MOLECULAR AND CELLULAR BIOLOGY, June 1991, p. 3180–3190
0270-7306/91/063180-11$02.00/0

Vol. 11, No. 6

Maturation of Polycistronic Pre-mRNA in *Trypanosoma brucei*: Analysis of *trans* Splicing and Poly(A) Addition at Nascent RNA Transcripts from the *hsp70* Locus

JIN HUANG AND LEX H. T. VAN DER PLOEG*

Department of Genetics and Development, College of Physicians and Surgeons, Columbia University, New York, New York 10032

Received 6 February 1991/Accepted 29 March 1991

Numerous protein-coding genes of the protozoan *Trypanosoma brucei* are arranged in tandem arrays that are transcribed polycistronically. The pre-mRNA transcripts are processed by *trans* splicing, leading to the addition of a capped 39-nucleotide (nt) miniexon and by poly(A) addition. We wished to determine the order of the RNA processing events at the *hsp70* locus and address the potential occurrence of cotranscriptional RNA processing. We determined the rate of transcriptional elongation at the *hsp70* locus in isolated nuclei, which measured between 20 and 40 nt/min. This low rate of RNA chain elongation allowed us to label the 3′ end of *hsp70* nascent RNA with a short (about 180-nt) ^{32}P tail. The structure of the labeled nascent *hsp70* RNA could then be analyzed by RNase T_1 and RNase T_1/RNase A mapping. We show that the *trans* splicing of *hsp70* pre-mRNA did not occur immediately after the synthesis of the 3′ splice acceptor site, and nascent RNA molecules that contained about 550 nt of RNA beyond the 3′ splice acceptor site still had not acquired a miniexon. In contrast, nascent RNA with a 5′ end that mapped to the polyadenylation site of the *hsp70* genes could be detected, indicating that maturation of the pre-mRNA in trypanosomes involves a rapid cleavage of the nascent *hsp70* RNA (within seconds after synthesis of the site) for poly(A) addition. Our data suggest that polycistronic pre-mRNA is unlikely to be synthesized in toto and rather appears to be processed cotranscriptionally by cleavage for poly(A) addition.

Most protein-coding genes in eukaryotes are interrupted by intervening sequences (introns) which are spliced from the precursor mRNA to generate mature mRNA by *cis* splicing (removal of introns from colinear transcripts; 18, 43, 48). In kinetoplastida and several nematode species, the exons of a single mRNA can be encoded by different genes located on separate chromosomes. In these cases, the exon-joining reaction is through *trans* splicing of a separately transcribed 5′-capped miniexon donor RNA (medRNA or spliced leader RNA) and an mRNA precursor (3, 4, 24, 57). In nematodes, both *cis* splicing and *trans* splicing play roles in mRNA maturation (24), while in trypanosomes and related kinetoplastidae, *trans* splicing is the only route of pre-mRNA maturation. Every mRNA of *Trypanosoma brucei* is presumed to be processed by *trans* splicing (9, 59), in which a 39-nucleotide (nt) miniexon, derived from the 5′ end of the medRNA, is spliced onto the 5′ end of each mRNA molecule. Conventional introns, i.e., those that disrupt protein-coding sequences, have not been found in trypanosomes.

The mechanism of *trans* splicing bears a striking resemblance to its *cis* counterpart in that the cleavage signals at the 5′ and 3′ splice junctions are conserved on the medRNA and pre-mRNAs, respectively. The analogy extends further with the identification of Y-shaped splicing intermediates involving 2′-5′ phosphodiester bonds that connect the introns of medRNA and the pre-mRNA (27, 40, 45, 51). The Y-shaped splicing intermediates are the structural equivalent of the lariat structures generated in *cis* splicing. Analogs of U2, U4, and U6 small nuclear RNA (snRNA) involved in *cis* splicing were also found in trypanosomes (37, 53), and they have recently been demonstrated to be essential for *trans* splicing (55). The medRNA itself has been proposed to play a catalytic role in *trans* splicing since it may be functionally equivalent to the U1 snRNA in *cis* splicing (7, 58). This notion is further supported by studies in which the spliced leader RNA of *Caenorhabditis elegans* was shown to be able to substitute for the human U1 snRNA, suggesting that the medRNA can function both as a substrate and as a catalyst in the *trans*-splicing reaction (6).

RNA splicing is generally considered to be a posttranscriptional event preceded by polyadenylation (25, 41, 46, 52). However, electron microscopic analyses revealed a different picture of RNA processing in which spliceosome assembly and splicing were shown to occur on growing, nascent RNA chains (1, 42). Nascent RNA processing suggests a role for the transcription complex or the nascent RNA chain in RNA maturation. To understand the control of RNA processing, knowledge of the timing of spliceosome assembly, the in vivo location of spliceosomes, differential intron removal in alternative splicing, and the parameters that determine the efficiency of splicing is essential.

In trypanosomes, many protein-coding genes are transcribed polycistronically from putative single upstream promoters, generating polycistronic pre-mRNA transcripts (13, 21, 38, 44, 47, 54, 62) that are processed by *trans* splicing with the medRNA. However, the precursor-product relationship of polycistronic pre-mRNA and the mature mRNA has not been clearly defined, and RNA from polycistronically transcribed protein-coding genes may be processed posttranscriptionally or cotranscriptionally. In either case, the processing is believed to be rapid, since the medRNA has a short half-life (about 4 min; 26, 27, 56).

It is obvious that cotranscriptional and/or posttranscriptional control is required to adjust the mRNA abundance of

* Corresponding author.

constitutively transcribed genes. Hence, the timing of the RNA processing events becomes essential in understanding the control of mRNA abundance. In this work, we studied RNA processing events of the RNA polymerase II generated heat shock protein of 70 kDa (*hsp70*) nascent RNA (14, 30). We show that miniexon addition at *hsp70* genes does not occur immediately after synthesis of the 3′ splice acceptor site. Moreover, synthesis of a further 550 nt of nascent RNA, beyond the 3′ splice acceptor site, occurred without miniexon addition. In contrast, evidence for cotranscriptional RNA processing for poly(A) addition was seen. This observation provides an explanation for the rarity of polycistronic *hsp70* pre-mRNA in total nascent RNA in trypanosomes.

MATERIALS AND METHODS

Construction of *hsp70* cDNA plasmids. A 209-nt cDNA construct, comprising 20 nt of the miniexon and 189 nt of the *hsp70* coding region, was made by using a 25-mer synthetic oligonucleotide (5′-CCGTGAAGGCAACGTAGGACGGC GT), complementary to the 5′ end of the *hsp70* coding sequence, to generate a short miniexon containing cDNA from poly(A)$^+$ trypanosome RNA. The antisense cDNA was then used as a template for the polymerase chain reaction with the same 25-mer *hsp70* oligonucleotide and a synthetic 20-nt miniexon sense-strand oligomer, which is identical to the last 20 nt of the miniexon (5′-ACAGTTTCTGTACTAT ATTG). The polymerase chain reaction-amplified 209-nt fragment was cloned into the *Sma*I site of plasmid Sp 64 (35). The DNA nucleotide sequence of the construct was confirmed by double-strand DNA sequencing. A 400-nt cDNA construct was made by extending the 209-nt construct by inserting a *Bst*EII-*Sac*II restriction enzyme fragment (the *Sac*II site was blunted by S1 nuclease) of the *hsp70* coding sequence (derived from the genomic *hsp70* clone h10; 14) into the *Bst*EII-*Hin*cII-digested 209-nt cDNA plasmid. The nucleotide sequence of this clone was determined. It contained 380 bp of the *hsp70* coding sequence, located downstream of the 3′ splice acceptor site, and 20 bp of the miniexon sequence. A 570-nt cDNA construct was also made by extending the *hsp70* coding region of the 209-nt plasmid construct. A *Bst*EII-*Eco*RI restriction enzyme fragment, shortened by Bal 31 treatment of the *Eco*RI restriction enzyme site, was ligated into the *Bst*EII-*Hin*cII-digested 209-nt Sp 64 cDNA clone. The length of the insert of this construct measured 570 bp, including 20 bp of the miniexon sequence, as confirmed by DNA nucleotide sequencing. DNA construct pHS-IG1 was generated by inserting a *Pst*I-*Bst*EII fragment (partially digested with *Pst*I; the *Bst*EII site was Klenow treated) containing the *hsp70* intergenic region from clone h10 into Sp 65 (*Pst*I-*Sma*I digested). An *hsp70* intergenic-region *Pst*I-*Eco*RI fragment (partially digested with *Pst*I) was trimmed by Bal 31 from the *Eco*RI site to nucleotide position 256 downstream of the 3′ splice acceptor site and introduced into *Pst*I-*Sma*I-digested Sp 65, generating pHS-IG2.

