they suggest that the spacing between these critical sequences is important.

Conservation of the key elements in a metacyclic VSG promoter. The expression site for the metacyclic AnTat 11.17 VSG gene of *Trypanosoma gambiense* was recently characterized in our laboratory (30). The transcription promoter was found to be located 1.9 kb upstream from the VSG gene, and the sequence of its core region is shown in Fig. 1 (Met). This sequence is identical to that of the promoter of an unusual monocistronic VSG transcription unit active in the bloodstream form (1). While the overall homology with typical bloodstream VSG promoters is low (45% identity), it is interesting that the critical elements, box 1, the central AG dinucleotide of box 2, and the start region, are conserved. In procyclic forms, a synthetic oligonucleotide covering the −69 to +6 region of this promoter was at least twice as active as the corresponding region of the bloodstream VSG promoter (Fig. 2, MET). Moreover, a 6-nucleotide alteration in the region corresponding to box 2 of the bloodstream VSG promoter also inactivated this promoter (Fig. 2, MET1).

Interchangeability of the box 2 regions from the VSG and ribosomal promoters. Although their sequences differ, the core regions of the procyclin and ribosomal promoters of *T. brucei* possess a structure resembling that of the VSG promoter, with two important boxes located at similar positions with respect to the start site (Fig. 2, boxed regions of VSG, PRO, and RIB). However, the level of activity of the −69 to

