The EMBO Journal vol.13 no.22 pp.5460–5469, 1994

A conserved upstream element is essential for transcription of the *Leishmania tarentolae* mini-exon gene

R.M.Saito, M.G.Elgort and D.A.Campbell[1]

Department of Microbiology and Immunology, University of California, Los Angeles, CA 90024, USA

[1]Corresponding author

Communicated by E.M.F.De Robertis

We demonstrate that the mini-exon genes of *Leishmania tarentolae* are individually transcribed by an enzyme pharmacologically identified as RNA polymerase II. To study transcription in these ancient eukaryotes, a stable transformation assay using an episomal mini-exon gene was developed. The introduced mini-exon gene, which had been marked with a 40 bp tag, yielded the predicted tagged transcript. An upstream *cis*-acting element that was essential for transcription of the mini-exon gene was identified by site-directed mutagenesis. Block substitution mutagenesis of the −1 to +9 and +10 to +19 regions of the exon results in 20- to 100-fold decreased levels of the tagged transcript in steady-state RNA. However, since these two mutations resulted in only a 2- to 3-fold decrease in nascent RNA levels, steady-state levels appear to be affected greatest by the stability of the resulting transcript. In contrast, mutation of the −67/−58 region resulted in undetectable levels of both steady-state and nascent RNA from the introduced gene. We conclude, therefore, that this upstream element, which is highly conserved in all *Leishmania* species, is a component of the mini-exon gene promoter.

Key words: kinetoplastid/promoter/proximal sequence element/RNA polymerase II/snRNA gene

Introduction

The mechanisms of transcription initiation within the trypanosomatid protozoa are poorly understood due to the lack of *in vitro* assays and the difficulties in identifying the transcriptional start sites of housekeeping genes, which may be many kilobases upstream (Gibson *et al.*, 1988; Bakalara *et al.*, 1992; Pays, 1993; Wong *et al.*, 1993). The representative kinetoplastids *Trypanosoma brucei* and *Crithidia fasciculata* possess the three standard classes of eukaryotic RNA polymerase (Grondal *et al.*, 1989; Köck and Cornelissen, 1991). The trypanosomatid 5S ribosomal RNA and tRNA genes are conventional in being transcribed by RNA polymerase III (Laird *et al.*, 1985; Hancock *et al.*, 1992; Fantoni *et al.*, 1994) and possessing the appropriate internal consensus promoter elements (Cordingley, 1985; Campbell *et al.*, 1989). The *T.brucei* U2 gene is also transcribed by RNA polymerase III (unlike the U2 genes of other eukaryotes, which are transcribed by RNA polymerase II) using, in part, extragenic regulatory elements that resemble the A and B boxes of tRNA genes (Fantoni *et al.*, 1994). The trypanosomatid ribosomal RNA (rRNA) genes also appear conventional in being transcribed by RNA polymerase I, which shows a high level of resistance to α-amanitin; however, the DNA sequences at the transcription start sites are not well conserved (White *et al.*, 1986; Grondal *et al.*, 1990; Rudenko *et al.*, 1991; Dietrich *et al.*, 1993; Martínez-Calvillo and Hernández, 1994). In contrast, the generation of mature RNAs from protein-encoding genes in these ancient organisms is characterized by a number of unusual features. First, mRNAs consist of two exons that are synthesized as discrete molecules and subsequently joined via a bimolecular *trans*-splicing pathway. Second, housekeeping genes are transcribed by a typical α-amanitin-sensitive RNA polymerase II, while genes encoding the variant surface glycoprotein (VSG) and procyclin/PARP are transcribed by an α-amanitin-resistant RNA polymerase (Kooter and Borst, 1984; König *et al.*, 1989; Rudenko *et al.*, 1989). The identification of transcription start sites has been complicated by the organization of protein-encoding genes into polycistronic units (Johnson *et al.*, 1987; Ben Amar *et al.*, 1991; Wong *et al.*, 1993). The only *cis*-acting elements identified for protein-encoding genes regulate the expression of procyclin/PARP and VSG genes by an α-amanitin-resistant polymerase (Brown *et al.*, 1991; Sherman *et al.*, 1991; Zomerdijk *et al.*, 1991).

In trypanosomatid protozoa, most if not all nuclear mRNAs possess a common 5′-exon (Cornelissen *et al.*, 1986; Walder *et al.*, 1986) that is acquired by *trans*-splicing (Murphy *et al.*, 1986; Sutton and Boothroyd, 1986). This common exon, termed mini-exon (Boothroyd and Cross, 1982; Van der Ploeg *et al.*, 1982) or spliced leader, is 39 nt in length and is derived from a primary transcript (mini-exon derived RNA, or medRNA) of ~95 nt in *Leishmania* sp. (Miller *et al.*, 1986; Fleischmann and Campbell, 1994). Mini-exon genes have been identified in a variety of kinetoplastid organisms. The sequence conservation between the genes from the different organisms is greatest within the exon, notably in two blocks from position +1 to +10 and position +20 to +39 (Cook and Donelson, 1987; Muhich *et al.*, 1987). Experiments to identify the RNA polymerase that transcribes the mini-exon gene in *T.brucei* have yielded conflicting results. On the basis of (i) the effects of 1,10-phenanthroline and manganese ions in the nuclear run-on assay, (ii) the presence of internal box A- and box B-like elements, and (iii) the presence of a T-track at the 3′-end of the transcribed region, Grondal *et al.* (1989) concluded that the mini-exon gene was transcribed by RNA polymerase III. In contrast, RNA polymerase II has been implicated as the responsible enzyme on the basis of the following treatments to the nuclear run-on/permeabilized cell assay:

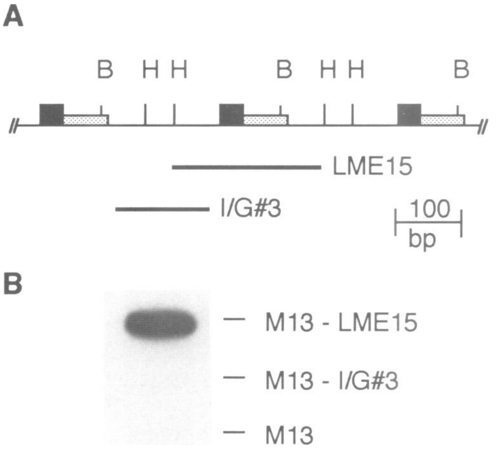

Fig. 1. Genomic organization and transcription of the *L.tarentolae* mini-exon genes. (**A**) Map showing the tandem organization of the mini-exon genes. Solid boxes represent the exon; shaded boxes represent the intron. Restriction sites are *Bss*HII (B) and *Hae*III (H). The location of hybridization probes LME15 and I/G#3 is indicated. (**B**) Nascent RNA, which was labelled by the nuclear run-on technique, was hybridized to the following single-stranded DNA probes: (i) M13-LME15, which contains the 259 nt antisense strand of pLME2 (Fleischmann and Campbell, 1994), (ii) M13-I/G#3, which contains the 204 nt antisense strand of the intergenic region and (iii) M13 vector.

α-amanitin (Laird *et al.*, 1985; Ben Amar *et al.*, 1988), Sarkosyl (Rudenko *et al.*, 1992) and tagetitoxin (cited in Köck and Cornelissen, 1991). In some studies, transcription of the mini-exon gene was more resistant to α-amanitin ($ID_{50} \approx 20$ µg/ml) than the control tubulin gene ($ID_{50} \approx 1$–5 µg/ml) (Laird *et al.*, 1985; Ullu and Tschudi, 1990; Ullu *et al.*, 1993).

In this study we demonstrate that the tandemly repeated mini-exon genes of *L.tarentolae* are independently transcribed by a Sarkosyl-sensitive, tagetitoxin-resistant RNA polymerase. By stable transformation of *L.tarentolae*, we determined the regulatory role of various regions on expression of a tagged mini-exon gene. We identified sequences within the transcribed region whose mutation greatly affects the stability of the resulting transcript. Furthermore, mutation of an upstream element, which is conserved within the *Leishmania* group, abolished transcription of the gene. We note sequence similarities of this essential region with the proximal sequence element (PSE; Mattaj *et al.*, 1985) regulatory element in the promoters of snRNA genes.