Preparation of nascent RNA from isolated trypanosome nuclei. Both bloodstream-form and insect-form *T. brucei* variant 118 clone 1 (29) were used to prepare nuclei according to a previously described procedure (22). Cells are passed through the cell disrupter and diluted in nuclei preparation buffer at 0°C within seconds of cell lysis, ensuring an immediate block in nascent RNA chain elongation. Nascent RNA was prepared after 6 min of incubation in the presence of [α-^{32}P]UTP or -GTP in a typical run-on transcription assay (23). The reaction volume was 200 μl containing 0.1 M Tris-HCl (pH 7.9), 50 mM NaCl, 60 mM KCl, 2 mM $MgCl_2$, 4 mM $MnCl_2$, 0.15 mM spermine, 0.5 mM spermidine, 25% glycerol, 2 mM dithiothreitol, 2 mM each ATP, CTP, and GTP, 10 μM UTP, 500 μCi of [^{32}P]UTP (3,000 Ci/mmol; NEN), and 10^9 nuclei. The resulting nascent RNA provides sufficient signal for 16 to 20 RNase protection assays. The reaction was terminated by adding an equal volume of 20 mM Tris-HCl (pH 7.5)–10 mM EDTA–1% sodium dodecyl sulfate–1 mg of proteinase K per ml, and the mixture was incubated for 20 min at 50°C. DNA in the nuclear lysate was sheared by passing the mixture 10 times through a 21-gauge needle. The lysate was extensively extracted with phenol-chloroform-isoamyl alcohol, passed through a Sephadex G-50 column, and again extracted with the same reagents. After precipitation in ethanol, the pellets were subject to RNase-free DNase I (Boehringer Mannheim) treatment in 10 mM Tris-HCl (pH 7.5)–5 mM $MgCl_2$–0.1 mM EDTA at 37°C for 15 min. The RNA was then extracted with phenol-chloroform and subsequently precipitated in ethanol. In some experiments, nascent RNA was size fractionated on 4% polyacrylamide gels and eluted as described previously (16). In other experiments, hybridization selection was carried out to improve the signal-to-noise ratio (27), in which the same antisense riboprobe was used for selection as for the RNase protection (nascent *hsp70* RNA was selected with clone pHS-IG1).

Methods to purify large nascent RNA and study its potential miniexon addition involved biotinylation of nascent RNA, labeled with ^{6}N-aminohexyl-ATP (Bethesda Research Laboratories) or allylamine-UTP (Enzo Biochem), followed by fractionation of the nascent (derivatized) RNA from nonbiotinylated (steady-state) RNA. However, the ^{6}N-aminohexyl-ATP-mediated biotinylation had a poor labeling specificity due to nonspecific labeling of RNA with ATP, presumably by a poly(A) polymerase, in addition to the low incorporation efficiency of this ATP derivative. Similarly, use of allylamine-UTP also led to nonspecific labeling, making our analysis of larger nascent RNA molecules impossible (20a).

We concluded that *trans* splicing did not occur at nascent *hsp70* RNA extending up to 550 nt beyond the 3′ splice acceptor site. The accuracy of this measurement has some limitations, since the relative contribution of nascent RNA from RNA polymerases that are further away from the 3′ splice acceptor site will become progressively less significant: these RNAs will have a lower specific activity, and the longest nascent RNAs analyzed will represent a fraction of the total RNA analyzed. However, even RNA polymerases that are for instance between 500 to 550 nt beyond the 3′ splice acceptor site will contribute as much as 7.5% of the intensity of the total signal. Therefore, since the protected fragments revealed unspliced RNA only, even the nascent RNAs from RNA polymerases that were located far beyond the 3′ splice acceptor site (for instance, between nt 500 and 550) must still represent unspliced molecules.

RNase protection analyses. Antisense riboprobes were synthesized as ^{32}P-labeled or unlabeled RNA with SP6 polymerase as described by Melton et al. (35). Unlabeled riboprobes were used to analyze ^{32}P-labeled nascent RNA, whereas the labeled SP6 RNA probes were used for the protection of total steady-state RNA. In general, an antisense riboprobe was coprecipitated with the target RNA molecules, and the pellets were dissolved in 20 μl of hybridization solution [10 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 6.9), 0.4 M NaCl, 1 mM EDTA, 80%

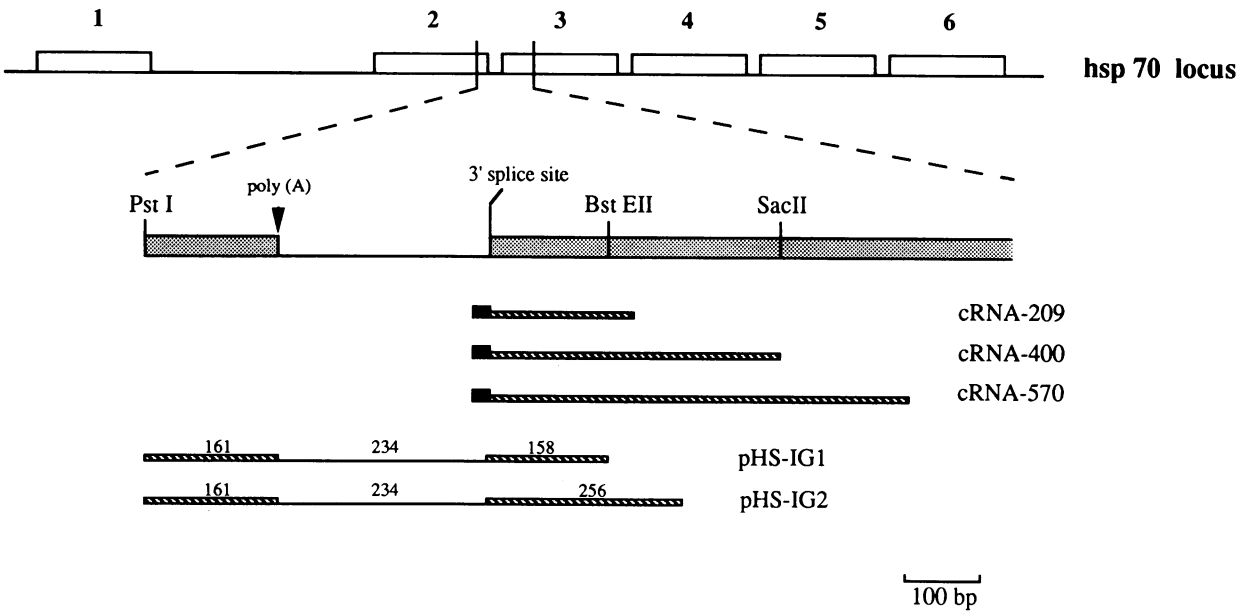

FIG. 1. Schematic representation of the *hsp70* locus of *T. brucei* and the synthetic cRNA probes. The polycistronic *hsp70* array (open boxes numbered 1 to 6) in *T. brucei* consists of five identical *hsp70* genes (genes 2 to 6) and one diverged cognate *hsp70* gene (gene 1). An enlarged region of one of the repeated *hsp70* genes is shown, containing, from 5' to 3', the 3' end of the preceding *hsp70* gene (stippled box), the intergenic region (thin line), and the 5' end of the next *hsp70* gene. The arrowhead points to the site of poly(A) addition of *hsp70* mRNA. The sizes and locations of the cRNA probes are shown below the physical map. All three of the antisense cRNA probes have the last 20 nt of miniexon complementary sequence (black boxes), and they differ from one another by the length of the *hsp70* coding exon downstream of the 3' splice acceptor site (hatched lines). pHS-IG1 and pHS-IG2 are antisense probes containing the intergenic region sequence (thin lines) flanked by sequences of the *hsp70* coding exons (hatched lines).

formamide]. Denaturation was at 85°C for 15 min, and annealing was at 45°C for 10 to 12 h. RNase T_1 digestion was performed in 10 mM Tris-HCl (pH 7.5)–0.3 M NaCl, with 200 U of the enzyme (Calbiochem) per ml, at 30°C for 1 h. The conditions for RNase T_1/RNase A incubation were similar except that 5 mM EDTA and 4 μg of RNase A (Sigma) per ml were included.

RNase H treatment of RNA. RNase H digestion of RNA molecules was carried out in the presence of complementary oligonucleotides. A 22-nt antisense miniexon oligomer (5'-CAATATAGTACAGAAACTGTTC) was used to remove miniexons from the 5' ends of steady-state RNA, and a 25-mer complementary to the 5' end of the *hsp70* coding sequence (5'-CCGTGAAGGCAACGTAGGACGGC GT) was used to mediate cleavage of *hsp70* RNA molecules in their coding sequence. Typically, a 15- to 20-fold excess of oligonucleotide was coprecipitated with nascent RNA in ethanol, and the pellets were dissolved in 20 mM Tris-HCl (pH 7.5)–10 mM $MgCl_2$–100 mM KCl–0.1 mM dithiothreitol–5% sucrose. After a 30-min incubation at 37°C, 2 to 3 U of RNase H (Bethesda Research Laboratories) was added, and digestion was allowed to proceed for 1 h at 37°C. Oligonucleotides were removed by a 15-min treatment with RNase-free DNase I. Upon extraction with phenol-chloroform-isoamyl alcohol, the nascent RNA was precipitated in ethanol and analyzed in the RNase T_1 protection assay.

