Nucleic Acids Research, 1998, Vol. 26, No. 8 **1985–1990**

A trypanosome metacyclic VSG gene promoter with two functionally distinct, life cycle stage-specific activities

Sheila V. Graham[+], Ben Wymer and J. David Barry*

Wellcome Unit of Molecular Parasitology, University of Glasgow, Anderson College, 56 Dumbarton Road, Glasgow G11 6NU, UK

Received November 21, 1997; Revised and Accepted March 4, 1998

ABSTRACT

In the mammalian bloodstream, African trypanosomes express the variant surface glycoprotein (VSG), continual switching of which allows evasion of the host immune response. Bloodstream VSG genes are transcribed from polycistronic bloodstream expression sites with promoters which are located 45–60 kb upstream. These promoters are not exclusively stage-regulated, being active in the insect midgut stage where VSG is not expressed. However, the metacyclic VSG (M-VSG) genes, a small subset activated when VSG synthesis begins in the metacyclic stage in the tsetse fly salivary glands, are transcriptionally activated specifically in that stage from promoters <3 kb upstream. Using deletion mapping and transient transfection, we show that the entire 1.22 M-VSG gene promoter region (171 bp) is required for full activity in metacyclic-derived trypanosomes. However, a subsidiary, bloodstream stage-specific activity is present in its 5′ half which directs transcription initiation very close to the initiation site used in metacyclic-derived trypanosomes. Our results imply that the M-VSG gene promoter is longer and more complex than other VSG gene promoters.

INTRODUCTION

Trypanosoma brucei is a protozoan parasite which causes serious disease in man and domestic livestock. It has two life cycle phases, one in mammalian hosts and the other in the tsetse fly vector (1). Trypanosomes evade the mammalian immune response by antigenic variation, the continual switching of the variant surface glycoprotein (VSG) which constitutes their surface coat (2). Although there are ~1000 VSG genes (3), in the bloodstream only one is expressed at a time. This is achieved by insertion of a VSG gene copy into one of up to 20 telomeric bloodstream expression sites. These sites are complex, polycistronic transcription units in which the VSG gene is co-transcribed with several expression site associated genes (ESAGs) (4) from the promoter some 45–60 kb upstream (5–9).

Regulation of gene expression in the order Kinetoplastida is unusual among eukaryotes in that it occurs largely post-transcriptionally (10,11). As most genes are organised in polycistronic transcription units and are co-transcribed from a common 5′ promoter, their differential expression requires regulation after transcription, and there is evidence for this occurring at most steps in the pathway between transcription and translation (11). For *T.brucei*, the only identified promoters for protein coding genes are those for the VSG (7,9,12–14) and procyclin/PARP genes (15–17). These promoters, which appear to direct transcription by RNA polymerase I (17–19), are constitutively active although their activity is modulated in different parasite life cycle stages. In particular, bloodstream expression sites are active in the procyclic stage in the insect midgut, where VSG is not expressed due to control by transcription attenuation close to the promoter (20,21). As with trypanosome ribosomal DNA promoters, VSG core promoters are ~70–80 bp long with two essential sequences of 5–10 bp, centred at ~30–40 and 60–70 bp upstream of the transcription initiation site (13,14,16,17,22). Other sequences at the initiation site appear to be important too (13,14,22). Additionally, full activity of the procyclin/PARP and ribosomal promoters requires upstream control elements (UCE), over 100 bp upstream of the core promoter region (16,19). Finally, the ribosomal, procyclin/PARP and bloodstream expression site promoters have functional similarities, since hybrid promoters with exchanged, important elements are still highly active (14,17).

VSG is first expressed in the metacyclic stage in the salivary glands of the tsetse fly as a pre-adaptation to life in the mammal (1). A small, specific set (<28) of metacyclic (M-) VSGs (23–25) is activated, using a mechanism distinct from that used in the bloodstream (26–28) and resulting in expression of a mixture of different antigen types in the metacyclic population, which is thought to facilitate establishment of infection in partially immune hosts in the field (29). M-VSG genes are telomeric and are activated *in situ* in the metacyclic stage (28); these metacyclic expression sites contain the M-VSG gene promoters. There are insufficient metacyclic trypanosomes in flies for direct study, so we have studied M-VSG gene expression in metacyclic-derived trypanosomes in the first 6–7 days of bloodstream infection, during which period these genes continue to be expressed *in situ*. We have shown previously that M-VSG genes are under exclusively transcriptional regulation during the parasite life cycle, the only example of true life cycle stage-specific control of gene expression by transcription initiation in the Kinetoplastida (30). We identified the 1.22 M-VSG gene promoter and showed

*To whom correspondence should be addressed. Tel: +44 141 330 4875; Fax: +44 141 330 5422; Email: j.d.barry@bio.gla.ac.uk

[+]Present address: Institute of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, UK

it to be active *in vivo* at the metacyclic stage, but inactive in both the bloodstream and procyclic stages (30). The promoter, including the transcription initiation site, lies within a 426 bp DNA fragment which can direct expression of a reporter gene in transient transfection (30). Regulation of this promoter depends on its chromosomal location, since relocation onto a plasmid or into the non-transcribed spacer region of rDNA led to its activation in bloodstream trypanosomes (31). Factors related to positioning at this telomeric locus may down-regulate the M-VSG gene promoter in the bloodstream.