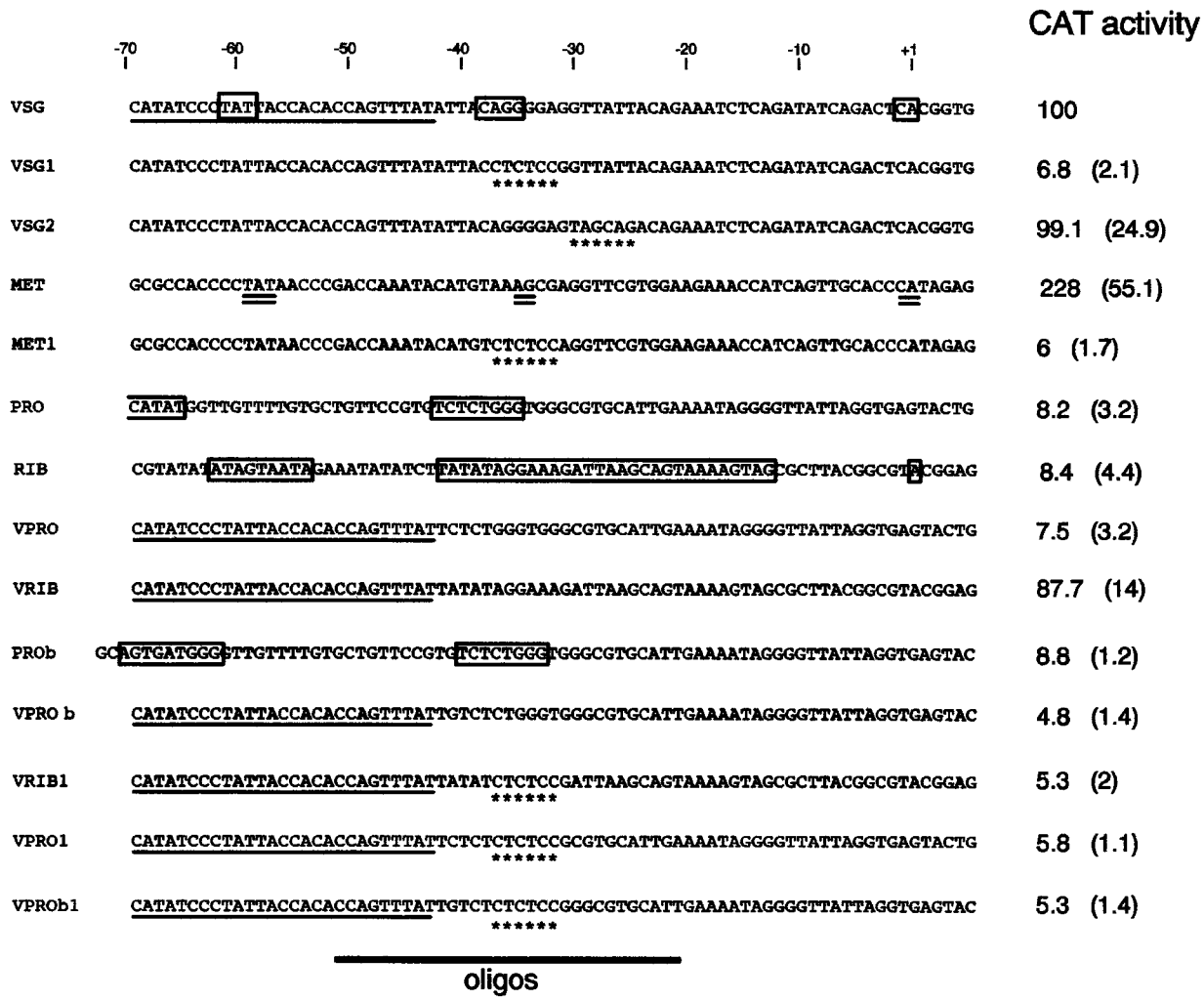

FIG. 2. Activity of the −69 to +6 region of the VSG, procyclin, ribosomal, and hybrid promoters. Synthetic oligonucleotides of both strands of this region were annealed and cloned between the *Sph*I and *Stu*I sites of pD5 (19). The CAT activity of these constructs was measured in procyclic forms and expressed as the percentage of activity of the VSG promoter (100% = 16 to 41% of conversion, depending on the experiments). The bloodstream VSG, metacyclic VSG, procyclin, and ribosomal promoters are referred to as VSG, MET, PRO, and RIB, respectively. The sequences are aligned on the transcription start site (VSG [23]; RIB and PRO [16]). The boxed regions have been defined as important for promoter activity (PRO [5, 28]; RIB [16]; VSG [Fig. 1 in this work]). The VSG1 and VSG2 sequences are mutated in box 2 and just next to box 2, respectively (asterisks). VPRO and VRIB contain the VSG promoter sequence from −69 to −43 (underlined) inserted in front of the −42 to +6 region of the procyclin and ribosomal promoters, respectively. VPRO1 and VRIB1 are mutated in the region corresponding to box 2 in the VSG promoter (asterisks). PROb is an alternative oligonucleotide representing the core region of the procyclin promoter with a different start site and the complete box 1 region of the procyclic acidic repetitive protein A locus. VPROb and VPROb1 are derived from PROb. The standard deviations of three independent experiments are indicated (in parentheses). Shown at the bottom are the extents of the box 2 oligonucleotide probes from the different promoters and their mutant versions, which were used for the protein binding experiments. In most cases, the noncoding strand—thus, the antisense of this sequence—was used.

+6 region of the VSG promoter was found to be 12-fold higher than that of the same region of the two others (Fig. 2, compare VSG, PRO, and RIB). In order to evaluate if the key elements of these promoters are related, synthetic oligonucleotides of the −69 to +6 region containing box 1 of the VSG promoter and different box 2 regions were placed upstream of the CAT gene. As shown in Fig. 2, the hybrid VSG-procyclin (VPRO) promoter was only weakly active (7.5% of the activity of VSG), while the VSG-ribosomal (VRIB) promoter was nearly as active as the wild-type VSG promoter (87.7% of the activity of VSG). In order to check if the low level of activity of the PRO and VPRO sequences was due to an improper choice of the transcription start site, the mapping of which is presently equivocal (5, 23, 28), two additional sequences, termed PROb and VPROb, were tested. These sequences differed not only in the choice of another start site but also in the facts that PROb contains the full box I of the procyclic acidic repetitive protein A promoter (5) and VPROb contains a different spacing between VSG box 1 and the rest of the promoter. As shown in Fig. 2, these sequences were as poorly active as the PRO and VPRO constructs. Significantly, the activity of the hybrid VRIB promoter was strongly reduced if this sequence was mutated in the region corresponding to that of box 2 of the VSG promoter (94% inhibition in VRIB1; 5 versus 87.7%). A similar experiment with VPRO and VPROb led to only a slight reduction of activity (Fig. 2, VPRO1 and VPROb1). Altogether, these data suggested that the box 2 regions in the VSG and RIB promoters have similar functions, while no definite conclusion could be drawn regarding the procyclin promoter.

Specific nuclear factors bind to the noncoding strand of box 2. We investigated the existence of nuclear factors able to bind to the VSG promoter. Different promoter probes (30-mer oli-