Determination of the rate of elongation. The overall rate of nascent RNA chain elongation by RNA polymerases I, II, and III in isolated nuclei was measured as described by Marzluff et al. (34). Nuclei were first incubated under the standard run-on conditions for 3 min without labeled nucleoside triphosphate; 45 μmol of [2,8-^{3}H]GTP was then introduced. The reaction was allowed to continue for another 3 min. RNA was recovered as described above and size fractionated on a 10 to 30% sucrose gradient (8). RNA molecules larger than 16S were combined and subjected to RNase T_1/T_2 digestion. The resulting nucleotides and nucleosides were separated on polyethyleneimine-cellulose plates (EM Science) with 200 mM LiCl. The GMP and guanosine spots were recovered from the plate and quantitated by liquid scintillation counting. The ratio of GMP and guanosine was thus determined, giving an estimate of the rate of RNA chain elongation. The rate of elongation at the *T. brucei* *hsp70* genes was determined with nascent RNA isolated from identical amounts of nuclei for each of the time points. Nuclear run-on and RNA size fractionation were performed as described above.

RESULTS

Lack of *trans* splicing on nascent *hsp70* RNA. Large polycistronic pre-mRNAs may be synthesized in toto before they are processed into multiple mature mRNA molecules via *trans* splicing and/or polyadenylation. Alternatively, *trans* splicing and polyadenylation may occur cotranscriptionally. To elucidate the mechanisms of mRNA maturation, we analyzed the processing events on nascent RNA derived from the *hsp70* locus. The presence of five identical *hsp70* genes in a tandem array (14, 30) increased the signal from nascent RNA, which was essential since (i) the *hsp70* locus is transcribed at a low efficiency and (ii) our attempts to increase the *hsp70* nascent RNA signals by a short heat shock failed (28, 31a).

We used RNase T_1 and RNase T_1/RNase A to map ^{32}P-labeled *hsp70* nascent RNA (RNA labeled as a result of RNA polymerase II-mediated chain elongation; 36, 50; see

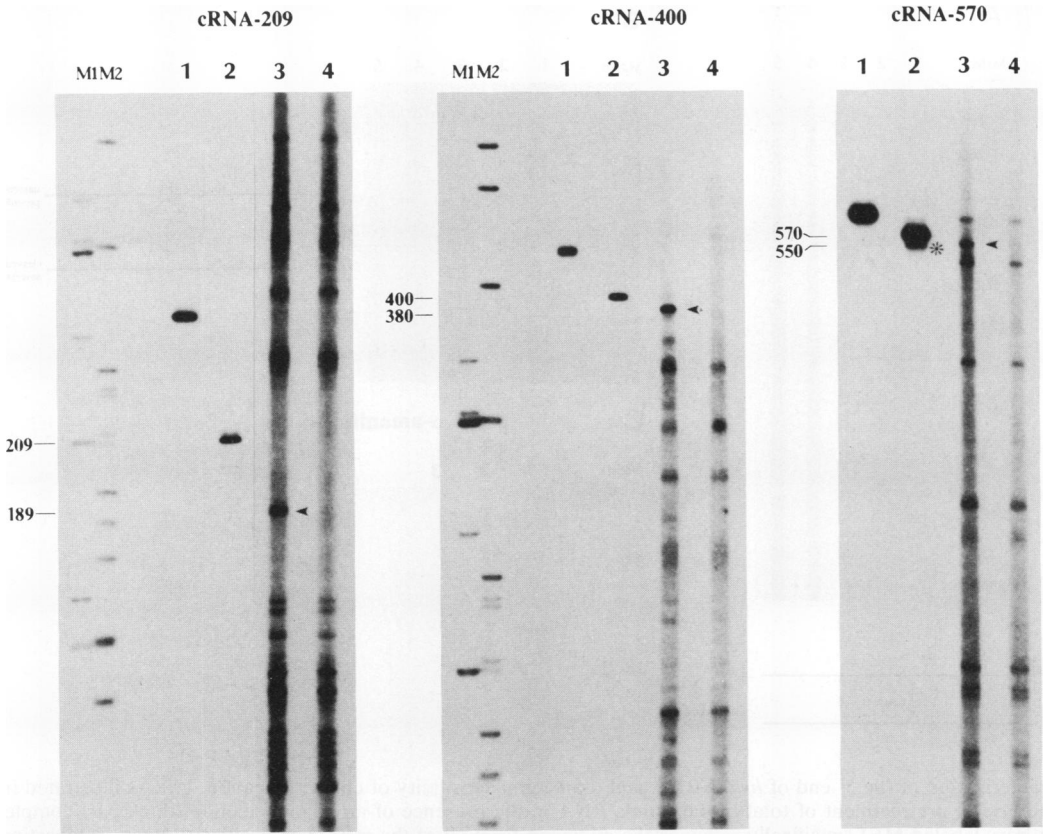

FIG. 2. RNase T_1 protection analyses of miniexon addition on *hsp70* nascent RNA, using antisense RNA probes cRNA-209, cRNA-400, and cRNA-570. Lanes: 1, size of undigested, labeled antisense cRNA probes; 2, RNase T_1 protection pattern of steady-state RNA annealed with a corresponding ^{32}P-labeled antisense cRNA probe; 3, result of RNase mapping of ^{32}P-labeled nascent RNA protected by unlabeled antisense cRNA probes; 4, control lane of RNase T_1 digestion of nascent RNA without addition of the unlabeled antisense cRNA. Arrowheads highlight RNase T_1-resistant, *hsp70*-specific nascent RNA fragments (lanes 3). Asterisks indicate protected fragments derived from the *hsp70* mRNA that uses a cryptic splice site for miniexon addition (visible only in long exposures [see also Fig. 3]). The sizes (in nucleotides) of spliced *hsp70* fragments from steady-state RNA (lanes 2) and unspliced RNase T_1-resistant *hsp70* bands from nascent RNA are indicated to the left of each panel. Size standards (lanes M1 and M2) are made from *Hae*III-digested M13mp18 and from *Hpa*II-digested pBR322, respectively (the same markers were used in all experiments). The autoradiographic intensity of the nonspecifically protected bands and their protection pattern are dependent on the RNase T_1 concentration and the temperature of the incubation. Fewer background bands occur with higher RNase T_1 concentrations or addition of RNase A (see also Fig. 4; data not shown). As expected, the protection pattern of these nonspecific bands can also be altered by size fractionation of nascent RNA (see Fig. 5). Variations in the pattern of the background bands between different experiments result from minor differences in the RNase T_1 digestion and minor differences between individual batches of size-fractionated nascent RNA. The low level of transcription at the *hsp70* genes makes the background bands prominent relative to the specific signal. In comparison with analysis of genes that are transcribed at a high rate, where the specific signal is more abundant (VSG; 15, 62), the analysis of *hsp70* gene transcription is therefore technically more demanding.

below for controls). Nascent RNA labeled with either [α-^{32}P]UTP or [α-^{32}P]GTP (data not shown) can thus be discriminated from mature mRNA, which will not be labeled. A typical labeling period of nascent RNA in our study is between 5 and 6 min, during which time about 180 nt of newly synthesized RNA will be added to the growing RNA chains by RNA polymerase II (see below). To distinguish *hsp70* nascent RNA from the heterogeneous pool of nascent RNA transcripts, we made antisense riboprobes (complementary RNAs or cRNAs) from a series of *hsp70* cDNA constructs, all having the last 20 nt of the miniexon spliced onto the major 3′ splice acceptor site and extending for various distances into the coding sequences located downstream of the 3′ splice acceptor site (Fig. 1). Hybrids formed between ^{32}P-labeled nascent RNA (size fractionated to be >350 nt; see below) and unlabeled antisense cRNA were then subjected to RNase T_1 or RNase T_1/RNase A digestion.

We studied nascent *hsp70* RNA, using unlabeled antisense *hsp70* cRNAs of 209 nt (cRNA-209), 400 nt (cRNA-400), and 570 nt (cRNA-570) as probes (Fig. 1). The results obtained with these different cRNA probes were compared to ensure the correct mapping of *hsp70*-derived nascent RNA. Two different types of protected fragments may result from the RNase T_1 digestion: (i) with unspliced nascent RNA, the RNA-RNA hybrids will be cleaved by RNase T_1 at a G residue at position -1 of the splice acceptor site in the pre-mRNA; and (ii) with spliced nascent RNA, a protected fragment that is 20 nt longer than unspliced RNA will result, because RNase T_1 will now cleave at a G residue at position -21 in the miniexon of the nascent RNA. Since the labeling of nascent RNA with ^{32}P will extend all of the growing RNA chains, additional nonspecific RNase T_1-resistant RNA molecules will also be detected. The latter results from structured RNA molecules that are resistant to RNase digestion.