Results

The tandemly repeated mini-exon genes are independently transcribed

Kooter *et al.* (1984) found that sequences encoding the mini-exon of *T.brucei* are transcribed at an ~750-fold higher level than the flanking sequences, thus demonstrating that the tandemly repeated genes are not transcribed in a polycistronic manner. The 'run-on' transcription assay was used to determine whether the tandemly repeated mini-exon genes of *L.tarentolae* (Figure 1A) are also independently transcribed. ^{32}P-labelled nascent RNA generated by this procedure was hybridized to constructs containing the mini-exon coding region (M13-LME15), the flanking region clone (M13-I/G#3) or the M13 vector (Figure 1B). Hybridization to the flanking region was equivalent to the vector contol and was at least 200-fold lower than hybridization of labelled RNA to the mini-exon coding region. These data indicate that transcription of each mini-exon gene proceeds independently and that the flanking region may be considered as a 'non-transcribed spacer'. The above result also suggests that each repeat will contain the relevant promoter element(s) for synthesis of the medRNA.

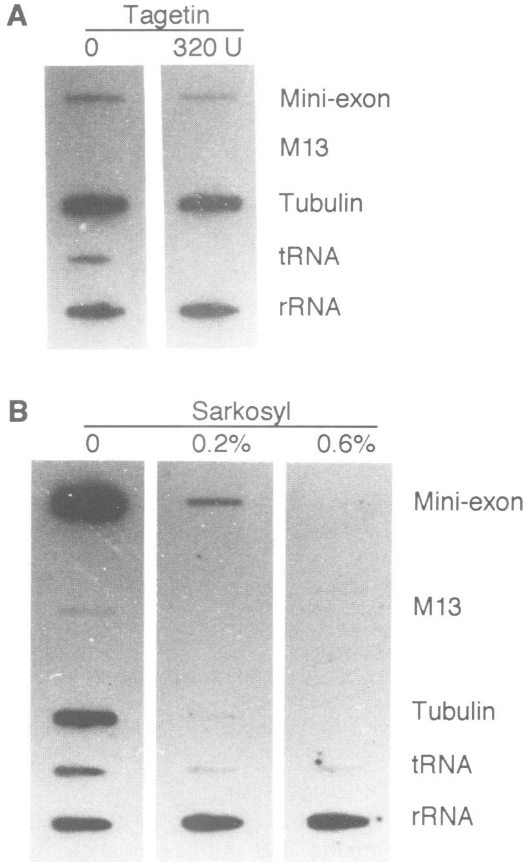

Fig. 2. Identification of the RNA polymerase that transcribes the *L.tarentolae* mini-exon gene. Lysolecithin-permeabilized *L.tarentolae* were treated with 320 U of tagetitoxin (**A**) or 0.2 and 0.6% Sarkosyl (**B**) and nascent RNA labelled by [^{32}P]GTP incorporation. Radiolabelled RNA was hybridized to single-stranded DNA probes for the mini-exon gene, the M13 vector and representative *Leishmania* genes that are transcribed by RNA polymerase I (rRNA), RNA polymerase II (tubulin) and RNA polymerase III (tRNA).

The mini-exon gene is transcribed by RNA polymerase II

To identify the RNA polymerase responsible for synthesis of the medRNA, we employed various RNA polymerase inhibitors in a run-on transcription assay. Studies with the inhibitor α-amanitin were unable to distinguish between RNA polymerases II and III, as had been observed with nuclear run-on assays in *C.fasciculata* (Köck and Cornelissen, 1991); however, they did eliminate a role for RNA polymerase I in the transcription of the mini-exon genes (data not shown).

In the presence of tagetitoxin, a specific inhibitor of

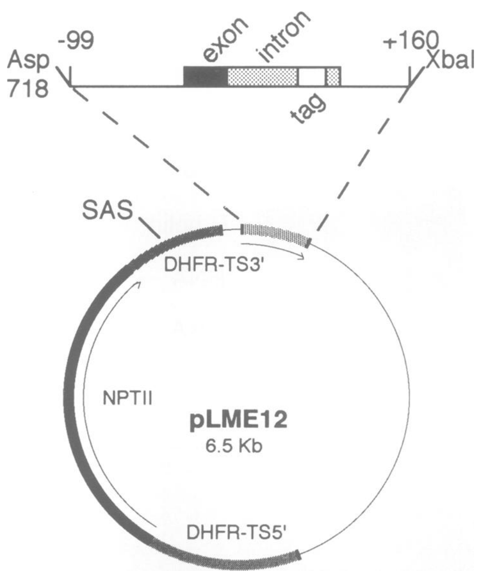

Fig. 3. Map of plasmid pLME12, an episomal vector for expression of the *L.tarentolae* mini-exon genes. The insertion site of the tagged *L.tarentolae* mini-exon gene is shown relative to the neomycin phosphotransferase gene (NPTII) and flanking *L.major* dihydrofolate reductase/thymidilate synthase (DHFR/TS) sequences of the pXδA vector. The major 'downstream' 5'-splice acceptor site (SAS; Kapler *et al.*, 1990b) is indicated. On the upper line, the positions of the exon, intron and tag within the mini-exon gene are indicated.

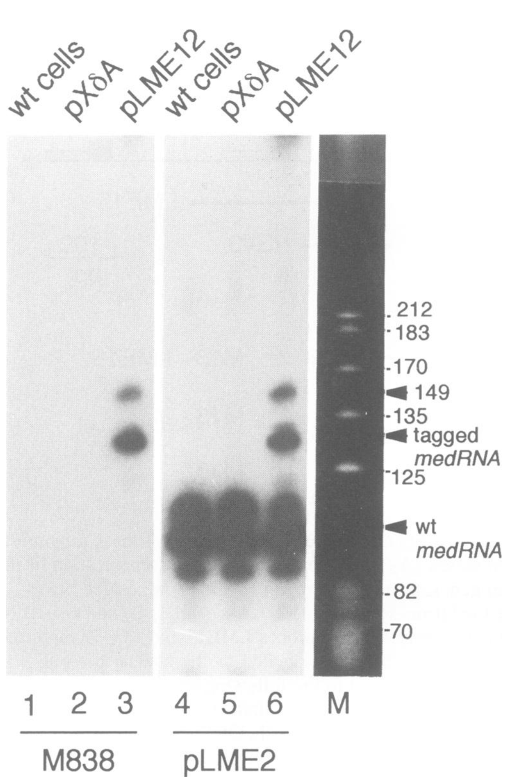

Fig. 4. Northern blot analysis of tagged medRNA transcribed from pLME12. RNA extracted from: wild type *L.tarentolae* (lane 1) or *L.tarentolae* transformed stably with pXδA (lane 2). pLME12 (lane 3), was hybridized first with the tag-specific oligonucleotide M838, stripped, then rehybridized with the pLME2 insert (lanes 4–6). Lane M represents ethidium bromide staining of the resolved small RNA species before transfer. Approximate sizes of the small rRNAs 1–6 (Schnare *et al.*, 1983) and leucine- and glycine-tRNAs (Campbell *et al.*, 1989) are indicated in nucleotides.

RNA polymerase III (Steinberg *et al.*, 1990), transcription of the rRNA, β-tubulin and mini-exon genes was unaffected (Figure 2A). In contrast, transcription of the tRNA genes was completely inhibited by the same concentration of tagetitoxin, suggesting that the mini-exon genes are not transcribed by RNA polymerase III.

Rudenko *et al.* (1992) were able to distinguish between the RNA polymerases of *T.brucei* in the presence of the detergent Sarkosyl. Treatment of permeabilized *L.tarentolae* with Sarkosyl (Figure 2B) had no effect on transcription of the rRNA genes. Transcription of the tRNA genes was reduced to the same level (22% of untreated) by both 0.2 and 0.6% Sarkosyl. In contrast, transcription of both β-tubulin and mini-exon genes was substantially reduced (to 3% of untreated) by 0.2% Sarkosyl and was completely eliminated by 0.6% Sarkosyl, suggesting that the mini-exon is transcribed by RNA polymerase II. Thus, in contrast to a previous report that concluded that the mini-exon gene is transcribed by RNA polymerase III (Grondal *et al.*, 1989), our experiments in *L.tarentolae* are consistent with the results of other groups who concluded that the mini-exon gene is transcribed by an RNA polymerase that most closely resembles the class II enzyme (Laird *et al.*, 1985; Rudenko *et al.*, 1992).

Stable expression of the tagged mini-exon gene

To study expression of the mini-exon gene in stably transfected cells, we tagged an *L.tarentolae* genomic fragment containing the majority of a mini-exon repeat and inserted the resulting construct into the *Leishmania* vector, pX (Figure 3). In addition to the medRNA coding region, the 259 bp *Hae*III fragment contains 99 and 54 bp, respectively, of upstream and downstream sequences. A unique 40 bp sequence was inserted at the *Bss*HII site in the intron of the transcribed region in order to discriminate the episome-derived medRNA (tagged medRNA) from the endogenous medRNA by both its unique sequence and size.