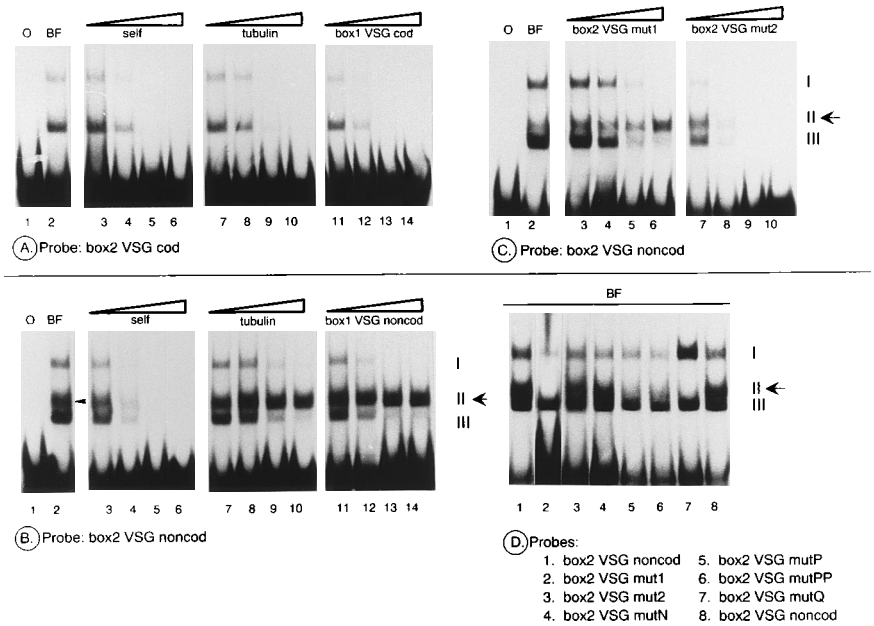

FIG. 3. Specific protein binding to the noncoding strand of the box 2 region of the VSG promoter. Bandshift assays were carried out with the indicated oligonucleotide probes (30-mers; Fig. 2). Binding was inhibited by increasing excesses (5-, 50-, 250-, and 2,500-fold) of the indicated cold oligonucleotides. All assays were performed with nuclear extracts from bloodstream forms. Retarded complexes are numbered with roman numbers. The specific complex (complex II) is indicated (arrow). The sequences of VSG mut1 and VSG mut2 are shown in Fig. 2, while those of mutN, mutP, mutPP, and mutQ are shown in Fig. 1. O, probe alone; BF, nuclear extracts from bloodstream forms without specific competitor.

gonucleotides; Fig. 1 and 2 [bottom]) were subjected to electrophoretic mobility shift assays after incubation with nuclear extracts from bloodstream and procyclic forms.

Several experiments were conducted with various double-stranded fragments from the promoter region. Only nonspecific complexes were observed, as judged by the disappearance of these complexes following the addition of excess amounts of irrelevant DNA, such as tubulin gene fragments (these results are not shown but resemble those presented in Fig. 3A).

The same experiments were performed with single-stranded oligonucleotides. As shown in Fig. 3, these probes formed both nonspecific and specific complexes with nuclear extracts from bloodstream forms. Single-stranded probes from the box 1 region and from the coding strand of box 2 showed only nonspecific binding, which was inhibited by an excess of either the same or irrelevant DNA (shown for the coding strand of box 2 in Fig. 3A). In contrast, a specific complex was obtained with the noncoding strand of box 2, as revealed by its resistance to competition by different irrelevant DNAs, such as tubulin DNA or the noncoding strand of box 1 (Fig. 3B, arrow). This complex was also resistant to competition by double-stranded DNA of the box 2 region (data not shown). It was clear that the factors responsible for the appearance of this band were proteins, since the band was not observed following incubation with proteinase K. In addition, the binding activity did not contain a detectable RNA component, since incubation with RNase A did not affect the results (data not shown).

In order to determine the functional significance of protein binding to the noncoding strand of box 2, competition experiments were performed with mutated versions of the noncoding strand of the box 2 region. Two mutant sequences, mut1 and mut2, were assayed. The mut1 sequence differs in 6 nucleotides centered on box 2, while mut2 contains 6 nucleotide changes in a region next to box 2 (Fig. 2). As mentioned above, the promoter activity was lost in mut1 (Fig. 2, VSG1) but conserved in mut2 (Fig. 2, VSG2). Accordingly, the specific complex with the noncoding strand of box 2 was not inhibited by mut1, while mut2 was able to compete (Fig. 3C). These results strongly suggested that formation of this complex is due to the binding of factors to box 2. A more direct demonstration was provided by the incubation of nuclear extracts with the mut1 and mut2 probes. As shown in Fig. 3D (lanes 1 to 3), the specific band (arrow) was observed with mut2 but not with mut1. This analysis was refined by the evaluation of protein binding to sequences containing single-, double-, and triple-nucleotide mutations centered on box 2. The results shown in Fig. 3D (lanes 4 to 7) indicated that mutP, mutPP, and mutQ (88, 75, and 50% inhibition of promoter activity, respectively) strongly and specifically affected the assembling of complex II, whereas mutN (38% inhibition of promoter activity) was less inhibitory. Competition experiments using these mutant sequences to challenge the binding on the wild-type box 2 VSG probe fully confirmed these conclusions (data not shown). Taken together, these data demonstrated that a specific protein complex with the noncoding strand of box 2 is involved in the activity of the VSG promoter.

The metacyclic and bloodstream VSG promoters bind the same proteins. The binding of proteins to the box 2 region of the noncoding strand of the metacyclic VSG promoter was found to be similar to that of the bloodstream VSG promoter (Fig. 4, lane 1). This binding was challenged by the addition of increasing amounts of oligonucleotides specific to the same region of the bloodstream VSG promoter, as well as with control oligonucleotides. The results clearly indicated that the same proteins were involved in the building of a specific complex (Fig. 4, complex II, arrows). The complex was inhibited equally by sequences from the metacyclic and bloodstream VSG promoters (lanes 2 to 4 and 8 to 10, respectively), except if box 2 was mutagenized (lanes 11 to 13). The reverse experiment fully confirmed these results: complex II on box 2 of the bloodstream VSG promoter was inhibited by the corresponding region of the metacyclic VSG promoter (data not shown).