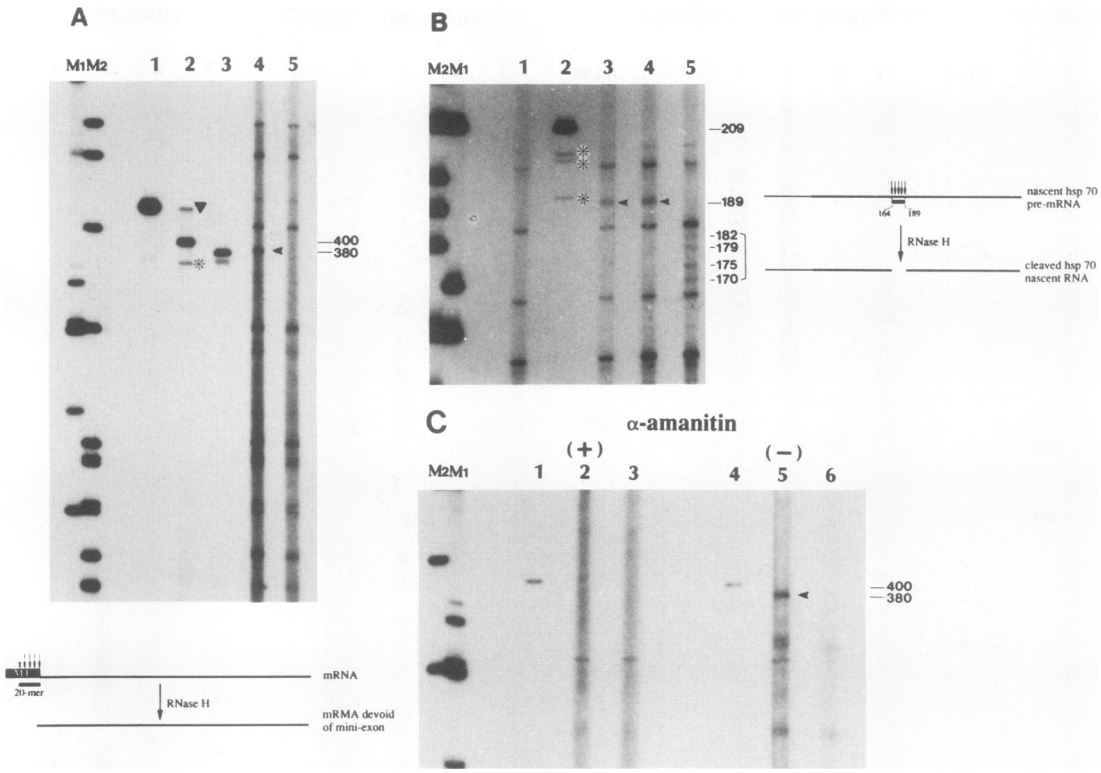

FIG. 3. Characterization of the 5′ end of *hsp70* RNA and α-amanitin sensitivity of chain elongation. (A) As illustrated in the schematic outline below, RNase H pretreatment of total trypanosome RNA in the presence of an oligonucleotide (thick bar) complementary to the miniexon (black box, labeled ME) specifically removes the miniexon from all of the mRNA molecules; RNase T_1 digestion performed on RNase H-treated steady-state RNA annealed with ^{32}P-labeled cRNA-400 (lane 3) showed a 380-nt protection fragment. For comparison, lane 2 shows the major protected fragment, which measured 400 nt in steady-state RNA that has not been treated with RNase H. The RNase T_1 protection pattern with ^{32}P-labeled nascent RNA, protected with unlabeled cRNA-400, comigrates (lane 4) to allow a comparison of the sizes of the different protected bands (lane 5 is the control for RNase T_1 digestion of ^{32}P-labeled nascent RNA without addition of the cRNA probe). Lane 1 shows the undigested ^{32}P-labeled input cRNA-400. A protected fragment resulting from partial RNase T_1 digestion is indicated by a triangle in lane 2. An RNase T_1-resistant band derived from protection of an mRNA with a miniexon added at a second alternate 3′ splice acceptor sites is marked with an asterisk (this band is not expected to change size due to the RNase H treatment since it cannot protect the full-length cRNA). (B) RNase T_1 protection of nascent RNA with cRNA-209 is outlined in the schematic at the right. Lanes: 1, RNase T_1 control on ^{32}P-labeled nascent RNA without the addition of cRNA; 2, RNase T_1 protection pattern with ^{32}P-labeled antisense cRNA-209 and total trypanosome mRNA; 3 to 5, RNase T_1 incubations done in the presence of unlabeled 209-nt cRNA. Lane 3 shows the protection with ^{32}P-labeled nascent RNA; lane 4 represents an incubation with ^{32}P-labeled nascent RNA pretreated with a synthetic antisense miniexon oligonucleotide and RNase H to show that the 189-nt band does not contain a miniexon; lane 5 shows an RNase H pretreatment of ^{32}P-labeled nascent RNA with an antisense coding-sequence oligonucleotide. The RNase T_1-protected fragments in lane 5 are bracketed. The protected fragment at 209 nt is the expected fragment containing the miniexon. The bands marked with asterisks result from protection of the cRNA probe with mRNA that has a miniexon added at an alternate 3′ splice acceptor site visible only if abundant amounts of SP6 counts are loaded to the lane (30). (C) Sensitivity of nascent *hsp70* RNA synthesis to α-amanitin. Nascent RNA elongated in the presence [(+)] of 40 μg of α-amanitin per ml or in the absence of the drug [(−)] was subjected RNase T_1 digestion with the cRNA-400 probe. Lanes: 1 and 4, size of the full-length protected band of 400 nt derived from hybrids between ^{32}P-labeled antisense SP6 cRNA and unlabeled steady-state RNA; 2 and 3, RNase T_1 protection patterns of nascent RNA from α-amanitin-treated nuclei in the presence and absence of unlabeled cRNA-400, respectively; 5 and 6, protection patterns of nascent RNA, prepared from nuclei without added α-amanitin, in the presence and absence of cRNA-400, respectively.

These nonspecific RNAs can be distinguished from specifically protected fragments by comparison of the patterns obtained in incubations with and without (unlabeled) *hsp70* cRNA probes (see legend to Fig. 2 for details).

Each of these three antisense cRNA probes protected a specific major band in nascent RNA measuring (Fig. 2, the band marked with an arrowhead in lane 3 of each panel) 189 nt with cRNA-209, 380 nt with cRNA-400, and 550 nt with cRNA-570 (lanes 4 contain the control incubations, without unlabeled *hsp70* cRNA). Occasionally, minor protected bands which resulted from nonspecific degradation of ^{32}P-labeled *hsp70* nascent RNA could also be detected. The protection patterns show that the major protected fragments were each 20 nt shorter than their full-length protected miniexon-containing cRNA molecules (lane 2 of each panel), indicative of unspliced nascent RNA. The size of the full-length spliced RNA was visualized by RNase T_1 digestion of ^{32}P-labeled SP6 cRNA (Fig. 2, lanes 1) hybridized with unlabeled trypanosome mRNA (Fig. 2, lanes 2; it is important to note that the nucleotide sequence of the full-length protected antisense miniexon containing cRNA and sense-strand miniexon-containing RNA predicts RNase T_1 digestion products of the same size; see below).

The protected *hsp70* nascent RNA was further characterized to confirm that the molecules lacked a miniexon. Total unlabeled trypanosome RNA was treated with RNase H in