Upon transformation of *Leishmania* by electroporation with the resulting plasmid, pLME12, we detected tagged medRNA transcripts in Northern blots of steady-state low molecular weight RNA. Hybridization with the tag-specific probe revealed no tagged medRNA in RNA extracted from wild type cells or cells transformed with the parent vector, pXδA (Figure 4). In contrast, RNA from cells transformed with pLME12 showed two bands of hybridization with the tag-specific probe (Figure 4, lane 3): a major ~131 nt transcript, which is consistent with the predicted size (~135 nt) of a correctly transcribed tagged medRNA, and a lower abundance ~149 nt transcript. Rehybridization of the filter with a mini-exon probe revealed approximately equal loading of the RNA (lanes 4–6). We conclude, therefore, that the tag sequence was not inserted into an essential control element, and that the cloned *Hae*III fragment appears to contain the necessary information to direct transcription of the mini-exon gene.

Attempts to identify the RNA polymerase that transcribes the episomal mini-exon gene were not successful; however, we were able to eliminate RNA polymerase I because synthesis of tag-specific RNA was sensitive to 25 μg/ml α-amanitin (data not shown).

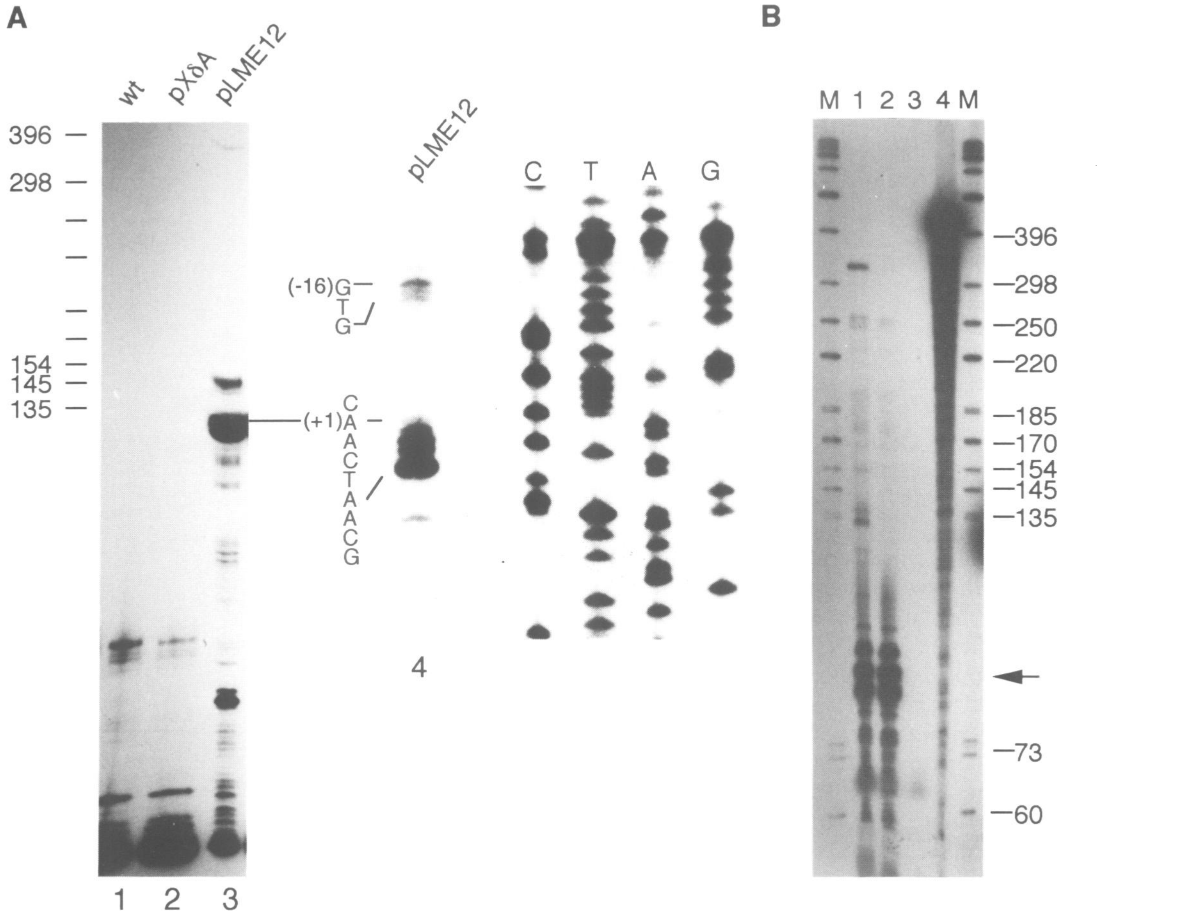

Fig. 5. Analysis of tagged medRNA transcripts by primer extension and RNase protection. (**A**) Primer extension products generated from RNA extracted from wild type *L.tarentolae* (lane 1) or *L.tarentolae* containing pXδA (lane 2), or pLME12 (lane 3), in the presence of γ-^{32}P-labelled oligonucleotide M838 and reverse transcriptase. Molecular size markers represent denatured, *Hin*fI-digested pGB117.1 (Bernards *et al.*, 1981). To determine the precise 5′-end of the tagged RNA molecules, primer extension products (lane 4) were electrophoresed adjacent to a homologous sequence reaction. DNA sequence was generated by the dideoxy termination method using the M838 oligonucleotide on the pLME12 plasmid. (**B**) RNase protection analysis of RNA extracted from pLME12-containing *L.tarentolae* (lane 1), wild type *L.tarentolae* (lane 2), and in the absence of added RNA (lane 3). The input ^{32}P-radiolabelled probe (lane 4) was generated by T7 RNA polymerase from an *Asp*718-linearized pLME12 template. Molecular size markers are denatured, *Hin*fI-digested pGB117.1

Mapping the termini of the tagged medRNA

Based on its size, the 131 nt tagged transcript identified by Northern blotting appears to represent an accurately synthesized medRNA. To determine whether the 131 nt transcript possesses the correct termini, we first mapped the 5′-end of the molecule by a primer extension assay using an oligonucleotide complementary to the tag (Figure 5A). Reverse transcription of RNA extracted from pLME12-containing cells yielded a major product of ~130 nt (lane 3), which was not observed in RNA extracted from wild type cells (lane 1) or cells stably transformed with the vector pXδA (lane 2). Greater resolution of the 130 nt product identified five discrete products (lane 4), which differed in size by a single nucleotide, and corresponded to positions +1 to +5 of the mini-exon. These staggered bands most likely represent premature termination of the reverse transcriptase at the modified 'cap 4' structure of the mini-exon (Freistadt *et al.*, 1987; Perry *et al.*, 1989). Thus, we infer that the majority of tagged medRNAs may be modified at the first four positions and that the 5′-terminus is most likely at the +1 position of the wild type mini-exon.

In addition, two larger, less abundant primer extension products of ~150 and ~350 nt were generated from pLME12-derived RNA (Figure 4, lane 3). The ~150 nt primer extension product possessed 3′-termini at about position −16 (Figure 5A, lane 4) and based on its size, may represent the 5′-terminus of the 149 nt RNA observed in Northern blots (Figure 4, lane 3). The 3′-end of the ~350 nt primer extension product correlates with the location of the *L.major* DHFR/TS 'downstream' splice acceptor site [SAS (Figure 3); Kapler *et al.*, 1990b]. The excised 350 nt product could be demonstrated, by RT-PCR, to possess mini-exon sequences at its 3′-end (data not shown), suggesting that it represents a stable *trans*-spliced product of the run-around transcription that has been demonstrated to occur from episomes in *Leishmania* (Curotto de Lafaille *et al.*, 1992).

To demonstrate that the tagged medRNAs are collinear with the input construct, we performed an RNase protection analysis (Figure 5B). Protection by the endogenous medRNA resulted in heterogeneously sized products of ~84 nt that were present in RNA from transfected and wild type RNA (arrow; lanes 1 and 2). Bands of ~132,

~154 and ~330 nt were protected by RNA from transfected cells (lane 1), but not by RNA from untransfected cells (lane 2) or by no RNA (lane 3). The 132 nt product is consistent with protection of an accurately transcribed, tagged, collinear medRNA. The 154 nt product is consistent with protection of the 149 nt product observed in Nothern blots and the 330 nt product is consistent with the protection by a run-around transcript.