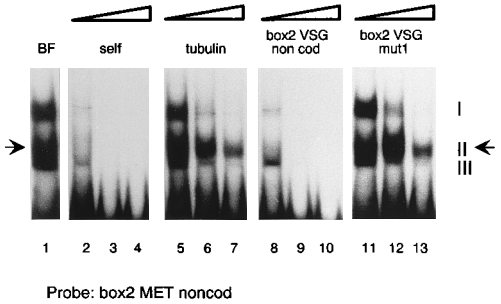

FIG. 4. The box 2 region of the bloodstream VSG promoter competes for protein binding to the metacyclic VSG promoter. Assays were carried out as described in the legend to Fig. 3, except that the competitions were performed with 50-, 250-, and 2,500-fold excesses of cold oligonucleotides. See the legend to Fig. 2 for the extents of the oligonucleotide competitors used. BF is as described in the legend to Fig. 3.

Common proteins bind to box 2 of the VSG and ribosomal promoters. As shown in Fig. 5 (lanes 1 to 3), a similar pattern of complexes, including that found to be specific in bloodstream forms (arrows), was observed with the box 2 probe of the VSG promoter and extracts from either bloodstream (BF) or procyclic (PF) nuclei. This observation suggested that the proteins that bind to box 2 of the VSG promoter are present in both the bloodstream and procyclic forms.

In order to determine if these proteins may bind to both the VSG and procyclin promoters, competition experiments were conducted on the box 2 complex using excess DNA from the region of the procyclin promoter corresponding to box 2 of the VSG promoter. The data presented in Fig. 5 (lanes 4 to 9) indicated that both strands of that region of the procyclin promoter were able to compete. These results were not observed with other probes from the procyclin promoter (data not shown) and were reproduced in several independent experiments. They suggest that both strands of the box 2-related region of the procyclin promoter bind specific proteins, in accordance with recent results obtained with procyclic nuclear extracts (6). In order to determine if these proteins bind to the region corresponding to box 2 of the VSG promoter, competition experiments were conducted with sequences mutated in this region, using two variant constructs with alternative start sites (VPRO1 and VPROb1 [Fig. 2]). The results were negative in both cases (data not shown). In conclusion, although these experiments revealed that at least some of the proteins binding to the VSG box 2 also bind to the procyclin promoter,

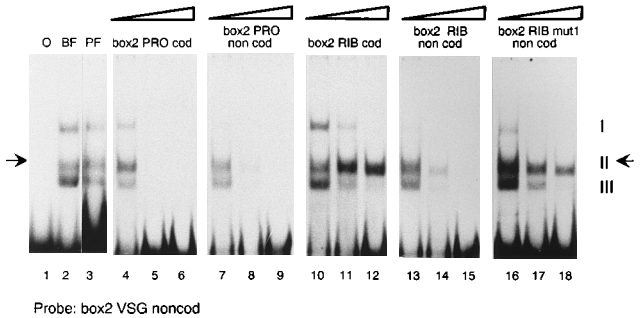

FIG. 5. The region corresponding to the VSG box 2 in the procyclin and ribosomal promoters competes for protein binding to the VSG promoter. Assays were carried out as described in the legend to Fig. 4, except that nuclear extracts from procyclic forms (PF) were used in lane 3. BF is as described in the legend to Fig. 3.

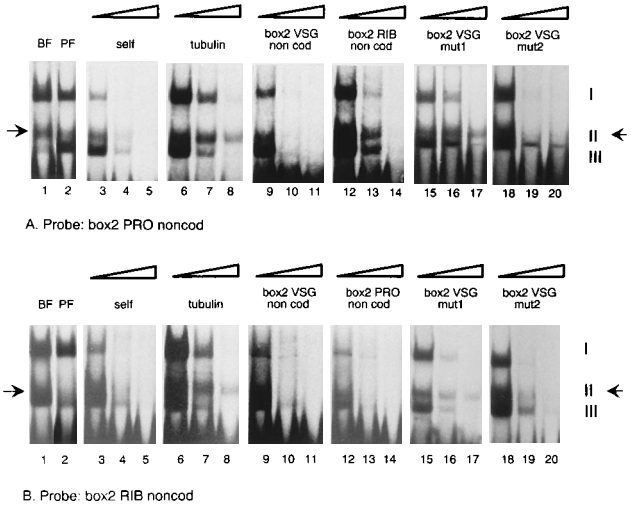

FIG. 6. Binding of similar proteins to the noncoding strand of box 2 from the VSG, procyclin, and ribosomal promoters. Assays were carried out as described in the legend to Fig. 4. BF and PF are as described in the legends to Fig. 3 and 5, respectively.

both competition of protein binding and promoter activity assays (Fig. 2) indicated that this binding does not occur in the region corresponding to that of box 2 of the VSG promoter, at least according to the two sequence alignments that we tested.