Since M-VSG gene expression continues for up to 7 days post-infection, our finding of 1.22 promoter function in bloodstream trypanosomes when dissociated from any stage-specific regulatory elements was not unexpected. We have now used this observation to map sequence elements which, though not involved in life cycle stage regulation, are important in basic promoter function. We find that the promoter is more complex than other trypanosome promoters, and that it contains an essential element which, on its own, can act as a promoter exclusively in bloodstream stage trypanosomes.

MATERIALS AND METHODS

Trypanosomes

A virulent, cloned line of *T.brucei* EATRO 795 (28) which retains fly transmissibility was subjected to standard methods for tsetse fly transmission, maintenance in mammalian hosts and harvesting from rodents (32), except that infected blood collected from animals was maintained at 37°C throughout subsequent procedures. Procyclic culture form trypanosomes, established by standard methods, were maintained in SDM-79 medium (33). Flies were fed on mice, from which ~2×10^6 metacyclic-derived trypanosomes/ml blood were harvested 5 days later.

Recombinant clones

All clones containing 1.22 promoter sequences were derivatives of p122sHD52CAT (30), where the chloramphenicol acetyltransferase (CAT) reporter gene is controlled by the 1.22 promoter within a 426 bp *Bam*HI/*Pst*I restriction fragment containing the 1.22 gene transcriptional start site. The initiation site in metacyclic-derived trypanosomes is either the G at 171 or the C at 172 bp downstream of the first G of the *Bam*HI site (Fig. 2C). It is not possible to be precise about this since the 5′ RT-PCR 5′ anchor ends in a G run, but there is a G in the promoter exactly at the start site. For the purposes of calculating distances and fragment sizes the initiation site has been set at the G 171 bp from the *Bam*HI site. Exonuclease III deletions were obtained using the Erase-a-Base system (Promega) exactly according to the protocol supplied. Plasmid p122Δ5′-120/+24 was prepared by PCR amplification of a 144 nt region whose 5′-end was 51 bp 3′ of the *Bam*HI site with a 3′-end 24 bp downstream of the 1.22 transcription initiation site using a *Sac*I-linked 5′ primer (5′-CGGCGAGCTCGCCGACATCCGTACCC-3′) and a *Sma*I-linked 3′ primer (5′-TGACCCGGGACGAATGTCGCTGC-GG-3′). Amplified products were cleaved with *Sac*I and *Sma*I and inserted into *Sac*I/*Sma*I-cut pHD52CAT from which the bloodstream expression site promoter had been removed. All clones and deletions were checked by sequencing. Construction of other reporter gene plasmid clones has been described previously (30). For convenience, the trimmed promoter inserts in deletion constructs are referred to in Results by the position of their terminal bases relative to the transcription initiation nucleotide used in metacyclic-derived trypanosomes.

DNA sequence analysis

Sequencing was carried out on denatured double-stranded plasmid DNA by the dideoxy chain termination method (Sequenase kit; Amersham International). The sequence of both strands of recombinant plasmids was obtained with the recommended primers for pBluescript or custom-synthesised specific primers. Computer analysis was carried out with the Genetics Computer Group sequence analysis software package.

Transient transfection

Transient transfection of metacyclic-derived trypanosomes was performed as described (30), except that trypanosomes were electroporated in blood which had been maintained at 37°C since harvesting. Transient transfection of bloodstream form trypanosomes was performed in a similar manner, except that trypanosomes were obtained in the buffy coat by centrifugation of blood for 10 min at 800 *g*, then cells were diluted in Zimmerman post-fusion medium (34) to 5×10^6/ml before electroporation. Procyclic trypanosomes, harvested at 6×10^6/ml, were transiently transfected exactly as described (30). CAT reactions were carried out for 2 h at 37°C, and assays were undertaken by xylene extraction (35). Transfections were performed in triplicate and results presented are an average from three experiments.

5′ RT-PCR

For analysis of the transcription initiation site used in p122sHD52CAT and p122Δ3′-75 (containing the –171 to –75 bp 3′ deletion fragment described in Results; see Fig. 1A), first strand CAT cDNAs were synthesised from total RNA isolated from bloodstream form *T.brucei* 5 h after electroporation and incubation in HMI-9 medium at 37°C. CAT transcripts were hybrid-selected prior to amplification, using CAT DNA fragments bound to nylon membrane to exclude any primary transcripts arising from the endogenous 1.22 M-VSG gene locus. 5′ RT-PCR was performed with a 5′ RACE kit (Life Technologies) exactly as described in the manufacturer's protocol. For p122sHD52CAT, first strand synthesis used a CAT-specific primer, 1.22CATGSPSG1 (5′-CCGGATGAG-CATTCATCAGGCGGGC-3′), whose 5′-end is located 215 bp downstream of the CAT translation initiation codon (30). cDNAs were purified from unincorporated nucleotides and then tailed with an oligo [dC] anchor. The first round of amplification was performed with oligonucleotide 122GSP1 (5′-GGCCAGGCTTGCGGAGTC-GCACG-3′) and the kit's anchor primer which contains an oligo [dG] anchor region attached to a universal amplification primer region. A second round of amplification was performed with oligonucleotide 122GSP2 (5′-CACGTCCGGGCCTCCGAGTAG-TTGG-3′) and the kit's universal amplification primer. A final round of amplification was undertaken with 122GSP3 (5′-CAAAGCTA-CAGCTCAAGCGCCAAGACG-3′) and the universal amplification primer. PCR amplification was performed for 30 cycles of 30 s at 94°C, 1 min at 67°C (first round of PCR), 63°C (second round of PCR) or 65°C (third round of PCR) and 1 min at 70°C. In the case of p122Δ3′-75, the primer for first stand cDNA synthesis was GSP3′CAT (5′-TACGCCCCGCCCTGCCACTCATCGC-3′), whose 5′-end is located at the CAT coding sequence stop codon. For 5′ RT-PCR, the first round of amplification was undertaken