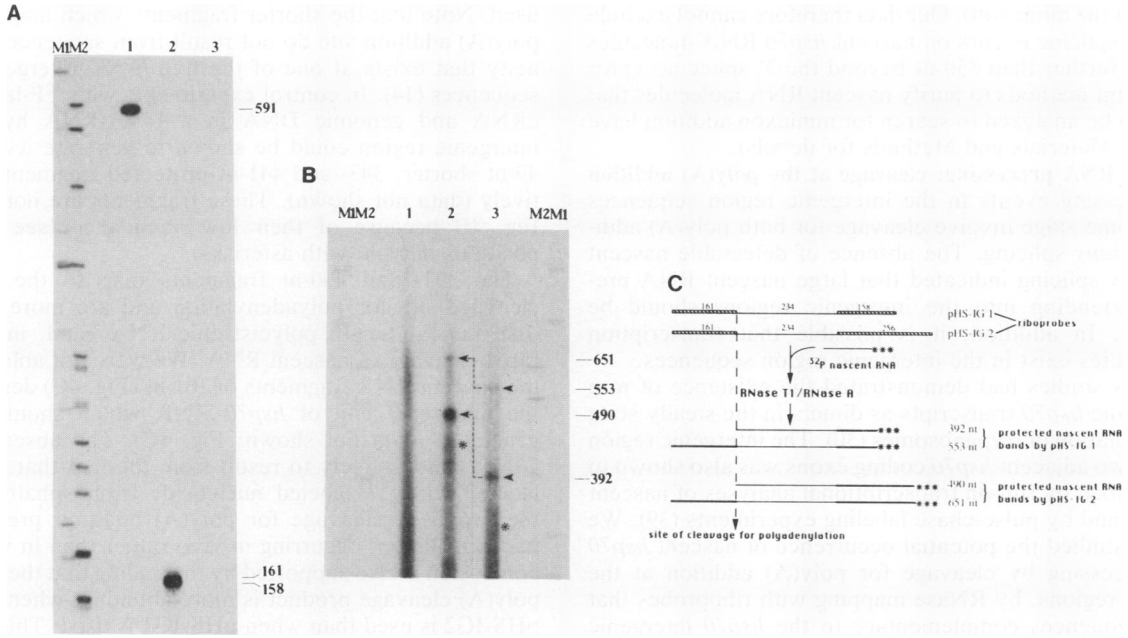

FIG. 4. Analyses of processing events in the intergenic region at nascent *hsp70* RNA. The two probes used, pHS-IG1 and pHS-IG2, are schematically outlined in Fig. 1 and panel C. (A) RNase T_1/RNase A mapping of the mature 5′ and 3′ termini of *hsp70* mRNA. Total trypanosome RNA was subject to RNase T_1/RNase A digestion upon annealing with a ^{32}P-labeled protection riboprobe, pHS-IG1. Lanes: 1, size of undigested pHS-IG1; 2, the two protected fragments mapping to the mature 5′ end (158 nt) and 3′ end (161 nt) of *hsp70* mRNA; 3, control reaction of RNase T_1/RNase A incubation without the addition of total trypanosome RNA. (B) ^{32}P-labeled nascent RNA recovered from hybridization selection was subjected to RNase T_1/RNase A mapping, with either pHS-IG1 or pHS-IG2 used for protection. Lanes: 1, control for RNase T_1/RNase A digestion of ^{32}P-labeled nascent RNA without the addition of unlabeled riboprobes; 2, protection pattern of nascent RNA with unlabeled pHS-IG2; 3, protection pattern of nascent RNA with unlabeled pHS-IG1. Both bands corresponding to the full-length protection (651 and 553 nt) and the bands that map to the cleavage site for polyadenylation (490 and 392 nt) are indicated with arrowheads, and the corresponding bands are aligned by dotted lines. The protected bands derived from pHS-IG1 (lane 3) are 98 nt shorter than those of pHS-IG2. Asterisks mark the predicted positions of nascent RNA molecules that should have resulted from one of the *hsp70* intergenic regions with a sequence heterogeneity. (C) Summary of results of RNase mapping. The protected nascent RNA bands are schematically depicted with tails of asterisks, which represent ^{32}P label incorporated into nascent RNA. The 553- and 651-nt bands correspond to protection of the full-length intergenic region, and the 392- and 490-nt bands represent cleaved nascent *hsp70* RNA, mapping the poly(A) addition site. The thick line illustrates the *hsp70* coding exon of nascent RNA, and the thin line represents the intergenic region. The riboprobes, pHS-IG1 and pHS-IG2, are depicted in Fig. 1.

the presence of an oligonucleotide which is complementary to the last 22 nt of the miniexon, to remove the miniexon. The RNA was then subjected to RNase T_1 digestion with ^{32}P-labeled cRNA-400 for protection. This treatment, which had removed the miniexon from the 5′ end of all mRNAs, including *hsp70* mRNA, resulted in a protected ^{32}P-labeled fragment (Fig. 3A, lane 3) that now comigrated with the 380-nt band of the ^{32}P-labeled nascent *hsp70* RNA (lane 4). RNase H incubations with the miniexon oligomer and the ^{32}P nascent RNA left the 380-nt band (Fig. 3A, lane 4) or the 189-nt band (obtained by protection with cRNA 209; Fig. 3B, lane 4) unaffected. In contrast, the 189-nt (Fig. 3B) and 380-nt (data not shown) *hsp70* nascent RNAs were sensitive to pretreatment of the nascent RNA with an *hsp70* coding-sequence oligonucleotide and RNase H. Pretreatment of ^{32}P-labeled nascent RNA with an antisense 25-mer synthetic *hsp70* coding-sequence oligonucleotide (positions +164 to +189 with respect to the splice site) generated RNase H cleavage products of 170 to 182 nt (Fig. 3B, lane 5), while the 189-nt band has entirely disappeared. These labeled protected fragments map within the range of predicted cleavage sites of RNase H, and the size heterogeneity of the protected bands results from preferential RNase H cleavages in the short DNA-RNA hybrids. From these data we conclude that the protected nascent RNA represents unspliced *hsp70* nascent RNA molecules.

The presence of these unspliced *hsp70* nascent RNA molecules could be abolished by the addition of 40 μg of α-amanitin per ml to the nuclei (Fig. 3C; compare lanes 2 and 3 with lanes 5 and 6). The drug α-amanitin specifically inhibits RNA polymerase II and interferes with transcription of *hsp70* genes (Fig. 3C, lanes 2 and 3), thereby showing that the labeled *hsp70* nascent RNA indeed resulted from RNA polymerase II-mediated chain elongation. The additional, nonspecifically protected bands seen in all the lanes containing ^{32}P-labeled nascent RNA represent RNase T_1-resistant molecules that are unrelated to the *hsp70* RNA. Since most of these RNA molecules persist in the presence of α-amanitin, they are probably rRNA transcripts (Fig. 3C and data not shown).

We analyzed nascent RNA derived from RNA polymerases that were positioned in the 5′ portion of the *hsp70* genes (up to 550 nt downstream of the 3′ splice acceptor site). This region represents only a small portion of the nascent *hsp70* RNA molecules (about one-quarter of the size of a single exon). With the methods described above, it is difficult to discriminate between larger spliced and unspliced molecules because the protected bands will differ by only 39 nt at most

(the size of the miniexon). Our data therefore cannot exclude that *trans* splicing occurs on nascent *hsp70* RNA molecules extending further than 550 nt beyond the 3′ splice acceptor site. Current methods to purify nascent RNA molecules that could then be analyzed to search for miniexon addition have failed (see Materials and Methods for details).

Nascent RNA processing: cleavage at the poly(A) addition site. Processing events in the intergenic region sequences must at some stage involve cleavage for both poly(A) addition and *trans* splicing. The absence of detectable nascent RNA *trans* splicing indicated that large nascent RNA precursors extending into the intergenic regions should be detectable. In addition, it is possible that transcription initiation sites exist in the intergenic region sequences.

Previous studies had demonstrated the existence of rare polycistronic *hsp70* transcripts as dimers in the steady-state RNA population of trypanosomes (30). The intergenic region between two adjacent *hsp70* coding exons was also shown to be transcribed by run-on transcriptional analyses of nascent RNA (30) and by pulse-chase labeling experiments (39). We therefore studied the potential occurrence of nascent *hsp70* RNA processing by cleavage for poly(A) addition at the intergenic regions, by RNase mapping with riboprobes that contain sequences complementary to the *hsp70* intergenic region (probes pHS-IG1 and pHS-IG2 in Fig. 1 and 4).

The ^{32}P-labeled antisense riboprobe, pHS-IG1, protected two major fragments of 158 and 161 nt in total steady-state RNA (Fig. 4A, lane 2). These fragments correspond to the sequences derived from the extreme 5′ and 3′ ends of mature *hsp70* mRNA, respectively (Fig. 4C). With use of the same riboprobe, the 553-nt full-length protected fragment can be seen only in an overexposed autoradiogram (data not shown; 30, 39). In contrast, *hsp70* intergenic region-derived specific fragments were more difficult to identify with ^{32}P-labeled nascent RNA, even though under the same conditions unspliced *hsp70* signals had been readily detected with the cRNA probes (see above). The reasons for this could be that (i) active nascent RNA processing takes place in the intergenic region; (ii) potential transcription initiation sites may exist in the intergenic regions, leading to RNA transcripts with potentially heterogeneous 5′ ends which could obscure intergenic region derived-signals; and (iii) degradation of nascent RNA may occur preferentially at the intergenic regions.

To improve the detection of *hsp70* intergenic region-derived nascent transcripts, we increased the number of trypanosome nuclei used in each assay fourfold, and we also adopted a hybridization selection procedure (see Materials and Methods for details) in order to enrich ^{32}P-labeled *hsp70* nascent RNA prior to performing the RNase T_1/RNase A mapping (RNase A needed to be added due to the presence of background bands that obscured the specific *hsp70* intergenic region-derived nascent RNA signal). Even though RNase A increases the specificity, its use decreased the sensitivity, explaining why it was omitted from the experiments presented in Fig. 2, 3, and 5. As a result of these modifications, we identified two protected fragments with each of the pHS-IG1 and pHS-IG2 riboprobes, which differ by 98 nt at their 3′ ends (Fig. 4B, lanes 2 and 3; Fig. 4C). We could only faintly identify the full-length fragments containing the intergenic regions, measuring 553 nt (pHS-IG1) and 651 nt (pHS-IG2), and we detected shorter fragments mapping to the poly(A) addition site of the *hsp70* mRNA (14; 392 nt with construct pHS-IG1 and 490 nt with pHS-IG2). The relative intensity of the shorter protected fragments was always much stronger (lane 2) when probe pHS-IG2 was used. Note that the shorter fragments which mapped to the poly(A) addition site do not result from sequence heterogeneity that exists at one of the five *hsp70* intergenic-region sequences (14). In control experiments with ^{32}P-labeled SP6 cRNA and genomic DNA in a DNA-RNA hybrid, this intergenic region could be shown to generate its predicted 49-nt shorter, 343- and 441-nt protected fragments, respectively (data not shown). These fragments are not visible in Fig. 4B because of their low abundance (see predicted positions marked with asterisks).

The 392- and 490-nt fragments map to the predicted cleavage site for polyadenylation and are more abundant than the full-length polycistronic RNA band, indicating a rapid cleavage of nascent RNA. We were not able to detect the nascent RNA fragments of 161 nt (Fig. 4C) derived from the mature 3′ end of *hsp70* RNA which should also be generated (data not shown; Fig. 4C). The absence of the 161-nt band is likely to result from the fact that it was not labeled with ^{32}P-labeled nucleoside triphosphate, because the proposed cleavage for poly(A) addition preceded the nuclei isolation, occurring in vivo rather than in vitro. This conclusion is also supported by the finding that the presumed poly(A) cleavage product is more abundant when construct pHS-IG2 is used than when pHS-IG1 is used. This intensity difference can reflect the additional time of exposure of the longer nascent RNA to the enzyme that performs the cleavage.