A similar approach was used to compare the VSG and ribosomal promoters. The results presented in Fig. 5 (lanes 10 to 15) indicated that only the noncoding strand of the ribosomal promoter sequence corresponding to the box 2 region of the VSG promoter was able to compete with that region. Furthermore, if mutated in this region (VRIB1 [Fig. 2]), the ribosomal promoter was rendered unable to compete (Fig. 5, lanes 16 to 18). This result was in accordance with those presented in Fig. 2, which indicated that the mutations in VRIB1 strongly affect promoter activity. In conclusion, both competition for protein binding and promoter activity assays indicated that similar proteins bind to a functionally important element (box 2) located in the same region of the VSG and ribosomal promoters.

Similar analyses were performed by using the procyclin and ribosomal promoters as probes. Both strands of the box 2-related region from the procyclin promoter were incubated with nuclear extracts from bloodstream forms. A specific complex was clearly observed with the noncoding strand (Fig. 6A, lanes 1 to 8, arrow). As expected, this complex was inhibited by an excess of DNA from the box 2 region of the VSG promoter (Fig. 6A, lanes 9 to 11). The results obtained with the coding strand (not shown) indicated that this sequence is able to form two specific complexes. A minor complex showed the same mobility as that observed with the noncoding strand, whereas a major complex appeared to be specific to the coding strand. Only the minor complex was inhibited by an excess of DNA from the VSG promoter. The box 2 region of the ribosomal promoter was also probed. It formed a specific complex with nuclear extracts from either the bloodstream or procyclic forms, although a long exposure of the autoradiograms was necessary to visualize this complex (Fig. 6B, lanes 1 to 8, arrow). This complex was inhibited by the noncoding strand of the box 2 region of the VSG and procyclin promoters (Fig. 6B, lanes 9 to 11 and 12 to 14, respectively).

The significance of the competition by VSG promoter DNA for the binding of proteins to the procyclin and ribosomal promoters was evaluated by assaying mutated versions of the

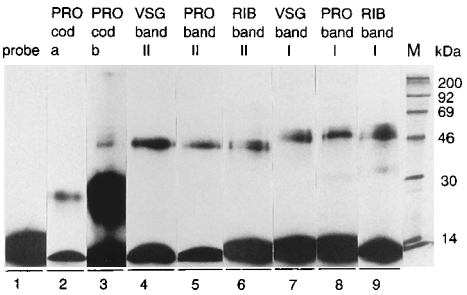

FIG. 7. A 40-kDa DNA-binding protein is present in specific complexes with the VSG, procyclin, and ribosomal promoters. The indicated UV cross-linked complexes (bands I and II in Fig. 5 and 6) were subjected to SDS-PAGE. PRO cod (lanes 2 and 3) refers to a similar analysis of the two (very close and technically not separable) specific complexes formed on the coding strand of box 2 from the procyclin promoter (a, short exposure; b, long exposure). Lane 1: probe alone; lane 10: molecular mass markers.

VSG promoter. The mutants 1 and 2 described above (Fig. 2, VSG1 and VSG2) were used for this purpose. The results shown in Fig. 6A and B (lanes 15 to 20 of each panel) indicated that the VSG promoter competes efficiently with the procyclin and ribosomal promoters unless box 2 is altered: competition clearly occurred at a lower level with mut1 than with mut2 (competition of complex II in lanes 15 to 20 should be compared both with the positive and negative controls in the same experiments [lanes 3 to 5 and 6 to 8, respectively] and with the results obtained with complexes I and III in the same lanes).

In conclusion, the results presented in Fig. 5 and 6 strongly suggest that the same proteins are involved in the formation of complexes with the box 2 region of the VSG and ribosomal promoters. The same conclusion applies to the procyclin promoter, but in this case a specific binding to both DNA strands was observed, and the location of the box 2-related element could not be determined.

In order to visualize directly the proteins binding to the three promoters, the DNA-protein complexes were cross-linked by UV irradiation and then subjected separately to electrophoresis. As shown in Fig. 7, different proteins were present in the different complexes. A major 30-kDa protein and a minor 40-kDa protein appeared to bind to the coding strand of box 2 of the procyclin promoter (Fig. 7, lanes 2 and 3). A 40-kDa protein was present in the specific complexes with the noncoding strand of box 2 from the VSG, procyclin, and ribosomal promoters (Fig. 7, lanes 4 to 6), while a 50-kDa protein was present in the nonspecific complexes (lanes 7 to 9). Thus, at least a 40-kDa single-stranded DNA-binding protein was associated with the functionally similar region of the noncoding strand of the three promoters. The additional binding of this protein to the coding strand of the procyclin promoter (Fig. 7, lane 3) may explain why the formation of complexes with this sequence was inhibited by the noncoding strand of box 2 from the VSG promoter (a summarizing scheme is shown in Fig. 8).