The tagged medRNA possesses a 5′ cap structure

In addition to the methylation of the four 5′-most nucleotides, the medRNA of *T.brucei* possesses a 7-methylguanosine cap structure that is most likely added during the early stages of transcriptional initiation. To demonstrate that the 5′-terminus of the 131 nt tagged *L.tarentolae* medRNA is due to transcription initiation and not a processing event, we determined that the tagged medRNA possesses a cap structure by using an oligonucleotide-directed RNase H assay (Figure 6). To achieve this, we employed RNA extracted from cells containing the Δ30/39 derivative of pLME12 (to be described later; Figure 7A), in which positions +30 to +39 of the mini-exon have been substituted. In the presence of an oligonucleotide complementary to positions +25 to +39, we observed an RNase H-dependent band of ~24 nt that could only be labelled by guanylyltransferase if previously decapped chemically (lane 4). This demonstrates that the tagged medRNA possessed a 5′-cap structure, which must be removed before labelling can occur. A second oligonucleotide/RNase H-dependent band, which could be capped after chemical decapping, was observed at ~260 nt (lane 4, arrowed). Consistent with previous results, this band is further evidence for *trans*-splicing of the run-around transcript. No capped species corresponding to the 149 nt RNA was observed. However, additional experiments are required to determine if its 5′-end was created by upstream transcription initiation or by an RNA processing event. Because the combination of primer extension, nuclease protection and capping studies strongly indicates that the 131 nt molecule represents the predicted tagged medRNA, we have followed this molecule in functional assays as a reporter of transcription from the episomal gene.

Mutagenesis of the 5′-end of the mini-exon reduces the level of the steady-state transcript

The high level of sequence conservation within the mini-exons of kinetoplastid protozoa and the function of the exon as a promoter element in *Ascaris* (Hannon *et al.*, 1990) suggested a similar regulatory function in *Leishmania*. To test this possibility, a series of 10 bp block substitutions (abbreviated to Δ here) in the region −31 to +49 were produced by oligonucleotide-directed mutagenesis. This linker-scanning approach, which introduced the sequence CCTCGAGGAA, produced six to 10 base substitutions in each case (Figure 7A). Plasmids bearing each mutation were introduced into *L.tarentolae* and stably transformed populations selected in the presence of G418. The effect of the mutations on the level of tagged medRNA in stable RNA populations was assayed by Northern blotting. Hybridization of the blot with a tag-specific probe (Figure 7B, left panel) revealed the presence of the 131 and 149 nt transcripts in all but one sample. Mutation of the region −1/+9 reduced the steady-state level of tagged

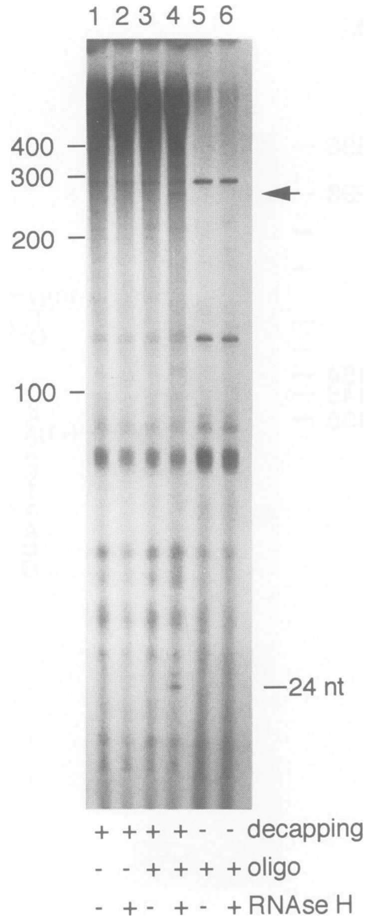

Fig. 6. Detection of a cap structure on the plasmid-encoded medRNA. An oligonucleotide-directed RNase H assay was used to specifically cleave RNA corresponding to the mutated exon (positions 30–39) transcribed from pLME12Δ30–39. RNA extracted from pLME12Δ30–39 stably transformed *L.tarentolae* was incubated with the vaccinia virus guanylyl transferase and [α-^{32}P]GTP following chemical decapping (lanes 1–4) or mock cap removal (lanes 5 and 6). The oligonucleotide 5′-TTCCTCGAGGCTGAA, which represents the antisense strand at the Δ30–39 mutation, was added to aliquots of both decapped and untreated RNA (lanes 3–6). The samples were incubated in the absence (lanes 1, 3 and 5) or presence (lanes 2, 4 and 6) of RNase H before resolution through a denaturing acrylamide gel. Molecular size markers (nt) are the BRL 100 nt ladder. Oligonucleotide specific bands of ~24 and ~260 nt (arrow) are indicated. Bands that are labelled by the capping enzyme in RNA that was not decapped (lanes 5 and 6) most likely represent 5S rRNA (125 nt) and 7SL RNA (273 nt), both of which possess a 5′ di- or triphosphate terminus (Lenardo *et al.*, 1985; Michaeli *et al.*, 1992).

transcript by at least 100-fold (Figure 7B, lane 5). Longer exposure of the autoradiograph and primer extension revealed very low levels of the two transcripts (data not shown). We also observed a 20-fold reduced level of the tagged transcripts with the mutation at positions 10/19 (Figure 7B, lane 6). In control experiments, we show that the reduced levels of tagged RNAs in the Δ−1/+9 and Δ10/19 samples did not result from unequal loading of RNA (Figure 7B, right panel, lanes 5 and 6) or from proportionately decreased levels of the plasmids within the cells (Figure 7C, lanes 5 and 6) relative to the pLME12 control (Figure 7C, lane 1), suggesting that the two mutations result in differences of either transcription rates or RNA stability. It is also noted that two mutations