Finally, the intensities of the bands obtained with the intergenic-region probes were at least 10-fold less than the level of unspliced nascent *hsp70* RNA that could be identified with the cRNA probes. We interpret this to mean that the majority of the processed nascent RNA intermediates are rapidly trimmed by exonuclease activities in isolated nuclei (20), thereby generating heterogeneous 5′ ends in the intergenic region which cannot be easily identified.

Determination of the in vitro RNA polymerase chain elongation rate in the *hsp70* locus. In trypanosomes, five *hsp70* genes are arranged in a tandem array of a 3.2-kb repeat unit. The validity of several of the observations described above is dependent on the extent of read-through by RNA polymerases in this array during the 6-min labeling period. Extensive read-through (several hundreds of nucleotides) might have resulted in the generation of newly synthesized stretches of labeled nascent RNA, explaining why we detected only nascent RNA without a miniexon.

The published overall in vitro rate of RNA polymerase-mediated chain elongation in eukaryotes is approximately 25 to 50 nt/min (10, 19, 60). The overall rate of nascent RNA chain elongation in trypanosomes, measured according to the method of Marzluff et al. (34), was 30 nt (see Materials and Methods for details; data not shown). However, the overall rate of elongation may not accurately reflect the rate of elongation at the *hsp70* locus, and we therefore developed a method for determining the rate of RNA polymerase II chain elongation in vitro at the *hsp70* locus. The approach taken was to measure the time needed for RNA polymerase II to reach and traverse a hypothetical window in the *hsp70* gene, between nt 400 and 520 downstream of the 3′ splice acceptor site (Fig. 5C). We selected for nascent RNA that was labeled in the first 400 nt downstream of the 3′ splice acceptor site and that was extended for more than 120 nt by using nascent RNA longer than 520 nt (Fig. 5B shows the size-selected RNA; Fig. 5C outlines the method). Since the instability of the intergenic region indicated that most RNAs were cleaved in the intergenic region, we simply size selected RNA longer than 520 nt, assuming that the majority of

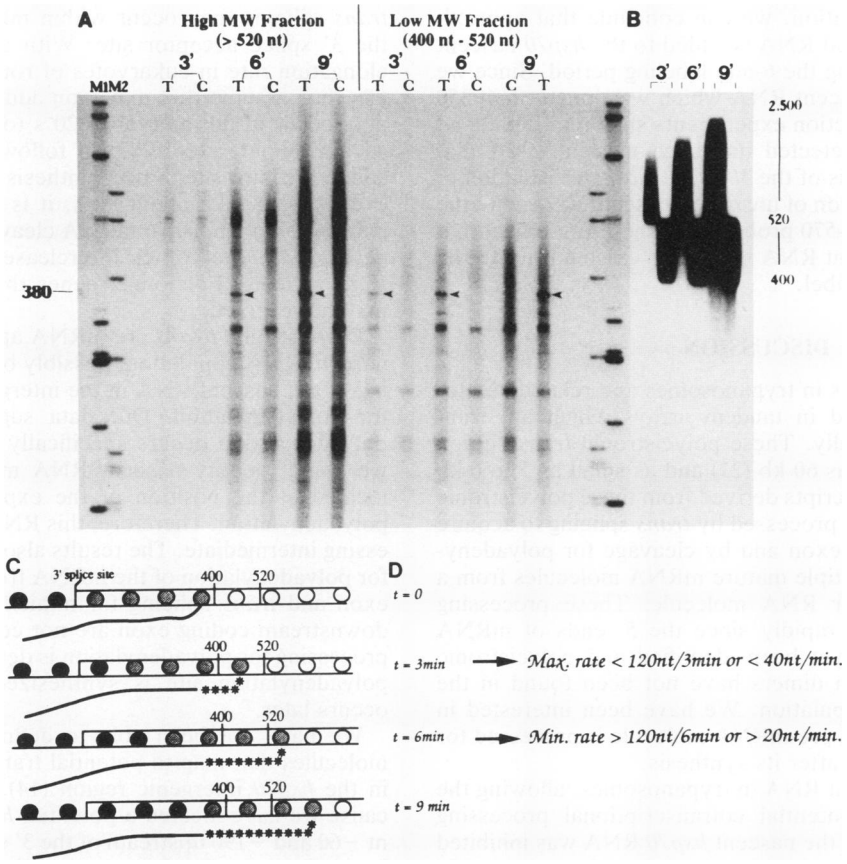

FIG. 5. Determination of the rate of elongation at the *hsp70* locus. (A) ^{32}P-labeled nascent RNA, recovered after 3, 6, and 9 min of elongation, from nuclei (lanes 3′, 6′, and 9′) was size fractionated in two separate molecular weight (MW) fractions ranging from 400 to 520 nt and from 520 to 2,500 nt (B). The ^{32}P-labeled nascent RNA was subjected to RNase T_1 digestion in the presence of unlabeled antisense cRNA-400 (lanes T) or without the addition of cRNA-400 (lanes C). The expected 380-nt protection fragment is indicated on the left side. (C) Schematic outline showing RNA chain elongation on the *hsp70* gene at the various time points. At the 0 time point, there are three classes of RNA polymerases: preceding the 3′ splice acceptor site (black), between nt 1 and 400 downstream of the splice acceptor site (gray), and beyond nt position 400 (white). The fourth gray RNA polymerase is ready to proceed into the 400- to 520-nt window. This RNA polymerase has been marked with a nascent RNA chain, and its incorporation of label is marked by asterisks. The rate of elongation is determined by following the gray RNA polymerase. At the 3-min time point, it fails to reach the 520- to 2,500-nt RNA fraction but can be found in the 400- to 520-nt fraction. At the 6-min time point, the gray RNA polymerase has reached the 520-nt mark (D), and at 9 min some of the gray RNA polymerases are beyond nt 520. Since the gray RNA polymerases started in the first 400 nt downstream of the 3′ splice acceptor site, their labeled nascent RNA will give rise to a protected fragment if incubated with the cRNA-400 and RNase T_1, while extension of nascent RNA and incorporation of label by any white RNA polymerases will not give rise to a protected fragment with the cRNA-400 probe, as the labeled portion of the RNA is not protected by cRNA and will be digested by RNase T_1.

the RNA had heterogeneously distributed 5′ ends located somewhere upstream of the 3′ splice acceptor site. We then determined whether this large RNA (over 520 nt in size) can efficiently protect the cRNA-400. Figure 5A shows that it was not possible to detect unspliced 380-nt *hsp70* nascent RNA in the RNA fraction that was longer then 520 nt if it had been extended by only a short, 3-min incubation. However, after 6 min of elongation, a faint protection signal could be seen, while after a 9-min elongation, abundant labeling of the 380-nt *hsp70* nascent RNA band was detected (bands marked with arrowheads). The control for this experiment was to compare these data with the protection pattern that resulted with shorter (400- to 520-nt) ^{32}P-labeled nascent RNA derived from the nuclear run-on experiments described above. In contrast to the larger RNA, the smaller (400- to 520-nt) nascent RNA fraction gave a labeled protected fragment after a 3-min elongation, the intensity of which increased after 6 and 9 min. From the experiment using the larger (>520-nt) RNA, we can conclude that the rate of elongation must be less than 40 nt/min (less than 120 nt in 3 min), at least 20 nt/min (minimally 120 nt in 6 min), or more than 13 nt/min (synthesis of at least 120 nt in 9 min). The shorter (400- to 520-nt) RNA shows that the rate of elongation must be close to 20 nt/min (efficient labeling in 6 min) and less than 40 nt/min (suboptimal labeling after 3 min). The rate of elongation at the *hsp70* locus in vitro is therefore roughly between 20 and 40 nt/min (comparable values were obtained in three other experiments). This value is of the same order as the overall value for elongation by RNA polymerases I, II, and III in trypanosome nuclei (30 nt/min; data not shown) and the previously published in vitro rates of RNA chain elongation measured in other eukaryotes (25 to 50 nt/min; 10, 11). We therefore assume that the rate of elongation at the *hsp70* genes is 30 nt/min.

The experiments confirm that in our nuclei the RNA polymerase traverses the template and elongates *hsp70*

nascent RNA. In addition, we can conclude that approximately 180 nt of labeled RNA is added to the *hsp70* nascent RNA transcripts during the 6-min labeling period. Since we used size-selected nascent RNA which was longer than 350 nt in the RNase protection experiments shown in Fig. 2 and 3, we cannot have detected unspliced nascent RNA that resulted from synthesis of the 3′ splice site after isolation of the nuclei. The detection of unspliced nascent RNA with the cRNA-400 and cRNA-570 probes must therefore reflect that these unspliced nascent RNA molecules existed prior to the incorporation of ^{32}P label.

DISCUSSION

Protein-coding genes in trypanosomes and related kinetoplastidae are arranged in tandem arrays which are transcribed polycistronically. These polycistronic transcription units can be as large as 60 kb (21) and as small as 5 to 6 kb (47). Pre-mRNA transcripts derived from these polycistronic transcription units are processed by *trans* splicing to acquire the 39-nt capped miniexon and by cleavage for polyadenylation to generate multiple mature mRNA molecules from a single large precursor RNA molecule. These processing events must proceed rapidly since the 5′ ends of mRNA precursors have not yet been identified and polycistronic precursors larger than dimers have not been found in the steady-state RNA population. We have been interested in determining whether a pre-mRNA molecule is processed for maturation, during or after its synthesis.