DISCUSSION

Virtually nothing concerning transcription initiation in protozoa is known. This is particularly true for trypanosomatids, the organization of whose genome into long polycistronic transcription units has hindered the mapping of transcription start regions (10, 24). Among the few promoters identified to date in these organisms are those for the genes of the two major surface stage-specific antigens of *T. brucei*, the VSG of the bloodstream form, and procyclin of the procyclic form. Although the expression of the transcription units driven by these respective promoters can be taken as an example for mutually exclusive developmental regulation, transient activity assays in episomal reporter constructs as well as run-on transcription assays and analysis of transcripts from the 5′ region of the units all indicated that the VSG and procyclin promoters are active at both stages of the parasite life cycle (19, 23, 31, 32). This paradox can be explained by the existence of major controls operating at the level of transcription elongation, thus, downstream from transcription initiation (23, 29). Therefore, the analysis of the activity of the VSG promoter, as well as the preliminary characterization of specific protein complexes with the VSG promoter, was possible in procyclic forms. This is worth noting, as the efficiency of transient activity assays in the bloodstream form is 100-fold lower than in the procyclic form, precluding detailed studies of the promoter at this developmental stage. We present here such an analysis. This should be compared with similar analyses of the procyclin promoter, already presented by different groups (5, 6, 16, 17, 28).

Previous studies have shown that the 88-bp region immediately upstream of the transcription start site can be considered as the basal element of the VSG promoter, with no evidence for any modulating region in the close upstream environment (19, 31, 32). Our results show that the core element is restricted to 75 bp, as a synthetic oligonucleotide of this size is necessary and sufficient for promoter activity. Within this sequence, two small boxes and the start site region are essential. In addition, the spacing between the crucial sequences appears to be very important. These results can be compared with those obtained with the procyclin and ribosomal promoters (5, 16). Although the sequences of these promoters are different, it is clear that the structural organizations of their core regions are similar. This conclusion is strongly supported by the fact that hybrid constructs containing the first half of the VSG core promoter and the second half of the ribosomal promoter are active. This confirms similar observations with hybrid constructs between the procyclin and ribosomal promoters (16) and strongly suggests that elements of these three promoters are functionally interchangeable. However, it should be stressed that upstream elements are required for full activity of the procyclin and ribosomal promoters, which is not the case for the VSG promoter.

The VSG promoter does not share consensus sequence elements with any known promoter. Our results indicate that the few elements shared with the procyclin and ribosomal promoters are not critical for promoter function. The only distinctive feature is the presence of three direct TATTAC repeats, but these repeats do not appear to be essential. We note that the very short sequence elements identified as vital for promoter

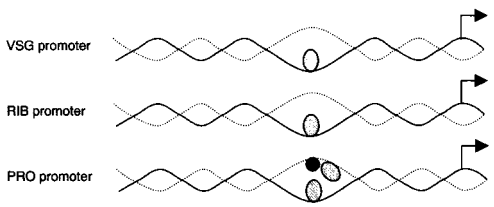

FIG. 8. Summary of the DNA binding assays. A common 40-kDa protein (hatched ellipse) binds to the noncoding strand of a crucial discrete element present approximately 35 bp upstream from the transcription start site (arrow) in the VSG and ribosomal (RIB) promoters, as well as to a close but unidentified region of the procyclin (PRO) promoter. In the latter case, the coding strand binds another specific protein (black dot), although some binding of the 40-kDa protein also occurs. The different elements of this conjectural scheme are not drawn to scale.

activity, TAT (box 1, positions −61 to −59) and CAGG (box 2, positions −38 to −35), are located close to the respective consensus positions of the CAAT and TATA boxes of eukaryotic promoters (−99 to −60 and −35 to −16, respectively) (8).

Attention should be drawn to the fact that the conclusions of this study are derived from transient assays with trypanosomes in the exponential phase of growth. Whether other transcriptional or promoter features dependent on the chromosome location or stage of the cell or parasitic cycle are needed is not known. Regarding this possibility, we have tested some critical and noncritical constructs in both bloodstream and procyclic forms, showing that they behave similarly. However, in this system the core promoter of a metacyclic VSG unit was active in both bloodstream and procyclic forms, while in vivo it is probably only active at the metacyclic stage (13, 21). Therefore, the possibility of a differential activity of these promoters in situ cannot be excluded. To date, several lines of evidence suggest that the transcription of the VSG gene expression sites may be modulated by nondiffusible, thus *cis*-acting, elements, probably linked to the location and/or composition of the chromosomes (4, 24, 29).

All our attempts to observe reproducibly a specific binding of proteins to double-stranded DNA fragments of the VSG promoter were unsuccessful. In this respect, our results contradict those of Janz et al. (17) with the procyclin promoter, perhaps because of the use of different experimental conditions. In striking contrast, a specific binding was clearly observed when nuclear extracts were incubated with a single-stranded probe, namely, the noncoding strand of box 2. These results confirm and extend previous observations with the procyclin promoter by another group (6). The significance of the interaction of promoter elements with single-stranded DNA-binding factors is presently unclear (see reference 6 and references therein). The cloning and characterization of these factors will probably help to resolve this issue.