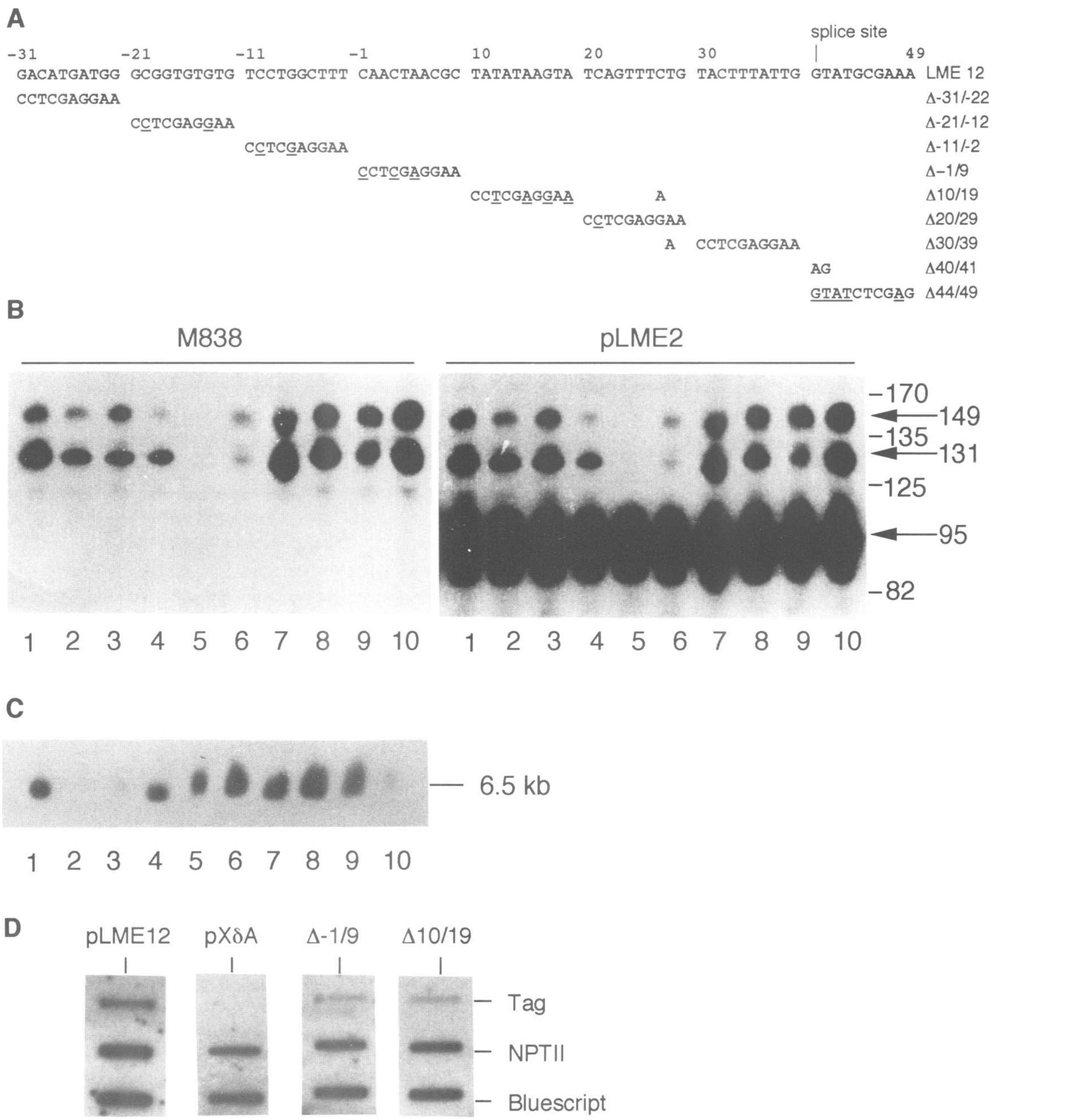

Fig. 7. Analysis of tagged medRNA transcription from pLME12-derived plasmids containing block substitution mutations. (**A**) Nucleotide changes introduced by oligonucleotide-directed mutagenesis are indicated under the sequence for the respective block substitution mutation. Underlined nucleotides are identical to those present in the wild type sequence (LME12). (**B**) Northern blot analysis of RNA from *L.tarentolae* transformed stably with plasmids containing the block substitution mutations. RNA that had been extracted from cells containing pLME12 (lane 1), Δ−31/−22 (lane 2), Δ−21/−12 (lane 3), Δ−11/−2 (lane 4), Δ−1/9 (lane 5), Δ10/19 (lane 6), Δ20/29 (lane 7), Δ30/39 (lane 8), Δ40/41 (lane 9) and Δ44/49 (lane 10) was resolved by polyacrylamide gel electrophoresis and transferred to a nylon membrane. The RNA was hybridized first with the tag-specific oligonucleotide M838 (left panel). After removal of the probe, the RNA was hybridized with the pLME2 insert (right panel). The Δ40/41 mutation was tested to determine if removal of the consensus splice donor site would stabilize the labile medRNA. The lack of increased levels of tagged medRNA in this sample suggests that even with the absence of processing by *trans*-splicing the medRNA has a very short half-life. (**C**) Southern blot analysis of DNA extracted from cells harbouring the mutation-containing plasmids. DNA was digested with *Xba*I. Samples are numbered identically to those in (B). Presence of the plasmid was indicated by hybridization with the radiolabelled tag-specific oligonucleotide M838. The lack of a signal in lane 2 is due to the unexpected loss of the *Xba*I site during subcloning; the uncut plasmid remained in the well (data not shown). (**D**) Analysis of nascent RNA from plasmid-derived sequences. ^{32}P-labelled nascent RNA from *L.tarentolae* containing the indicated plasmids was hybridized to single-stranded probes that are specific for the tag, the NPTII gene or the denatured pBluescript. Bound labelled RNA was quantified using a Molecular Dynamics PhosphorImager.

(Δ−31/−22 and Δ44/49) apparently result in an increased level of tagged transcripts (Figure 7B, lanes 3 and 10) relative to the plasmid copy number (Figure 7C, lanes 3 and 10).

To discriminate between effects on transcription versus RNA stability, the two mutations (Δ−1/+9 and Δ10/19) were tested for relative levels of labelled nascent RNA in the run-on assay. Both mutations show a 2- to 3-fold reduction in transcription of tag-specific sequences relative to the positive control, pLME12 (Figure 7D). In the negative control (cells transformed with the vector pXδA), no background hybridization to tag sequences is observed. In addition, approximately equivalent levels of transcription are observed in all four samples from other plasmid-specific sequences (NPTII and pBluescript vector), which indicates no loss of general transcriptional competence from the plasmids. Since the stable RNA levels from the Δ−1/+9 and Δ10/19 clones are reduced >100- and 20-fold respectively, we conclude that the major effect of these two mutations is on RNA stability.

Upstream sequences are essential for transcription of the mini-exon

Because the exon does not function as an essential promoter element in *L.tarentolae* and because most RNA polymerase II promoter elements lie upstream of the respective gene, we have searched the 5′-flanking region of the mini-exon for promoter elements. First, we generated a plasmid in which the region −99 to −29 was deleted (δ−99/−29). Steady-state RNA from cells containing this plasmid contained no detectable tagged medRNAs (Figure 8A, lane 2) by Northern blot, even after longer exposure (data not shown).

Comparative DNA sequence analysis of the mini-exon gene from multiple human-pathogenic *Leishmania* species (Fernandes *et al.*, 1994) revealed a highly conserved block located within the δ−99/−29 region. In *L.tarentolae*, this conserved block is located at positions −67 to −58 (Fleischmann and Campbell, 1994). We therefore determined whether this conserved block was essential for transcription of the mini-exon gene. The region −67/−58 was mutated by introducing transversions (T↔G and A↔C) within the conserved block. The resulting plasmid was stably transformed into *L.tarentolae* and RNA was assayed for the presence of tagged medRNAs. No tagged medRNA was detected (Figure 8A, lane 3). For both upstream mutations, the absence of tagged medRNAs is not due to unequal RNA loading (Figure 8A, lanes 5 and 6) or differences in the plasmid copy number (Figure 8B). Because neither Δ−67/−58 nor δ−99/−29 mutations are contained within the primary medRNA transcript, it is likely that the absence of tagged transcripts is due to inactivation of transcription rather than effects on RNA stability.

Δ−*67/−58 mutation inactivates transcription of the mini-exon gene*

To test the hypothesis that the Δ−67/−58 mutation inactivated transcription, we analysed nascent RNA corresponding to the tag (Figure 8C). Cells transformed with the plasmid pLME12 contain significant levels of labelled, tag-specific nascent RNA; in contrast, no tag-specific hybridization is observed in labelled RNA extracted from cells containing the vector only (pXδA). Cells containing the Δ−67/−58 mutation show no detectable level of tag-specific nascent RNA, which indicates that transcription of the gene has been inactivated. In each case, approximately equivalent levels of transcription from the NPTII and Bluescript sequences are observed, suggesting that the Δ−67/−58 mutation does not affect transcription of other pX sequences. Because this mutation inactivated transcription of only the tagged mini-exon gene, we conclude that the conserved region serves as all or part of an essential promoter element.

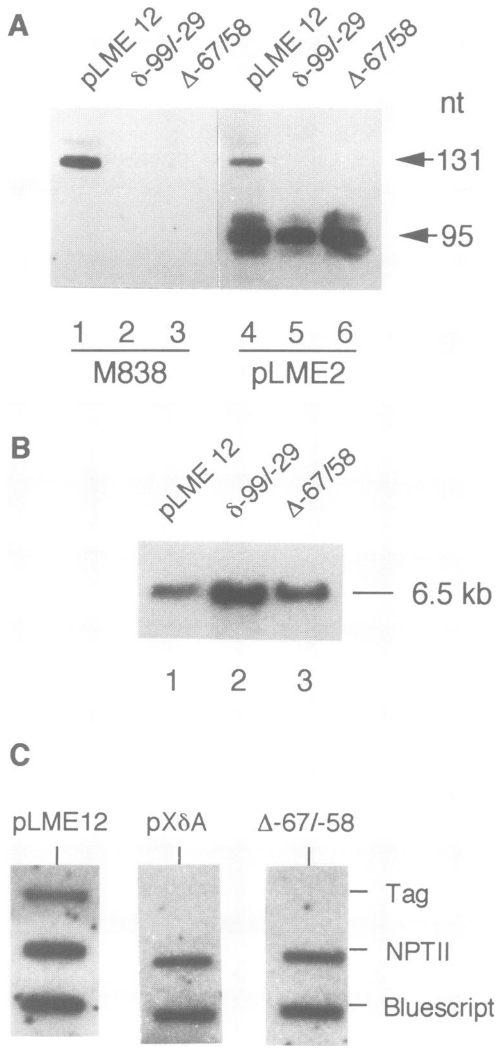

Fig. 8. Effect of an upstream deletion or Δ−68/−58 mutation on tagged medRNA transcription. (**A**) Northern blot analysis of RNA extracted from *L.tarentolae* transformed with pLME12 (lane 1), δ−99/−29 (lane 2) or Δ−67/−58 (lane 3). The blot was first hybridized with the tag-specific oligonucleotide M838 (left panel), stripped by boiling in water, and then hybridized with the complete mini-exon gene repeat (right panel). (**B**) Southern blot analysis of DNA extracted from *L.tarentolae* transformed with pLME12 (lane 1), δ−99/−29 (lane 2), or Δ−67/−58 (lane 3). DNA was digested with *Xba*I and electrophoresed through 0.7% agarose gel. Following transfer to a nylon membrane, the DNA was hybridized with the tag-specific oligonucleotide M838. The plasmid-borne copies of the mini-exon gene are contained on a 6.5 kb fragment. (**C**) Analysis of nascent RNA from plasmid derived sequences. ^{32}P-labelled nascent RNA from *L.tarentolae* containing the indicated plasmids was hybridized to single-stranded probes that are specific for the tag, the NPTII gene, or to denatured pBluescript. Bound labelled RNA was quantified using a Molecular Dynamics PhosphorImager.