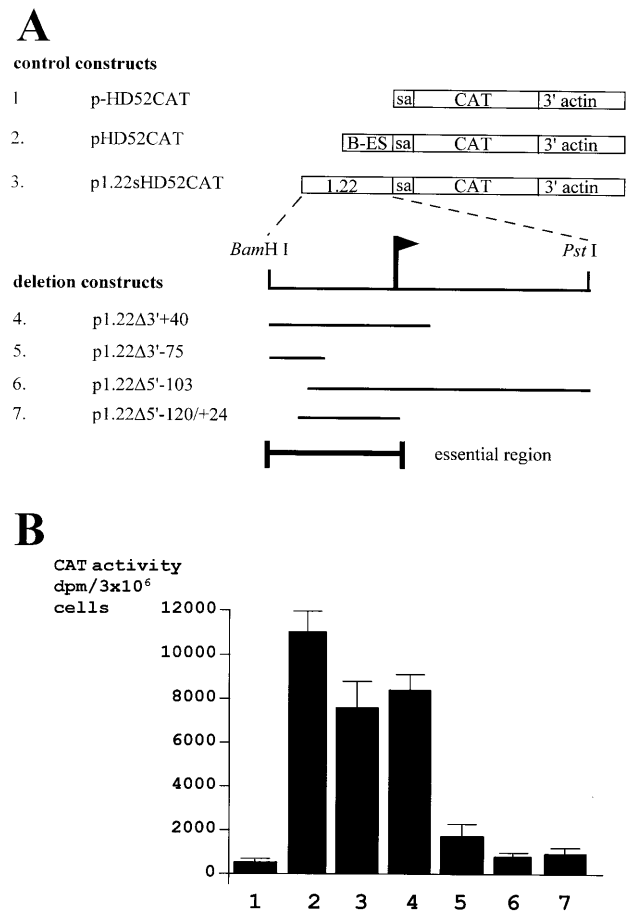

Figure 1. Functional analysis of selected promoter deletions in metacyclic-derived trypanosomes. Promoter deletions were assayed for ability to direct transcription of the CAT reporter gene in transient transfection of metacyclic-derived cells. (**A**) Schematic illustrations of the plasmids. 1, p-HD52CAT, promoterless negative control; 2, pHD52CAT, bloodstream expression site promoter positive control; 3, p122sHD52CAT, the parent test construct. For the experimental constructs 4 (p122Δ3′+40), 5 (p122Δ3′–75), 6 (p122Δ5′–103) and 7 (p122Δ5′–120/+24), only the extent of the promoter fragment contained in the reporter gene construct is shown against a simple map of the 426 bp *Bam*HI/*Pst*I fragment (see Fig. 2). The thick horizontal bar with end bars is the deduced essential region for promoter activity. Abbreviations: sa, splice acceptor region; CAT, chloramphenicol acetyltransferase coding sequence; 3′ actin, actin 3′ untranslated region; B-ES, 221 bloodstream VSG expression site promoter; 1.22, 426 bp *Bam*H I/*Pst* I 1.22 promoter fragment. (**B**) Assay in metacyclic-derived trypanosomes. The bar chart shows CAT activity in disintegrations per minute per 3×10^6 cells for each construct in transient transfection of day 5 metacyclic-derived trypanosomes. Values for each bar are the means and standard deviations of triplicate assays from one representative experiment.

with 1.22CATSPSG1 and the anchor primer at 63°C as above, then two subsequent rounds of amplification were carried out with actinGSP1 (5′-GGCAGCAACGAGACCTTACGT-3′) at 55°C and actinGSP2 (5′-GAGCGGAGACTGCAATGCAGAG-3′) at 57°C, both paired with the universal primer. Reactions were performed in thin walled tubes in a final volume of 10 µl containing 10 µM each primer, 1 mM dATP, dCTP, dTTP and dGTP, 45 mM Tris–HCl, pH 8.8, 11 mM $(NH_4)_2SO_4$, 4.5 mM $MgCl_2$, 6.7 mM 2-mercaptoethanol, 4.4 µM EDTA, pH 8.0, 113 µg/ml BSA and 2 U *Taq* DNA polymerase (Applied Biosystems). PCR products were resolved by gel electrophoresis in 1.5% agarose, cloned using the 'T-vector' system (Promega) and sequenced by the dideoxy-chain termination method (Sequenase kit).