The specific DNA-binding proteins, in particular, a 40-kDa protein, are present at both stages of the parasite life cycle. In addition, these proteins appear to also bind to the noncoding strand of the box 2 region of the procyclin and ribosomal promoters, despite the sequence differences among the three promoters. Significantly, when mutated in the region corresponding to box 2 of the VSG promoter, the ribosomal promoter was inactivated. Therefore, box 2 of the VSG and ribosomal promoters appears to bind common proteins which are important for activity. These observations indicate that despite their overall sequence difference, these promoters are related. In this respect, it is worth noting that within the limits of box 2, as defined in the two promoters by both activity assays and specific protein binding, the nucleotide sequences are very similar (TACAGG versus TATAGG; thus, 5′-CCT[G/A]TA-3′ in the noncoding strand). Presumably this sequence is involved in the binding of common factors. If it is shared between the promoters for such different genes, the DNA-binding proteins identified probably have no specific regulatory function. This conclusion is in keeping with the evidence that these promoters are constitutively active in situ (23). Our results also indicate that the procyclin promoters bind the same factors, since competition for protein binding on the procyclin, VSG, and ribosomal promoters was found to occur at the same region. However, in the case of the procyclin promoter, several observations complicated the interpretation of the results. (i) In agreement with a previous report (6), we found that both strands of this promoter bind specific proteins, which is clearly different from the observations with the VSG and ribosomal promoters. (ii) On the basis of two different sequence alignments between the procyclin and VSG promoters, we were unable to map box 2 of the procyclin promoter, since procyclin promoter sequences mutated in the region corresponding to box 2 of the VSG promoter were both still active and able to compete for the formation of a specific complex. (iii) In contrast to the case of the ribosomal promoter, hybrid constructs between the VSG box 1 and the rest of the procyclin promoter (VPRO and VPROb) were only weakly active, suggesting improper spacing between the different elements. A likely explanation for points ii and iii is that box 2 of the procyclin promoter is not located at the same place as those of the VSG and ribosomal promoters in our sequence alignments, and thus, this element remains to be identified. On the basis of the discussion above, it may be relevant to note that in the procyclin promoter the sequence AATAGG is present approximately 20 nucleotides downstream from the putative box 2 region. However, this sequence falls outside of the region that we probed for protein binding, and in addition, it was not defined as important in previous studies (5, 28).

Our results suggest that some elements of the VSG promoter must be single stranded in vivo. It is tempting to speculate that this constraint is involved in the control of the activity of the different VSG promoters in the bloodstream form. In keeping with this hypothesis, it was observed that, in situ, the active VSG expression site is selectively sensitive to single-strand-specific endonuclease digestion (14). Speculation regarding the possibility that local DNA denaturation is dependent on telomeric chromatin interaction with the nuclear or nucleolar matrix has already been presented (14).

ACKNOWLEDGMENTS

We thank D. Nolan and D. Perez-Morga for comments on the manuscript.

This work was supported by the Belgian FRSM and FRC-IM, by a research contract with the Communauté Française de Belgique (ARC), and by the Agreement for Collaborative Research between ILRAD (Nairobi) and Belgian Research Centres. L.V. is a research associate of the Belgian National Fund for Scientific Research.

REFERENCES

1. **Alarcon, C. M., H. Jin Son, T. Hall, and J. E. Donelson.** 1994. A monocistronic transcript for a trypanosome variant surface glycoprotein. Mol. Cell. Biol. **14:**5579–5591.
2. **Baltz, T., D. Baltz, C. Giroud, and J. Crokett.** 1985. Cultivation in a semi-defined medium of animal infective forms of *Trypanosoma brucei*, *T. equiperdum*, *T. evansi*, *T. rhodesiense*, and *T. gambiense*. EMBO J. **4:**1273–1277.
3. **Bellofatto, V., and G. A. M. Cross.** 1989. Expression of a bacterial gene in a trypanosomatid protozoan. Science **244:**1167–1169.
4. **Borst, P., J. H. Gommers-Ampt, M. J. L. Ligtenberg, G. Rudenko, R. Kieft, M. C. Taylor, P. A. Blundell, and F. Van Leeuwen.** 1993. Control of antigenic variation in African trypanosomes. Cold Spring Harbor Symp. Quant. Biol. **58:**105–114.
5. **Brown, S. D., J. Huang, and L. H. T. Van der Ploeg.** 1992. The promoter for the procyclic acidic repetitive protein (PARP) genes of *Trypanosoma brucei* shares features with RNA polymerase I promoters. Mol. Cell. Biol. **12:**2644–2652.
6. **Brown, S. D., and L. H. T. Van der Ploeg.** 1994. Single-stranded DNA-protein binding in the procyclic acidic repetitive protein (PARP) promoter of *Trypanosoma brucei*. Mol. Biochem. Parasitol. **56:**109–122.
7. **Brun, R., and M. Schoenenberger.** 1979. Cultivation and *in vitro* cloning of procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Acta Trop. **36:**289–292.
8. **Bucher, P., and E. N. Trifonov.** 1986. Compilation and analysis of eukaryotic pol II promoter sequences. Nucleic Acids Res. **14:**10009–10026.
9. **Chung, H. M., M. G. S. Lee, and L. H. T. Van der Ploeg.** 1992. RNA polymerase I-mediated protein-coding gene expression in *Trypanosoma brucei*. Parasitol. Today **8:**414–418.
10. **Clayton, C.** 1992. Developmental regulation of gene expression in *Trypanosoma brucei*. Prog. Nucleic Acid Res. Mol. Biol. **43:**37–66.
11. **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.
12. **Gottesdiener, K., H. M. Chung, S. D. Brown, M. G. S. Lee, and L. H. T. Van**