Discussion

We showed that the tandemly repeated genomic copies of the *L.tarentolae* mini-exon gene are transcribed individually and not as part of a polycistronic unit, and that they are transcribed by an RNA polymerase II-like enzyme that is resistant to tagetitoxin and sensitive to Sarkosyl. Using a stable transformation assay we showed that conserved exon sequences do not function as a promoter element. However, a highly conserved upstream element was essential for transcription of the episome-borne tagged mini-exon gene.

Identification of the RNA polymerase that transcribes the mini-exon gene of kinetoplastid protozoa has yielded conflicting data. In this study, use of the polymerase inhibitors tagetitoxin and Sarkosyl strongly indicates transcription by RNA polymerase II, as has been observed by others (Laird *et al.*, 1985; Ben Amar *et al.*, 1988; Rudenko *et al.*, 1992). In *T.brucei* discrepancies regarding the relative α-amanitin sensitivities of the mini-exon and tubulin genes may be due to the different techniques used in labelling the nascent RNA (Laird *et al.*, 1985; Ben Amar *et al.*, 1988; Ullu and Tschudi, 1990; Ullu *et al.*, 1993). It is also possible that the tubulin gene may not be the most appropriate template for comparative analysis: the elongation properties of RNA polymerase II transcription complexes differ between mRNA and snRNA templates (Spencer and Groudine, 1990). Thus, although α-amanitin acts by binding directly to the RNA polymerase, its effect upon elongation (Cochet-Meilhac and Chambon, 1974) may be observed due to the different transcription rates along the *T.brucei* mini-exon and tubulin genes.

The alternative hypothesis, that RNA polymerase III transcribes the mini-exon gene (Grondal *et al.*, 1989), can be addressed further. First, the proposed tRNA gene box A-like element does not function as a promoter element because it can be mutated without affecting transcription of the mini-exon gene (Figure 7B, lanes 7 and 8; Agami *et al.*, 1994). Secondly, the proposed tRNA box B-like element is unlikely to be functional due to a change in the invariant A residue at position 6 of the GTTCRAnnCC consensus (reviewed in Geiduschek and Tocchini-Valentini, 1988). Thirdly, the presence of at least four T residues flanking the 3′-end ot the transcribed region is not conclusive evidence for transcription by RNA polymerase III. For example, a run of four T residues is found at the end of the transcribed region of the *Ascaris* SL RNA gene, which is transcribed by RNA polymerase II (Nilsen *et al.*, 1989). Like the U-rich snRNA genes, many mini-exon genes contain internal runs of at least four T-residues that could act as transcription terminators for RNA polymerase III, a plausible reason why most snRNAs are synthesized by RNA polymerase II (Hernandez, 1992). Finally, the effects of 1,10-phenanthroline and manganese in the nuclear run-on assay appear to discriminate between the different RNA polymerases; however, the same conditions suggest that the VSG gene is transcribed by an α-amanitin-resistant RNA polymerase II. More recent data suggest that RNA polymerase I is the α-amanitin resistant enzyme that transcribes the VSG and procyclin/PARP genes (Clayton, 1992; Rudenko *et al.*, 1992; Chung *et al.*, 1993). Clearly, direct approaches are required for unequivocal identification of the RNA polymerase(s) that transcribe the mini-exon and VSG/procyclin genes.

The promoter element upstream of the *L.tarentolae* mini-exon gene is highly conserved in 14 other *Leishmania* species (Miller *et al.*, 1986; Fernandes *et al.*, 1994) and yields the consensus sequence TGACRCGYGG that shows a 5/10 identity with the conserved portion (TSACCnTnAS; Hernandez, 1992) of the human snRNA gene PSE. The PSE of vertebrate snRNA genes is typically 10–17 nt long and located at least 40 bp upstream of the gene. Using a similar transfection assay to that described here, Agami *et al.* (1994) have recently reported that mutation of the region −70 to −30 inactivated transcription of a tagged mini-exon gene in *L.amazonensis*. The −57/−30 region is 50% identical between *L.tarentolae* and *L.amazonensis*, thus the conserved 5′ motif and variable 3′ sequences are consistent in both structure and position with the PSE of snRNA genes (Hernandez, 1992).

Agami *et al.* (1994) identified a second promoter element for the *L.amazonensis* mini-exon gene located between positions −10 and +10. In *L.tarentolae*, mutation of the region −1 to +9 affects RNA stability to a greater extent than transcription of the gene, and mutation of the region −11 to −2 does not influence expression of the tagged medRNA. The latter discrepancy may result from differences in the nucleotide substitutions.

We have evidence that the tagged medRNA from the pLME12Δ30/39 plasmid is accurately *trans*-spliced onto a ubiquitin mRNA (J.Fleischmann, S.Wong, M.G.Elgort and D.A.Campbell, in preparation), so the choice of tag sites may also have yielded important information on sequences necessary for the *trans*-splicing process. Participation of the tagged medRNA in the *trans*-splicing reaction implies that the staggered primer extension products (Figure 5A) probably represent faithful modifications of the 'cap 4' structure, since this is essential for *trans*-splicing (Ullu and Tschudi, 1991; McNally and Agabian, 1992). The transfection assay described here will therefore allow us to characterize further the role of the upstream promoter in transcription initiation and perform a functional analysis of the *trans*-splicing pathway.

Materials and methods

Recombinant DNA clones

The modified transformation vector pXδA was derived from the *Leishmania* shuttle vector pX (LeBowitz *et al.*, 1990) by removal of the *Asp*718 site at position 4266. pX was cleaved with *Asp*718, end-filled with the Klenow fragment of DNA polymerase I and resealed with DNA ligase. pLME2, which contains one *L.tarentolae* mini-exon gene, was made by subcloning the 259 bp *Hae*III fragment (EMBL database accession number X73121) into the *Sma*I site of the pUC18 plasmid vector. pLME3, which contains the tagged *L.tarentolae* mini-exon gene, was created by the insertion of a 40 bp double-stranded linker M841/M838 (5′-AATTCGGAACACCCGGTCACGTGGCCTACACCTATG-AATT-3′/5′-AATTCCATAGGTGTAGGCCACGTGACCGGGTGTTCCGAATT-3′) into the unique *Bss*HII site, which had been end-filled by the Klenow fragment of DNA polymerase I, located at position 85 of the medRNA. The 0.3 kb *Sac*I−*Xba*I fragment of pLME3 was cloned into pBluescript (pLKC3) and the *Sma*I–*Xba*I fragment of pXδA (pLME12: the *Sac*I site was made blunt by the action of T4 DNA polymerase in the absence of dNTPs). M13-LME15 is the single-stranded version of pLME3 (non-sense strand cloned in M13 phage). M13-I/G#3 is a single-stranded clone representing the non-sense strand of a non-transcribed region. A single 204 bp non-transcribed region was PCR amplified using the specific oligonucleotides 5′-agactcggtaccGTGGCGGCGCAGTGT-3′ and 5′-tctagactcgagCCCATCATGTCCGGG-3′. Lower case letters represent restriction sites added to facilitate cloning.

Site-directed mutagenesis

Block substitution mutagenesis (Δ series) was performed with specific oligonucleotides on the pLME15 template using an Amersham mutagenesis kit. Oligonucleotides used for mutagenesis contained 10 or 15 bases of complementary sequence on either side of the 10 base block mutation, which introduced an *Xho*I site. Mutations were confirmed by DNA sequencing (Sequenase^{TM}; US Biochemicals). M13 clones containing the appropriate mutation were made double-stranded using the specific mutation oligonucleotide in a DNA polymerase reaction with the presence of DNA ligase. Released fragments (*Asp*718–*Xba*I) were cloned into the *Asp*718–*Xba*I vector fragment of pLME12. Double-stranded clones were re-screened by restriction analysis for the introduction of the unique *Xho*I site at the correct position.