Table 1. Functional assays of promoter deletion fragments in bloodstream trypanosomes

Deletion series	Fragment tested	% CAT activity
Long 5′ deletions	–171 to +255	100
	–66 to +255	8 ± 5
	–29 to +255	9 ± 6
	+62 to +255	9 ± 6
	+169 to +255	8 ± 5
	promoterless	9 ± 6
Short 5′ deletions	–150 to +255	17 ± 7
	–136 to +255	18 ± 7
	–120 to +255	18 ± 8
	–103 to +255	19 ± 7
	promoterless	18 ± 6
3′ deletions	–171 to +40	94 ± 8
	–171 to –16	95 ± 4
	–171 to –75	83 ± 5
	–171 to –110	19 ± 6
	promoterless	20 ± 2
5′ + 3′ deletion	–120 to +24	18 ± 4

Each value is the mean percentage CAT activity and standard deviation from three separate transient transfection experiments, each performed in triplicate. Although absolute CAT values varied between experiments, as also noted by others (44,45), percentage activities for each construct, relative to the positive control pHD52CAT set at 100%, were always very similar. 'Promoterless' means the promoterless plasmid p-HD52CAT was used.

RESULTS

Sequence elements responsible for 1.22 promoter activity in bloodstream trypanosomes

The 1.22 transcription initiation site is 171 bp from the 5′-end of the 426 bp *Bam*HI/*Pst*I fragment derived from the 1.22 locus (30). This fragment, the smallest previously shown to drive reporter gene expression, is contained in plasmid p122sHD52CAT upstream of a CAT gene flanked by actin RNA processing signals. We deleted the 426 bp fragment in both directions, initially in a series of large 5′ deletions across the promoter region. Table 1 shows that none of these deletion mutants gave activity much above background, indicating that sequences over 66 bp upstream of the transcription initiation site were important for promoter activity in bloodstream trypanosomes. Background CAT activity was much higher in bloodstream than in procyclic stage transient transfection, but was very similar to the level obtained using p-HD52CAT, the promoterless construct, and pGEM3, the non-recombinant parent plasmid. We next tested a series of short 5′ deletions (Table 1). Once again, only the undeleted construct gave high CAT activity, with even the shortest deletion reducing activity to background, indicating that sequences relatively far away from the transcriptional start site were necessary for full activity at this life cycle stage. Deletions from the 3′-end (Table 1) showed that sequences downstream of the initiation site were not associated with significant promoter activity. Surprisingly, however, deletion of sequences around the transcription initiation site used in metacyclic-derived trypanosomes also had no effect on CAT expression (–171 to –16 fragment, Table 1) and further deletion up to 75 bp 5′ of the initiation site still yielded an active promoter. Deleting a further 35 bp (–171 to –110 fragment) greatly reduced activity, implying that sequences between 75 and 110 bp 5′ of the initiation site were essential for promoter activity in