We analyzed nascent RNA in trypanosomes, allowing the characterization of potential cotranscriptional processing events. ^{32}P labeling of the nascent *hsp70* RNA was inhibited by the addition of a low concentration of the drug α-amanitin (40 μg/ml), indicating that synthesis of the *hsp70* nascent RNA was mediated by RNA polymerase II (32). In addition, labeling of the nascent RNA with either [^{32}P]UTP or [^{32}P]GTP (data not shown) gave identical results, showing little or no effect of the uridyl terminal transferase activity, which could have artificially labeled the nascent RNA molecules (61). We also measured the elongation rate of run-on transcription in the *hsp70* locus in isolated nuclei and determined this rate to be between 20 and 40 nt/min. This value was sufficiently low to allow labeling of nascent RNA with a short ^{32}P tail for the further analyses of processing events that occurred prior to isolation of the nuclei.

The results of the RNase mapping, using a set of *hsp70* cRNA probes of various lengths, revealed only unspliced nascent RNA, showing that when maximally 550 nt of nascent *hsp70* RNA located downstream of its 3′ splice acceptor site was synthesized, *trans* splicing had not yet occurred. These observations have not ruled out the hypothetical model of *trans* splicing at the nascent RNA level, since the size of each *hsp70* coding exon is 2,226 nt and we have only analyzed the occurrence of *trans* splicing during the synthesis of roughly the first 550 nt that are located downstream of the 3′ splice acceptor site. The data, however, indicate that *trans* splicing does not occur immediately after the synthesis of the 3′ splice acceptor site. Presumably, it is the time required for spliceosome assembly (5, 12, 17) and/or the effects of steric hindrance due to the close proximity of the RNA polymerase and the 3′ splice acceptor site, which determines when *trans* splicing occurs with respect to transcription. Spliceosome assembly and the subsequent intron removal in *cis* splicing have been estimated to proceed within a 3-min interval in vivo. Measurements of the half-life of medRNA (26, 27, 56) indicate that *trans* splicing may occur within minutes after synthesis of the 3′ splice acceptor site. With an in vivo transcription elongation rate in eukaryotes of roughly 25 nt/s (1, 11), we can thus assume that miniexon addition at the *hsp70* genes may occur in an interval of 20 s (or 550 nt; this report) to several minutes (26, 27, 56) following synthesis of the 3′ splice acceptor site. Since synthesis of a single *hsp70* coding exon should take about 90 s, it is conceivable that in the maturation of an *hsp70* mRNA cleavage for polyadenylation occurs first, after which the released RNA is processed by *trans* splicing. For longer genes the order of these events may be reversed.

Polycistronic *hsp70* pre-mRNA appeared to be rare in the nascent RNA population, possibly because of rapid processing of the nascent RNA in the intergenic region by cleavage for polyadenylation. Our data suggest that cleavage for polyadenylation occurs specifically at nascent RNA, since we could identify nascent RNA molecules with a 5′ end located at the position of the expected cleavage site for polyadenylation. Therefore, this RNA may represent a processing intermediate. The results also show that 3′ processing for polyadenylation of the mRNA from the preceding coding exon and *trans* splicing for miniexon addition at the next downstream coding exon are not coupled events, since the processing for polyadenylation is detectable shortly after the polyadenylation site is synthesized, while *trans* splicing occurs later.

We have not been able to define the 5′ ends of RNA molecules that map to potential transcription initiation sites in the *hsp70* intergenic region (14). This is surprising, because we have located a potential *hsp70* promoter between nt −60 and −190 upstream of the 3′ splice acceptor site (19a) in transient transfection assays with *hsp70* intergenic regions fused to the chloramphenicol acetyltransferase gene in insect-form trypanosomes. We interpret this to mean that these potential intergenic region initiation sites either are not used or are used at a low frequency in vivo.

We speculate that the potential cotranscriptional cleavage for polyadenylation could serve a role in regulating mRNA abundance if differential control of the processing step occurs. A similar speculative argument could be made for *trans* splicing, utilizing differential, co- or posttranscriptional *trans* splicing to control mRNA abundance of genes that are transcribed constitutively.

Finally, the finding of a potential processing intermediate caused by cleavage for polyadenylation of pre-mRNA shortly (within seconds) after its synthesis suggests a cotranscriptional site-specific cleavage event for mRNA maturation. This view is consistent with previous results in which adenovirus type 2 pre-mRNA was studied, which showed that cleavage for polyadenylation preceded transcription termination and may be completed within 1 min of transcription of the site (33, 41). The mechanism for polyadenylation in trypanosomes is still unclear, and these organisms lack the usual AAUAAA consensus sequence (2) found in the vicinity of the poly(A) addition site in other eukaryotes. Our data indicate that in trypanosomes, transcription termination is not a prerequisite for the cleavage for polyadenylation and polycistronic pre-mRNA is likely to be processed cotranscriptionally by cleavage for poly(A) addition.

ACKNOWLEDGMENTS

We thank Steven Brown, Hui-Min Chung, Keith Gottesdiener, Sylvia Le Blancq, Mary Lee, Gloria Rudenko, Demetrios Vassilatis, and Michael Weiden for critical reading of the manuscript.

This work was supported by NIH grant AI 21784 to L.H.T.V.D.P. and by a grant from the John D. and Catherine T. MacArthur Foundation. L.H.T.V.D.P. is a Burroughs Wellcome Scholar in Molecular Parasitology.

REFERENCES

1. **Beyer, A. L., and Y. N. Osheim.** 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. Genes Dev. **2**:754–765.
2. **Birnstiel, M. L., M. Busslinger, and K. Strub.** 1985. Transcription termination and 3′ processing: the end is in site! Cell **41**:349–359.
3. **Boothroyd, J. C.** 1985. Antigenic variation in African trypanosomes. Annu. Rev. Microbiol. **39**:475–502.
4. **Borst, P.** 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annu. Rev. Biochem. **55**:701–732.
5. **Brody, E., and J. Abelson.** 1985. The "spliceosome": yeast pre-mRNA associates with a 40S complex in a splicing-dependent reaction. Science **228**:963–967.
6. **Bruzik, J. P., and J. A. Steitz.** 1990. Spliced leader RNA sequences can substitute for the essential 5′ end of U1 RNA during splicing in a mammalian *in vitro* system. Cell **62**:889–899.
7. **Bruzik, J. P., K. Van Doren, D. Hirsch, and J. A. Steitz.** 1988. Trans-splicing involves a novel form of small nuclear ribonucleoprotein particles. Nature (London) **335**:559–562.
8. **Campbell, D. A., D. A. Thornton, and J. C. Boothroyd.** 1984. Apparent discontinuous transcription of *Trypanosoma brucei* variant surface antigen genes. Nature (London) **311**:350–355.
9. **Cornelissen, A. W. C. A., M. P. Verspieren, J. J. Toulme, B. W. Swinkels, and P. Borst.** 1986. The common 5′ terminal sequence on trypanosome mRNAs: a target for anti-messenger oligodeoxynucleotides. Nucleic Acids Res. **14**:5605–5614.
10. **Cox, R. F.** 1976. Quantitation of elongating form A and B RNA polymerases in chick oviduct nuclei and effects of estradiol. Cell **7**:455–465.
11. **Darnell, J. E.** 1982. Variety in the level of gene control in eukaryotic cells. Nature (London) **297**:365–367.
12. **Frendeway, D., and W. Keller.** 1985. Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. Cell **42**:355–367.
13. **Gibson, W. C., B. W. Swinkels, and P. Borst.** 1988. Post-transcriptional control of the differential expression of phosphoglycerate kinase gene in *Trypanosoma brucei*. J. Mol. Biol. **201**:315–325.
14. **Glass, D. J., R. I. Polvere, and L. H. T. van der Ploeg.** 1986. Conserved sequences and transcription of the hsp 70 gene family in *Trypanosoma brucei*. Mol. Cell. Biol. **6**:4657–4666.
15. **Gottesdiener, K., H. M. Chung, S. Brown, M. G.-S. Lee, and L. H. T. van der Ploeg.** 1991. Characterization of VSG gene expression site promoters and identification of promoter-associated DNA rearrangement events. Mol. Cell. Biol. **11**:2467–2480.
16. **Grabowski, P. J., R. A. Padgett, and P. A. Sharp.** 1984. Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate. Cell **37**:415–427.
17. **Grabowski, P. J., S. R. Seiler, and P. A. Sharp.** 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell **42**:345–353.
18. **Green, M. R.** 1986. Pre-mRNA splicing. Annu. Rev. Genet. **20**:617–708.
19. **Hofer, E., and J. E. Darnell.** 1981. The primary transcription unit of the mouse B-major globin gene. Cell **23**:585–593.

19a. **Huang, J., M. G.-S. Lee, and L. H. T. van der Ploeg.** Unpublished data.

20. **Huang, J., and L. H. T. van der Ploeg.** 1988. A 5′ exoribonuclease and RNA ligase of *T. brucei*. Nucleic Acids Res. **16**:9567–9585.

20a. **Huang, J., and L. H. T. van der Ploeg.** Unpublished data.