der Ploeg. 1991. Characterization of VSG gene expression site promoters and promoter-associated DNA rearrangement events. Mol. Cell. Biol. **11:** 2467–2480.
13. **Graham, S. V., K. R. Matthews, P. G. Shiels, and J. D. Barry.** 1990. Distinct developmental stage-specific activation mechanisms of trypanosome VSG genes. Parasitology **101:**361–367.
14. **Greaves, D. R., and P. Borst.** 1987. *Trypanosoma brucei* variant-specific glycoprotein gene chromatin is sensitive to single-strand specific endonuclease digestion. J. Mol. Biol. **197:**471–483.
15. **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.
16. **Janz, L., and C. Clayton.** 1994. The PARP and rRNA promoters of *Trypanosoma brucei* are composed of dissimilar sequence elements that are functionally interchangeable. Mol. Cell. Biol. **14:**5804–5811.
17. **Janz, L., M. Hug, and C. Clayton.** 1994. Factors that bind to RNA polymerase I promoter sequences of *Trypanosoma brucei*. Mol. Biochem. Parasitol. **65:**99–108.
18. **Jefferies, D., and E. Pays.** Unpublished data.
19. **Jefferies, D., P. Tebabi, and E. Pays.** 1991. Transient activity assays of the *Trypanosoma brucei* variant surface glycoprotein gene promoter: control of gene expression at the posttranscriptional level. Mol. Cell. Biol. **11:**338–343.
20. **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.
21. **Lenardo, M. J., K. M. Esser, A. M. Moon, L. H. T. Van der Ploeg, and J. E. Donelson.** 1986. Metacyclic variant surface glycoprotein genes of *Trypanosoma brucei* subsp. *rhodesiense* are activated in situ, and their expression is transcriptionally regulated. Mol. Cell. Biol. **6:**1991–1997.
22. **Murphy, N. B., A. Pays, P. Tebabi, H. Coquelet, M. Guyaux, M. Steinert, and E. Pays.** 1987. A *Trypanosoma brucei* repeated element with unusual structural and transcriptional properties. J. Mol. Biol. **195:**855–871.
23. **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.
24. **Pays, E., L. Vanhamme, and M. Berberof.** 1994. Genetic controls for the expression of surface antigens in African trypanosomes. Annu. Rev. Microbiol. **48:**25–52.
25. **Roditi, I., and T. W. Pearson.** 1990. The procyclin coat of African trypanosomes. Parasitol. Today **6:**79–82.
26. **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 in *Trypanosoma brucei*. Mol. Cell. Biol. **10:**3492–3504.
27. **Rudenko, G., M. G. S. Lee, and L. H. T. Van der Ploeg.** 1991. The PARP and VSG genes of *Trypanosoma brucei* do not resemble RNA polymerase II transcription units in sensitivity to the addition of Sarkosyl in nuclear run-on assays. Nucleic Acids Res. **20:**303–306.
28. **Sherman, D. R., L. Janz, M. Hug, and C. Clayton.** 1991. Anatomy of the PARP promoter of *Trypanosoma brucei*. EMBO J. **10:**3379–3386.
29. **Vanhamme, L., M. Berberof, D. Le Ray, and E. Pays.** 1995. Stimuli of differentiation regulate RNA elongation in the transcription units for the major stage-specific antigens of *Trypanosoma brucei*. Nucleic Acids Res. **23:**1862–1869.
30. **Vanhamme, L., C. Do Thi, V. Olislagers, P. Tebabi, A. Pays, and E. Pays.** Unpublished data.
31. **Zomerdijk, J. C. B. M., R. Kieft, P. G. Shiels, and P. Borst.** 1991. Alpha-amanitin resistant transcription units in trypanosomes: a comparison of promoter sequences for a VSG gene expression site and for the ribosomal RNA genes. Nucleic Acids Res. **19:**5153–5158.
32. **Zomerdijk, J. C. B. M., M. Ouellette, A. L. M. A. ten Asbroek, R. Kieft, A. M. M. Bommer, C. E. Clayton, and P. Borst.** 1990. The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. EMBO J. **9:**2791–2801.