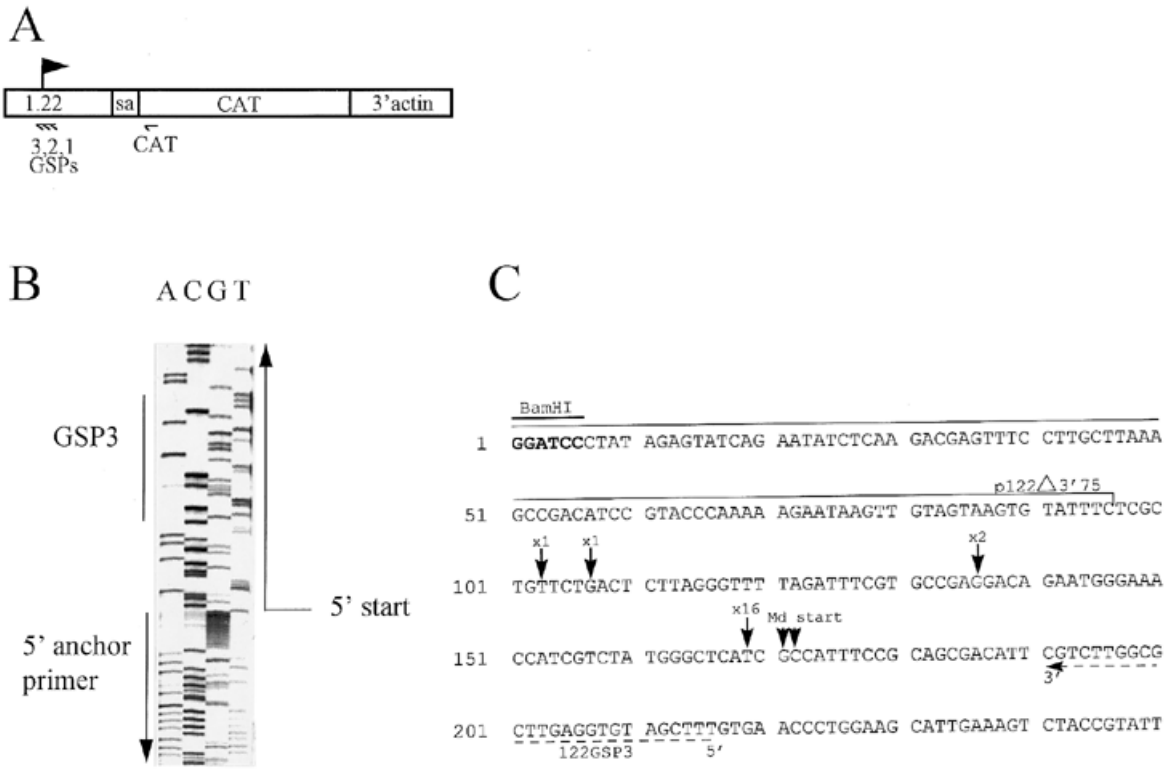

Figure 2. Transcription initiation sites for the 1.22 promoter in bloodstream trypanosomes. (**A**) Schematic of the plasmid p122sHD52CAT and the approximate location of the 5′ RT-PCR primers (GSPs 1,2,3). Abbreviations: sa, splice acceptor region; CAT, CAT coding sequence or specific primer; 3′ actin, actin 3′ untranslated region; 1.22, 426 bp *Bam*HI/*Pst*I 1.22 promoter fragment. (**B**) Sequence autoradiograph of the junction of the 5′ RT-PCR anchor primer ending with the oligonucleotide G anchor, followed by the start of the primary transcript for the CAT gene in transient transfections with p122sHD52CAT. '5′ start' indicates the initiation T in 16 of the 20 clones sequenced and the direction of transcription is indicated by the arrow. The 5′ anchor primer is indicated and GSP3 shows the location of the RT primer. (**C**) Sequence at the 5′-end of the 426 bp promoter fragment. The transcription initiation site in metacyclic-derived trypanosomes (Md start) is indicated by two adjacent arrowheads. Vertical arrows indicate initiation sites, and '$\times n$' shows their frequency in the 20 PCR clones analysed. The location and orientation of the GSP3 primer are indicated with a dashed arrow. The extent of the deletion p122Δ3′-75 is overlined.

bloodstream trypanosomes. The –120 to +24 fragment, which covers a region that in other trypanosome promoters encompasses core elements, yielded only background CAT levels in our experiments, again suggesting that the bloodstream stage activity resided in the 5′ portion of the 1.22 promoter. Combining all the data, activity in bloodstream stage trypanosomes appears to be located in the 5′ portion of the 1.22 promoter.

The 5′ promoter activity is bloodstream stage specific

We tested selected deletions in transient transfection of day 5 metacyclic-derived trypanosomes. Figure 1 shows that the 426 bp *Bam*HI/*Pst*I fragment (p1.22sHD52CAT) gave ~70% of the CAT activity observed with the positive control, the 221 bloodstream expression site promoter, as we reported before (30). The fragment containing the transcription initiation site flanked upstream by 171 bp and downstream by 40 bp gave essentially full activity. Similar to our observations in bloodstream trypanosomes, the fragments containing, respectively, from –103 to +255 bp (p122Δ5′-103) and –120 to +24 bp (p122Δ5′-120/+24) gave only background CAT activity. However, this time the fragment from –171 to –75 (p122Δ3′-75), which strongly directed CAT expression in bloodstream trypanosomes, gave only low activity. As in the bloodstream experiments, some variation in CAT activity was observed between experiments, but the activities relative to the positive control, pHD52CAT, were always very similar. Taken together, these results suggest that a region encompassing the transcription initiation site and stretching to 171 bp upstream is essential for full promoter activity at the metacyclic life cycle stage. A similar series of transient transfections was undertaken in procyclic trypanosomes but this time all the deletion fragments were recloned into reporter gene constructs with procyclin/PARP RNA processing signals [derived from p5′parpCAT3′parp (30,36)], to avoid difficulties inherent in the use of actin RNA processing signals at this life cycle stage (37). All constructs were checked by sequencing. The wild-type, and all the deleted, promoter constructs gave merely background CAT activity in procyclic trypanosomes.

Transcription initiation on the 1.22 promoter in bloodstream trypanosomes

As the promoter activity observed in the bloodstream trypanosome transient transfections appeared to be due to sequences residing at some distance upstream of the initiation site used in metacyclic-derived trypanosomes, we used 5′ RT-PCR to determine the initiation site used in the bloodstream trypanosomes. Bloodstream stage cells were electroporated with the positive control plasmid, p122sHD52CAT, and CAT transcripts were reverse transcribed with a CAT-specific primer, then PCR amplification was undertaken with three 3′ nested 1.22 promoter-specific oligonucleotides (Fig. 2A). Two independent transient transfection experiments each yielded a single PCR product that hybridised to a 1.22 promoter