21. **Johnson, P. J., J. M. Kooter, and P. Borst.** 1987. Inactivation of transcription by UV irradiation of *T. brucei* provides evidence for a multicistronic transcription unit including a VSG gene. Cell **51**:273–281.
22. **Kooter, J. M., T. De Lange, and P. Borst.** 1984. Discontinuous synthesis of mRNA in trypanosomes. EMBO J. **3**:2387–2392.
23. **Kooter, J. M., H. J. Van der Spek, R. Wagter, C. E. d'Oliveira, F. Van der Hoeven, P. J. Johnson, and P. Borst.** 1987. The anatomy and transcription of a telomeric expression site for variant-specific surface antigens in *T. brucei*. Cell **51**:261–272.
24. **Krause, M., and D. Hirsch.** 1987. A trans-spliced leader sequence on actin mRNA in *C. elegans*. Cell **49**:753–761.
25. **Lai, C.-J., R. Dhar, and G. Khoury.** 1978. Mapping of the spliced and unspliced late lytic SV40 RNA. Cell **14**:971–982.
26. **Laird, P. W., A. L. M. A. ten Asbroek, and P. Borst.** 1987. Controlled turnover and 3′ trimming of the trans-splicing precursor of *Trypanosoma brucei*. Nucleic Acids Res. **24**:10087–10103.
27. **Laird, P. W., J. C. B. M. Zomerdijk, D. De Korte, and P. Borst.** 1987. *In vivo* labelling of intermediates in the discontinuous synthesis of mRNAs in *Trypanosoma brucei*. EMBO J. **6**:1055–1062.
28. **Lee, M. G.-S., R. I. P. Polvere, and L. H. T. van der Ploeg.** 1990. Evidence for segmental gene conversion between a cognate hsp 70 gene and the temperature sensitively transcribed hsp 70 genes of *Trypanosoma brucei*. Mol. Biochem. Parasitol. **41**:213–220.
29. **Lee, M. G.-S., and L. H. T. van der Ploeg.** 1987. Frequent independent duplicative transpositions activate a single VSG gene. Mol. Cell. Biol. **7**:357–364.
30. **Lee, M. G.-S., and L. H. T. van der Ploeg.** 1990. Transcription of the heat shock 70 locus in *Trypanosoma brucei*. Mol. Biochem. Parasitol. **41**:221–232.
31. **Lee, M. G.-S., and L. H. T. van der Ploeg.** 1990. Homologous recombination and stable transfection in the parasitic protozoan *Trypanosoma brucei*. Science **250**:1583–1589.

31a. **Lee, M. G.-S., and L. H. T. van der Ploeg.** Unpublished data.

32. **Lindell, T. J., F. Weinberg, P. W. Morris, R. G. Roeder, and W. J. Rutter.** 1970. Specific inhibition of nuclear RNA polymerase II by alpha-amanitin. Science **170**:447–449.
33. **Manley, J. L., P. A. Sharp, and M. L. Gefter.** 1982. RNA synthesis in isolated nuclei. Processing of adenovirus serotype 2 late messenger RNA precursors. J. Mol. Biol. **159**:581–599.
34. **Marzluff, W. F., C. Pan, and D. L. Cooper.** 1978. RNA synthesis in myeloma cells synchronized by isoleucine starvation. Nucleic Acids Res. **5**:4176–4193.
35. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
36. **Mory, Y. Y., and L. G. Malcolm.** 1977. Synthesis of messenger RNA-like molecules in isolated myeloma nuclei. Nucleic Acids Res. **4**:1739.
37. **Mottram, J., K. L. Perry, P. M. Lizardi, R. Luhrmann, N. Agabian, and R. G. Nelson.** 1989. Isolation and sequence of four small nuclear U RNA genes of *Trypanosoma brucei* subsp. *brucei*: identification of the U2, U4, and U6 RNA analogs. Mol. Cell. Biol. **9**:1212–1223.
38. **Muhich, M. L., and J. C. Boothroyd.** 1988. Polycistronic transcripts in trypanosomes and their accumulation during heat shock: evidence for a precursor role in mRNA synthesis. Mol. Cell. Biol. **8**:3837–3846.
39. **Muhich, M. L., and J. C. Boothroyd.** 1989. Synthesis of trypanosome hsp 70 mRNA is resistant to disruption of trans-splicing by heat shock. J. Biol. Chem. **264**:7107–7110.
40. **Murphy, W. J., K. P. Watkins, and N. Agabian.** 1986. Identification of a novel Y-branch structure as an intermediate in trypanosome mRNA processing: evidence for trans-splicing. Cell **47**:517–525.
41. **Nevens, J. R., and J. E. Darnell.** 1978. Steps in the processing of Ad2 mRNA: poly(A)$^+$ nuclear sequences are conserved and poly(A) addition precedes splicing. Cell **15**:1477–1493.
42. **Osheim, Y. N., O. L. Miller, and A. L. Beyer.** 1986. RNP particles at splice junction sequences on Drosophila chorion transcripts. Cell **43**:143–151.
43. **Padgett, R. A., P. J. Grabowski, M. M. Konarska, S. Seiler, and P. A. Sharp.** 1986. Splicing of messenger RNA precursors.

Annu. Rev. Biochem. **55**:1119–1150.
44. **Pays, E., H. Coquelet, P. Tebabi, and M. Steinert.** 1990. *Trypanosoma brucei*: constitutive activity of the VSG gene and procyclin promoters. EMBO J. **9**:3145–3151.
45. **Ralph, D., J. Huang, and L. H. T. van der Ploeg.** 1988. Physical identification of branched intron side-products of splicing in *Trypanosoma brucei*. EMBO J. **7**:2539–2545.
46. **Roop, D. R., J. L. Nordstrom, S. Y. Tsai, M. J. Tsai, and B. W. O'Malley.** 1978. Transcription of structural and intervening sequences in the ovalbumin gene and identification of potential ovalbumin mRNA precursors. Cell **15**:671–685.
47. **Rudenko, G., S. Le Blancq, J. Smith, G.-M. S. Lee, A. Rattray, and L. H. T. van der Ploeg.** 1990. Alpha-amanitin resistantly transcribed PARP genes located in an unusually small polycistronic transcription unit: PARP promoter analysis by transient DNA transfection of *Trypanosoma brucei*. Mol. Cell. Biol. **10**:3492–3504.
48. **Sharp, P. A.** 1987. Splicing of messenger RNA precursors. Science **235**:766–771.
49. **Shea, C., M. G. S. Lee, and L. H. T. van der Ploeg.** 1987. VSG gene 118 is transcribed from a contransposed Pol I-like promoter. Cell **50**:603–612.
50. **Smith, M. M., A. E. Reeve, and R. C. C. Huang.** 1978. Analysis of RNA initiated in isolated mouse myeloma nuclei using purine nucleoside 5′ [gamma-S] triphosphates as affinity probes. Cell **15**:615–626.
51. **Sutton, R. E., and J. C. Boothroyd.** 1986. Evidence for trans splicing in trypanosomes. Cell **47**:527–535.
52. **Tilghman, S. M., P. J. Curtis, D. C. Tiemeier, P. Leder, and C. Weismann.** 1978. The intervening sequence of a mouse β-globin gene is transcribed within the 15S -globin mRNA precursor. Proc. Natl. Acad. Sci. USA **75**:1309–1313.
53. **Tschudi, C., F. Richards, and E. Ullu.** 1986. The U2 analogue of *Trypanosoma brucei gambiense*: implications for a splicing mechanism in trypanosomes. Nucleic Acids Res. **14**:8893–8903.
54. **Tschudi, C., and E. Ullu.** 1988. Polygene transcripts are precursors to calmodulin mRNAs in trypanosomes. EMBO J. **7**:455–463.
55. **Tschudi, C., and E. Ullu.** 1990. Destruction of U2, U4, or U6 small nuclear RNA blocks trans-splicing in trypanosome cells. Cell **61**:459–466.
56. **Ullu, E., and C. Tschuci.** 1990. Permeable trypanosome cells as a model system for transcription and trans-splicing. Nucleic Acids Res. **18**:3319–3326.
57. **Van der Ploeg, L. H. T.** 1986. Discontinuous transcription and splicing in trypanosomes. Cell **47**:479–480.
58. **Van Doren, K., and D. Hirsch.** 1988. Trans-spliced leader RNA exists as small nuclear ribonucleoprotein particles in *Caenorhabditis elegans*. Nature (London) **335**:556–559.
59. **Walder, J. A., P. S. Eder, D. M. Engman, S. T. Brentano, R. Y. Walder, D. S. Knutzon, D. M. Dorfman, and J. E. Donelson.** 1986. The 35-nucleotide splice leader sequence is common to all trypanosome messenger RNA. Science **233**:569–571.
60. **Weber, J., W. Jelinek, and J. E. Darnell.** 1977. The definition of a large viral transcription unit late in Ad2 infection of Hela cells: mapping of nascent RNA molecules labeled in isolated nuclei. Cell **10**:611–616.
61. **White, T. C., and P. Borst.** 1987. RNA end-labeling and RNA ligase activities can produce a circular rRNA in whole cell extracts from trypanosomes. Nucleic Acids Res. **15**:3275–3290.
62. **Zomerdijk, J. C. B. M., M. Ouelette, A. L. M. A. Ten Asbroek, R. Kieft, M. M. A. 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.