Mutagenesis of the −67/−58 region was performed by PCR with the specific oligonucleotide 5′-TCGCCGCGTGCCGCAACTATGTGACC-CACACAGACGCCCC-3′ and the 'reverse' sequencing primer from the pLME3 template. The PCR products were cleaved with *Asp*718 and *Xba*I and inserted into the *Asp*718–*Xba*I vector fragment of pLME12.

Stable transformation of L.tarentolae

Drug sensitive *L.tarentolae* (UC strain) were colony purified on BHI medium (Difco) containing 10 μg/ml haemin and 1% agarose. Cultured cells were transfected as described by Kapler *et al.* (1990a) except that cells were selected in liquid culture. Approximately 5×10^7 cells and 20 μg of plasmid DNA were subjected to electroporation conditions of 2.5 kV, 25 μF and 400 W using 0.4 cm gap cuvettes (in the Bio-Rad Gene Pulser). Resistant cells were selected in 199 medium containing 25 μg/ml G418 (Geneticin, Gibco-BRL).

RNA analysis

Total RNA was isolated from cultured *L.tarentolae* by SDS lysis, proteinase K digestion, acidifed phenol extraction and ethanol precipitation. Ten micrograms of RNA were resolved by electrophoresis through an 8 M urea–8% polyacrylamide gel followed by electrophoretic transfer to nylon membrane (Nytran, Schleicher and Schuell) in 12 mM Tris pH 7.5, 6 mM NaOAc, 5 mM EDTA. The blot was rinsed once in 2× SSC and UV cross-linked (Stratagene, UV Stratalinker 1800). Prehybridization was carried out at 65°C for 1 h in 3× SSC, 0.5% SDS, 1× Denhardt's solution and 100 μg/ml sheared herring sperm DNA. Hybridization to probes was performed at 65°C for 16 h in the same buffer. Tag-specific probes were prepared by 5′-end labelling oligonucleotide M838 with [γ-^{32}P]ATP and T4 polynucleotide kinase. The mini-exon probe was labelled by random hexamer priming (Feinberg and Vogelstein, 1983). Filters were washed at 65°C in 0.1× SSC and exposed to X-ray film. Quantification of autoradiographic bands was performed using a SciScan 2000 scanning densitometer. Primer extension analysis was performed as described by Smale and Baltimore (1989) using oligonucleotide M838.

RNase protection assays were performed on crude cell lysates using the Lysate Ribonuclease Protection kit (United States Biochemical). Riboprobe was produced by T7 RNA polymerase using *Asp*718-linearized pLKC3 as a template in the presence of [α-^{32}P]GTP (3000 Ci/mmol, ICN). A cell lysate of 10^7 cells/ml was prepared in 4 M guanidine thiocyanate, 25 mM sodium citrate and 0.5% Sarkosyl. Hybridization occurred at 55°C for 14 h followed by RNase A (10 U) and RNase T1 (0.5 U) digestion at 37°C for 30 min. Samples were separated through an 8 M urea–5% polyacrylamide gel and dried prior to exposure to X-ray film.

Nuclear run-on assays

Nuclear run-on assays were performed essentially as described by Ullu and Tschudi (1990). The RNA was elongated in $\sim10^8$ lysolethicin permeabilized cells in the presence of 200 μCi of [α-^{32}P]GTP (3000 Ci/mmol, ICN) for 20 min at 30°C. The nascent RNA was isolated by SDS lysis and acidified phenol extraction. Cells receiving polymerase inhibitor treatment were treated with Sarkosyl (*N*-lauroylsarcosine, Sigma) or tagetitoxin (Tagetin, Epicentre Biotechnologies) for 5 min at 4°C prior to RNA elongation.

Filters were prepared by applying 2 μg of M13 DNA per gene probe to Nytran membranes (S&S) with a slot blot apparatus, followed by UV cross-linking. DNA probes were as follows: oligonucleotide M838; a 1.2 kb *Hin*dIII–*Sal*I fragment of the neomycin phosphotransferase (NPTII) gene cloned in M13; alkali-denatured pBluescript; a 1.3 kb *Eco*RI fragment of the *L.tarentolae* rRNA gene cloned in M13; a 0.6 kb *Bam*HI–*Xho*I fragment of the *L.enriettii* β-tubulin gene (Landfear *et al.*, 1986) in M13; and a 0.43 kb *Hae*III fragment containing the *T.brucei* glutamine, valine and lysine tRNAs in M13 (Campbell, 1989). Hybridization of labelled RNA to blots was performed in 3× SSC, 0.5% SDS, 1× Denhardt's solution, 100 μg/ml sheared herring sperm DNA at 65°C for 16 h. Filters were washed in 0.1× SSC at 65°C and exposed to X-ray film.

Capping/RNase H

Radiolabelled RNA was produced by enzymatic addition of a 5′ [α-^{32}P]GTP. RNA was decapped by $NaIO_4$ oxidation and β-elimination essentially as described by Rose and Lodish (1976). Briefly, 10 μg of total RNA were incubated in 10 mM $NaIO_4$ at room temperature for 1 h. The 5′-cap was removed by 0.165 M aniline treatment. The enzymatic addition of the RNA cap was previously described elsewhere (Zwierzynski and Buck, 1990). Decapped and mock treated RNA were capped with vaccinia virus guanylyl transferase (Gibco-BRL) in the following reaction: 50 mM Tris–HCl (pH 8.0), 1.25 mM $MgCl_2$, 6 mM KCl, 2.5 mM DTT, 2 U guanylyl transferase and 100 μCi [α-^{32}P]GTP (3000 Ci/mmol, ICN). Approximately 2 μg of 5′ radiolabelled RNA were allowed to hybridize with 2 μg of oligo GE28s (5′-TTCCTCGAGGCTG-AA) at 32°C. RNA–DNA hybrids were digested by RNase H (Gibco-BRL) in a reaction containing 40 mM Tris–HCl (pH 8.0), 100 mM NaCl, 4 mM $MgCl_2$, 1 mM DTT and 3.5 U RNase H. The reaction was performed at 32°C for 1 h. The digestion products were resolved through an 8 M urea–6% polyacrylamide gel, which was then dried and exposed to film.

Acknowledgements

We are deeply indebted to our M&I colleagues, whose generosity and support have contributed significantly to this work. We also wish to thank Larry Feldman for the tag oligonucleotides and use of the electroporator; Steve Beverley for the gift of pX; Scott Landfear for pLEb3; Sandy Wong, Steve Smale and Mike Carey for critical comments on the manuscript; John Boothroyd for suggesting the use of an RNase H assay. This work was supported by a Burroughs Wellcome New Investigator Award in Molecular Parasitology (D.A.C.) and by the Academic Senate of the Los Angeles Division of the University of California. R.M.S. is supported by the USPHS National Research Service Award GM-07104.

References

Agami,R., Aly,R., Segula,H. and Shapira,M. (1994) *Nucleic Acids Res.*, **22**, 1959–1965.

Bakalara,N., Kendall,G., Michels,P.A.M. and Opperdoes,F.R. (1992) *EMBO J.*, **11**, 3808.

Ben Amar,M.F., Pays,A., Tebabi,P., Dero,B., Seebeck,T., Steinert,M. and Pays,E. (1988) *Mol. Cell. Biol.*, **8**, 2166–2176.

Ben Amar,M.F., Jefferies,D., Pays,A., Bakalara,N., Kendall,G. and Pays,E. (1991) *Nucleic Acids Res.*, **19**, 5857–5862.

Bernards,A., Van der Ploeg,L.H.T., Frasch,A.C.C., Borst,P., Boothroyd,J.C., Coleman,S. and Cross,G.A.M. (1981) *Cell*, **27**, 497–505.

Boothroyd,J.C. and Cross,G.A.M. (1982) *Gene*, **20**, 281–289.

Brown,S.D., Huang,J. and Van der Ploeg,L.H.T. (1992) *Mol. Cell. Biol.*, **12**, 2644–2652.

Campbell,D.A. (1989) *Nucleic Acids Res.*, **22**, 9479.

Campbell,D.A., Suyama,Y. and Simpson,L. (1989) *Mol. Biochem. Parasitol.*, **37**, 257–262.