Figure 3. Alignment of promoter sequences. The AnTat 1.3A sequence is the closest to a consensus bloodstream VSG expression site core promoter, in comparison with the 221 and 118 promoters (7,9,12). AnTat 11.17 is the closest to a consensus putative metacyclic VSG core promoter in comparison with the MVAT4, 5 and 7 putative promoters (14,42). 1.22 distal is the sequence 100–168 bp upstream of the 1.22 transcriptional start site. Sequences are aligned by Clustal W analysis, introducing gaps to maximize alignment, and identical nucleotides are boxed in a grey background. For the first two promoter classes, sequences shown to be important for function are boxed and the transcription initiation base is emboldened; 1.22 initiates 100 bp further downstream.

probe (data not shown). Ten amplified DNAs were cloned from each reaction and sequencing revealed that 16 gave the same initiation site, a T residue 2 or 3 nt upstream of the transcription initiation site identified in metacyclic-derived trypanosomes. It was not possible precisely to locate the initiation nucleotide, since the 5′ anchor used in the 5′ RT-PCR ends in a run of G residues and there is a G nucleotide in the promoter sequence at the start site. Figure 2B shows the sequence of one of the 16 T residue start site clones, with the anchor primer and, immediately adjacent, the initiation site and downstream region present in the primary transcript that templated the RT-PCR. Of the remaining four RT-PCR clones, the initiation sites in two were at 34 or 35 nt upstream (again, the presence of G residues in the promoter sequence makes it impossible to determine which is the exact site of initiation) and in the other two at 65 and 69 nt upstream respectively (Fig. 2C). However, none of these sites was present in the –171 to –75 fragment that drove transcription in the bloodstream form (Table 1). These results show that, despite promoter activity in bloodstream transient transfection apparently being due to sequences over 100 nt 5′ of the initiation site, the actual transcription initiation site on the reporter gene constructs is often similar in bloodstream trypanosomes and in metacyclic-derived trypanosomes, which do not display dependence solely on this 5′ region (30). However, they are not identical.

Using a similar approach (see Materials and Methods), we attempted to determine the transcription initiation site used in the plasmid p122Δ3′-75, which contains 96 nt from the 5′-end of the promoter fragment (the –171 to –75 fragment) and drives high levels of reporter gene expression in bloodstream trypanosomes. It was likely that transcription initiating within this fragment would do so close to its 3′-end. A product was obtained of around the correct size, containing the anchor primer at the 5′-end and the final actin splice acceptor-specific product at the 3′-end. However, a run of 14 T residues in the actin splice acceptor region was absent, indicating that errors in the short product had been introduced during amplification, perhaps due to polymerase slippage on this particular sequence motif, and therefore it was not possible to define a specific start within the 1.22 promoter fragment.

Conservation of VSG promoter elements in the 1.22 5′ promoter fragment

Figure 3 compares the sequence of the 1.22 distal promoter region (68 bp running downstream from the *Bam*HI site) with that of the AnTat 11.17 M-VSG gene putative promoter (14), as a representative of a class of non-stage regulated M-VSG gene putative promoters, and with that of the AnTat 1.3A VSG gene as a typical example of a bloodstream expression site (B-ES) promoter (12). By Clustal W analysis there is significant homology among the three: 37% between 1.3A and 11.17; 11% between 1.3A and 1.22 and 55% between 11.17 and 1.22. In particular, there is a high degree of conservation among all three promoter types over the functional elements (14) (boxed sequences). Although it remains to be demonstrated that the conserved 1.22 elements are essential for function, at least in bloodstream transient transfection, the smallest 5′–3′ deletion of the 1.22 promoter displayed no activity and this lacked 17 bp from the *Bam*HI site, leaving a promoter region without the highly conserved TAT-containing region which is crucial in VSG constitutive promoter activity (14). There are no significant homologies between these non-stage-regulated VSG gene promoters and the remaining 100 bp in the 3′ proximal portion of the 1.22 promoter.

DISCUSSION

In Kinetoplastida, most genes are organised in polycistronic transcription units and the promoters characterised thus far are active, at different levels, throughout the parasite life cycle (10,11). The best known case is in the procyclic stage, where no VSG is expressed, yet the bloodstream VSG gene expression site promoters are active in their own loci at a low level (14,20,21) and can display high activity in plasmids during transient transfection (7,12). The lack of VSG expression is due to transcription attenuation close to the promoter itself. The 1.22 M-VSG gene promoter is the only known promoter in Kinetoplastida which displays true life cycle stage regulation exclusively at the level of transcription initiation, being inactive in its normal chromosomal location in both bloodstream and procyclic trypanosomes (30,31). We believe that its down-regulation in each life cycle stage may involve different mechanisms. In the bloodstream, it is capable of high activity in a chromosome-internal position or on an episomal vector, indicating that telomere positioning may normally contribute to its silencing in this life cycle stage. In procyclic trypanosomes, on the other hand, it is always inactive regardless of location, indicating involvement of *cis*-acting sequences within the promoter itself (30,31). We have now shown that the promoter is complex, with a discrete, bloodstream stage-specific promoter activity located in a region distal from the metacyclic-specific transcription initiation site. It is important to note that this novel activity is probably never present under normal conditions *in vivo*, having become unmasked in these experiments by our removing it from normal bloodstream-specific down-regulatory mechanisms. Thus, as anticipated, the novel promoter activity was repressed in metacyclic-derived trypanosomes, where proper activity required the full 171 bp promoter fragment immediately 5′ of the transcription initiation site.

It is intriguing that, despite its superficial sequence similarity with other bloodstream promoters which direct transcription initiation immediately adjacent, the distal element directs initiation much further away, very close to the site used in metacyclic-derived trypanosomes. Involvement of sequences in the 3′ portion of the promoter is suggested, and it may be that the proximal portion does direct transcription initiation but also requires the subsidiary, distal element for full activity in the metacyclic stage. The proximal

element has up to 10–13 bp homology, over the initiation site, with other eukaryotic TATA-less promoter initiator elements (38,39). Since the distal element can become activated in the bloodstream stage when removed from its telomeric locus, it must be subject to locus-associated down-regulatory mechanisms. That the 1.22 promoter is much longer than other trypanosome promoters, and the requirement for the presence of both segments to yield full activity in metacyclic-derived trypanosomes, indicate that both interact with transcription factors. Further support for this hypothesis would require pair-wise deletions of the 5′ and 3′ promoter portions and assay in the appropriate life cycle stage.

There are a number of reports of other putative M-VSG gene promoters, but their major structural and functional differences from the 1.22 promoter, and the fact that they have not been shown to function in fly-transmitted trypanosomes, leave their significance open to question. The putative promoters for the MVAT4, MVAT5 (40,41), MVAT 7 (42) and AnTat 11.17 (14) genes do not have the strict life cycle stage regulation of the 1.22 promoter, and their sequences resemble those of the distal segment of the 1.22 promoter and the bloodstream VSG expression site promoter. The MVAT 4, 5 and 7 putative promoters are capable of activation in bloodstream trypanosomes, possibly only with low frequency as this has been seen only after very extensive selection, and their presence, and indeed their sequences, bear some resemblance to a promoter activity that was reported at a similar position upstream of the 118 bloodstream VSG gene (43) but which is subsidiary to the main bloodstream expression site promoter much further upstream. It appears that, within several kilobases upstream of telomeric VSG genes, there is a tendency for presence of pseudopromoters.

Our data suggest, however, that metacyclic stage-specific promoter activity requires a more complex structure. Although the sequence of the distal 1.22 element resembles the MVAT4, MVAT5 and AnTat 11.17 putative promoters, it is functionally distinct, being inactive in procyclic transient transfection.Our data suggest a model in which telomere-associated derepression and stage-specific transcription factors combine to provide metacyclic life-cycle stage specific promoter activity. In the procyclic stage, at least one critical *trans*-acting factor is absent (or a negative regulatory factor is present), preventing transcription initiation. In bloodstream trypanosomes, the distal part of the promoter is capable of activation by bloodstream stage-specific transcription factors, but is prevented from doing so by a locus-associated mechanism such as telomere silencing. It remains to be established how the 1.22 promoter normally retains activity in this chromosomal location in early bloodstream trypanosomes following fly transmission. One possible explanation is that there is an intermediate state between the initiation of infection and the onset of the bloodstream system for VSG expression site switching system at days 6–7 post-infection, before telomere associated repression becomes a dominant regulatory mechanism.

ACKNOWLEDGEMENTS

We wish to thank Christine Clayton for provision of reporter gene constructs and advice during the early part of this work. We thank Pat Blundell for critical reading of the manuscript. This work was supported by the Wellcome Trust.