Chung,H.-M.M., Lee,M.G.-S., Dietrich,P., Huang,J. and Van der Ploeg,L.H.T. (1993) *Mol. Cell. Biol.*, **13**, 3734–3743.

Clayton,C.E. (1992) *Prog. Nucleic Acids Res.*, **43**, 37–66.

Cochet-Meilhac,M. and Chambon,P. (1974) *Biochim. Biophys. Acta*, **353**, 160–184.

Cook,G.A. and Donelson,J.E. (1987) *Mol. Biochem. Parasitol.*, **25**, 113–122.

Cordingley,J.S. (1985) *Mol. Biochem. Parasitol.*, **17**, 321–330.

Cornelissen,A.W.C.A., Verspieren,M.P., Toulme,J.-J., Swinkels,B.W. and Borst,P. (1986) *Nucleic Acids Res.*, **14**, 5605–5614.

Curotto de Lafaille,M.A., Laban,A. and Wirth,D.F. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 2703–2707.

Dietrich,P., Soares,M.B., Affonso,M.H.T. and Floeter-Winter,L.M. (1993) *Gene*, **125**, 103–107.

Fantoni,A., Dare,A.O. and Tschudi,C. (1994) *Mol. Cell. Biol.*, **14**, 2021–2028.

Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.

Fernandes,O., Murthy,V.K., Kurath,U., Degrave,W.M. and Campbell,D.A. (1994) *Mol. Biochem. Parasitol.*, **66**, 261–271.

Fleischmann,J. and Campbell,D.A. (1994) *Gene*, **144**, 45–51.

Freistadt,M.S., Cross,G.A.M., Branch,A. and Robertson,H.D. (1987) *Nucleic Acids Res.*, **15**, 9861–9879.

Geiduschek,E.P. and Tocchini-Valentini,G.P. (1988) *Annu. Rev. Biochem.*, **57**, 873–914.
Gibson,W.C., Swinkels,B.W. and Borst,P. (1988) *J. Mol. Biol.*, **201**, 315–325.
Grondal,E.M.J., Evers,R., Kosubek,K. and Cornelissen,A.W.C.A. (1989) *EMBO J.*, **8**, 3383–3389.
Grondal,E.M.J., Evers,R. and Cornelissen,A.W.C.A. (1990) *Nucleic Acids Res.*, **18**, 1333–1338.
Hancock,K., LeBlanc,A.J., Donze,D. and Hajduk,S.L. (1992) *J. Biol. Chem.*, **267**, 23963–23971.
Hannon,G.J., Maroney,P.A., Ayers,D.G., Shambaugh,J.D. and Nilsen,T.W. (1990) *EMBO J.*, **9**, 1915–1921.
Hernandez,N. (1992) In Yamamoto,K. and McKnight,S.L. (eds), *Transcriptional Regulation.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 281–313.
Johnson,P.J., Kooter,J.M. and Borst,P. (1987) *Cell*, **51**, 273–281.
Kapler,G.M., Coburn,C.M. and Beverley,S.M. (1990a) *Mol. Cell. Biol.*, **10**, 1084–1094.
Kapler,G.M., Zhang,K. and Beverley,S.M. (1990b) *Nucleic Acids Res.*, **18**, 6399–6408.
Kooter,J.M. and Borst,P. (1984) *Nucleic Acids Res.*, **12**, 9457–9472.
Kooter,J.M., De Lange,T. and Borst,P. (1984) *EMBO J.*, **3**, 2387–2392.
Köck,J. and Cornelissen,A.W.C.A. (1991) *Mol. Microbiol.*, **5**, 835–842.
König,E., Delius,H., Carrington,M., Williams,R.O. and Roditi,I. (1989) *Nucleic Acids Res.*, **17**, 8727–8739.
Laird,P.W., Kooter,J.M., Loosbroek,N. and Borst,P. (1985) *Nucleic Acids Res.*, **13**, 4253–4266.
Landfear,S.M., Miller,S.I. and Wirth,D.F. (1986) *Mol. Biochem. Parasitol.*, **21**, 235–245.
LeBowitz,J.H., Coburn,C.M., McMahon-Pratt,D. and Beverley,S.M. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 9736–9740.
Lenardo,M.J., Dorfman,D.M. and Donelson,J.E. (1985) *Mol. Cell. Biol.*, **9**, 2487–2490.
Martínez-Calvillo,S. and Hernández,R. (1994) *Gene*, **142**, 243–247.
Mattaj,I.W., Lienhard,S., Jiricny,J. and De Robertis,E.M. (1985) *Nature*, **316**, 163–167.
McNally,K.P. and Agabian,N. (1992) *Mol. Cell. Biol.*, **12**, 4844–4851.
Michaeli,S., Podell,D., Agabian,N. and Ullu,E. (1992) *Mol. Biochem. Parasitol.*, **51**, 55–64.
Miller,S.I., Landfear,S.M. and Wirth,D.F. (1986) *Nucleic Acids Res.*, **14**, 7341–7360.
Muhich,M.L., Hughes,D.E., Simpson,A.M. and Simpson,L. (1987) *Nucleic Acids Res.*, **15**, 3141–3153.
Murphy,W.J., Watkins,K.P. and Agabian,N. (1986) *Cell*, **47**, 517–525.
Nilsen,T.W., Shambaugh,J., Denker,J., Chubb,G., Faser,C., Putnam,L. and Bennett,K. (1989) *Mol. Cell. Biol.*, **9**, 3543–3547.
Pays,E. (1993) In Broda,P.M., Oliver,S.G. and Sims,P. (eds), *The Microbial Genome.* Cambridge University Press, London, pp. 99–132.
Perry,K.L., Watkins,K.P. and Agabian,N. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 8190–8194.
Rose,J.K. and Lodish,H.F. (1976) *Nature*, **262**, 32–37.
Rudenko,G., Bishop,D., Gottesdiener,K. and Van der Ploeg,L.H.T. (1989) *EMBO J.*, **13**, 4259–4263.
Rudenko,G., Chung,H.M., Pham,V.P. and Van der Ploeg,L.H.T. (1991) *EMBO J.*, **10**, 3387–3397.
Rudenko,G., Lee,M.G.-S. and Van der Ploeg,L.H.T. (1992) *Nucleic Acids Res.*, **20**, 303–306.
Schnare,M.N., Spencer,D.F. and Gray,M.W. (1983) *Can. J. Biochem. Cell Biol.*, **61**, 38–45.
Sherman,D.R., Janz,L., Hug,M. and Clayton,C. (1991) *EMBO J.*, **10**, 3379–3386.
Smale,S.T. and Baltimore,D. (1989) *Cell*, **57**, 103–113.
Spencer,C.A. and Groudine,M. (1990) *Oncogene*, **5**, 777–785.
Steinberg,T.H., Mathews,D.E., Durbin,R.D. and Burgess,R.R. (1990) *J. Biol. Chem.*, **265**, 499–505.
Sutton,R.E. and Boothroyd,J.C. (1986) *Cell*, **47**, 527–535.
Ullu,E. and Tschudi,C. (1990) *Nucleic Acids Res.*, **18**, 3319–3326.
Ullu,E. and Tschudi,C. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 10074–10078.
Ullu,E., Matthews,K.R. and Tschudi,C. (1993) *Mol. Cell. Biol.*, **13**, 720–725.
Van der Ploeg,L.H.T., Liu,A.Y.C., Michels,P.A.M., De Lange,T., Borst,P., Majumder,H.K., Weber,H., Venneman,G.H. and Van Boom,J. (1982) *Nucleic Acids Res.*, **10**, 3591–3604.
Walder,J.A., Eder,P.S., Engman,D.M., Brentano,S.T., Walder,R.Y., Knutzon,D.S., Dorfman,D.M. and Donelson,J.E. (1986) *Science*, **233**, 569–571.
White,T.C., Rudenko,G. and Borst,P. (1986) *Nucleic Acids Res.*, **14**, 9471–9489.
Wong,S., Morales,T.H., Neigel,J.E. and Campbell,D.A. (1993) *Mol. Cell. Biol.*, **13**, 207–216.
Zomerdijk,J.C.B.M., Kieft,R., Shiels,P.G. and Borst,P. (1991) *Nucleic Acids Res.*, **19**, 5153–5158.
Zwierzynski,T.A. and Buck,G.A. (1990) *Nucleic Acids Res.*, **18**, 4197–4206.

Received on January 5, 1994; revised on August 11, 1994