REFERENCES

1 Vickerman,K. (1985) *Br. Med. Bull.* **41**, 105–114.
2 Cross,G.A.M. (1990) *Annu. Rev. Immunol.* **8**, 83–110.
3 Van der Ploeg,L.H.T., Valerio,D., de Lange,T., Bernards,A., Borst,P. and Grosveld,F.G. (1982) *Nucleic Acids Res.* **10**, 5905–5923.
4 Pays,E., Vanhamme,L. and Berberof,M. (1994) *Annu. Rev. Microbiol.* **48**, 25–52.
5 Johnson,P.J., Kooter,J.M. and Borst,P. (1987) *Cell* **51**, 273–281.
6 Pays,E., Tebabi,P., Pays,A., Coquelet,H., Revelard,P., Salmon,D. and Steinert,M. (1989) *Cell* **57**, 835–845.
7 Zomerdijk,J.C.B.M., Ouellette,M., Ten Asbroek,A.L.M.A., Kieft,R., Bommer,A.M.M., Clayton,C.E. and Borst,P. (1990) *EMBO J.* **9**, 2791–2801.
8 Crozatier,M., Van der Ploeg,L.H.T., Johnson,P.J., Gommers-Ampt,J. and Borst,P. (1990) *Mol. Biochem. Parasitol.* **42**, 1–12.
9 Gottesdiener,K., Chung,H., Brown,S.D., Lee,M.G.-S. and van der Ploeg,L.H.T. (1991) *Mol. Cell. Biol.* **11**, 2467–2480.
10 Clayton,C. (1992) *Prog. Nucleic Acid Res. Mol. Biol.* **43**, 37–65.
11 Graham,S.V. (1995) *Parasitol. Today* **11**, 217–223.
12 Jefferies,D., Tebabi,P. and Pays,E. (1991) *Mol. Cell. Biol.* **11**, 338–343.
13 Zomerdijk,J.C.B.M., Kieft,R., Duyndam,M., Shiels,P.G. and Borst,P. (1991) *Nucleic Acids Res.* **19**, 1359–1368.
14 Vanhamme,L., Pays,A., Tebabi,P., Alexandre,S. and Pays,E. (1995) *Mol. Cell. Biol.* **15**, 5598–5606.
15 Sherman,D.R., Janz,L., Hug,M. and Clayton,C. (1991) *EMBO J.* **10**, 3379–3386.
16 Brown,S.D., Huang,J. and Van der Ploeg,L.H.T. (1992) *Mol. Cell. Biol.* **12**, 2644–2652.
17 Janz,L. and Clayton,C. (1994) *Mol. Cell. Biol.*, **14**, 5804–5811.
18 Chung,H.-M., Lee,M.G.-S. and Van der Ploeg,L.H.T. (1992) *Parasitol. Today* **8**, 414–418.
19 Zomerdijk,J.C.B.M., Kieft,R., Shiels,P.G. and Borst,P. (1991) *Nucleic Acids Res.* **19**, 5153–5158.
20 Pays,E., Coquelet,H., Tebabi,P., Pays,A., Jefferies,D., Steinert,M., Koenig,E., Williams,R.O. and Roditi,I. (1990) *EMBO J.* **9**, 3145–3151.
21 Rudenko,G., Blundell,P.A., Taylor,M.C., Kieft,R. and Borst,P. (1994) *EMBO J.* **13**, 5470–5482.
22 Pham,V.P., Qi,C.C. and Gottesdiener,K.M. (1996) *Mol. Biochem. Parasitol.* **75**, 241–254.
23 Crowe,J.S., Barry,J.D., Luckins,A.G., Ross,C.A. and Vickerman,K. (1983) *Nature* **306**, 389–391.
24 Esser,K.M., Schoenbechler,M.J. and Gingrich,J.B. (1982) *J. Immunol.* **129**, 1715–1718.
25 Turner,C.M.R., Barry,J.D., Maudlin,I. and Vickerman,K. (1988) *Parasitol.* **97**, 269–276.
26 Turner,C.M.R., Barry,J.D. and Vickerman,K. (1986) *Parasitol.* **92**, 67–73.
27 Tetley,L., Turner,C.M.R., Barry,J.D., Crowe,J.S. and Vickerman,K. (1987) *J. Cell Sci.* **87**, 363–372.
28 Graham,S.V., Matthews,K.R., Shiels,P.G. and Barry,J.D. (1990) *Parasitol.* **101**, 361–367.
29 Barry,J.D., Graham,S.V., Matthews,K.R., Shiels,P.G. and Shonekan,O.A. (1990) *Biochem. Soc. Trans.* **18**, 708–710.
30 Graham,S.V. and Barry,J.D. (1995) *Mol. Cell. Biol.* **15**, 5945–5956.
31 Graham,S.V., Wymer,B. and Barry,J.D. (1998) *Mol. Cell Biol.* **18**, 1137–1146.
32 Hajduk,S.L., Cameron,C.R., Barry,J.D. and Vickerman,K. (1981) *Parasitol.* **83**, 595–607.
33 Brun,R. and Schonenberger,M. (1979) *Acta Tropica* **36**, 289–292.
34 Carruthers,V.B., Van der Ploeg,L.H.T. and Cross,G.A.M. (1993) *Nucleic Acids Res.* **21**, 2537–2538.
35 Seed,B. and Sheen,J.-Y. (1988) *Gene* **67**, 271–277.
36 Matthews,K.R., Tschudi,C. and Ullu,E. (1994) *Genes Dev.* **8**, 491–501.
37 Hausler,T. and Clayton,C. (1996) *Mol. Biochem. Parasitol.* **76**, 57–71.
38 Quon,D.V.K., Delgadillo,M.G., Khachi,A., Smale,S.T. and Johnson,P.J. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4579–4583.
39 Javahery,R., Khachi,A., Lo,K., Zenzie-Gregory,B. and Smale,S.T. (1994) *Mol. Cell. Biol.* **14**, 116–127.
40 Alarcon,C.M., Son,H.J., Hall,T. and Donelson,J.E. (1994) *Mol. Cell. Biol.* **14**, 5579–5591.
41 Nagoshi,Y.L., Alarcon,C.M. and Donelson,J.E. (1995) *Mol. Biochem. Parasitol.* **72**, 33–45.
42 Kim,K.S. and Donelson,J.E. (1997) *J. Biol. Chem.* **272**, 24637–24645.
43 Shea,C., Lee,M.GS. and Van der Ploeg,L.H.T. (1987) *Cell* **50**, 603–612.
44 Berberof,M., Vanhamme,L., Tebabi,P., Pays,A., Jefferies,D., Welburn,S. and Pays,E. (1995) *EMBO J.*, **14**, 2925–2934.
45 Biebinger,S., Rettenmaier,S., Flaspohler,J., Hartmann,C., Pena-Diaz,J., Wirtz,E., Hotz,H., Barry,J.D. and Clayton,C. (1996) *Nucleic Acids Res.* **24**, 1202–1211.