Cell, Vol. 47, 517–525, November 21, 1986, Copyright © 1986 by Cell Press

Identification of a Novel Y Branch Structure as an Intermediate in Trypanosome mRNA Processing: Evidence for *Trans* Splicing

William J. Murphy, Kenneth P. Watkins, and Nina Agabian
Department of Biomedical and Environmental Health Science
School of Public Health
University of California
Berkeley, California 94720
Naval Bioscience Laboratory
140 Warren Hall
Berkeley, California 94720

Summary

We present evidence that addition of the 35 nucleotide spliced leader (SL) to the 5′ end of T. brucei mRNAs occurs via *trans* RNA splicing. A 100 nucleotide fragment of the 135 base SL RNA (100-mer) is revealed by S1 nuclease analysis of total and poly(A)$^+$ RNA. This 100-mer is not detected by Northern hybridization analysis, indicating that it does not exist free in the cell. The 5′ end of the 100-mer maps precisely to the conserved splice junction sequence of the SL RNA. Purified debranching enzyme releases this 100-mer RNA as a free, 100 nucleotide species. This indicates that the 100-mer is covalently linked to poly(A)$^+$ RNA by a 2′-5′ phosphodiester bond, that the branched intermediate has a discontinuous intron or Y structure (rather than a lariat), which is expected of a *trans*-spliced mRNA, and that the SL RNA is indeed the donor of the SL sequence to trypanosome mRNAs.

Introduction

The study of gene expression in trypanosomes, such as Trypanosoma brucei, has led to the observation that messenger RNAs in these protozoa are composed of two distinct RNA components, the 35 base spliced leader (SL), common to virtually every trypanosome mRNA, and the target messenger RNA itself (Van der Ploeg et al., 1982; Boothroyd and Cross, 1982; Parsons et al., 1984; De Lange et al., 1984a). Surprisingly, however, this SL sequence is not contiguously encoded with any known structural gene (Nelson et al., 1983; De Lange et al., 1983), but rather is transcribed as part of a unique small RNA (135 nucleotides) that has the SL at its 5′ end (Milhausen et al., 1984; Campbell et al., 1984; Kooter et al., 1984). The gene that encodes this SL RNA is repeated about 200 times in the trypanosome genome (Nelson et al., 1983; De Lange et al., 1983; Michiels et al., 1983) and is transcribed at a high rate in isolated nuclei (Kooter et al., 1984). In addition, the SL RNA possesses a 7-methylguanosine cap, indicative of RNA polymerase II transcripts (Perry et al., submitted). These observations implicate the role of the SL RNA as the donor of the SL sequence to trypanosome structural mRNAs and further indicate that some process of RNA joining is responsible for generating the mature mRNAs.

In higher eukaryotes and yeast, exon RNAs are joined to form mature mRNAs by splicing, coincident with the removal of intervening sequences. This process of RNA splicing has been reconstituted in vitro and has become better understood in recent years (Hardy et al., 1984; Krainer et al., 1984; Ruskin et al., 1984; Padgett et al., 1984; Lin et al., 1985; Newman et al., 1985; Perkins et al., 1986; Cheng and Abelson, 1986). RNA splicing proceeds via a precise cleavage at the 5′ exon–intron junction of the pre-mRNA followed by (or concurrent with) the formation of a 2′-5′ phosphodiester linkage between the 5′ end of the intron and a specific adenosine nucleotide located 18 to 40 bases upstream from the 3′ splice junction (Ruskin et al., 1984; Padgett et al., 1984). The result is an intron–exon lariat structure. While the sequence surrounding this 2′-5′ branch point is highly conserved in yeast (the TACTAAC box; Langford and Gallwitz, 1983; Pikielny et al., 1983; Teem et al., 1984), it is less conserved in higher eukaryotes (Ruskin et al., 1984; Padgett et al., 1984; Zeitlin and Efstratiadis, 1984; Keller and Noon, 1984; Reed and Maniatis, 1985; Ruskin et al., 1985). Subsequent to the formation of the lariat intermediate, the RNA is specifically cleaved at the 3′ splice junction, releasing the free intron–lariat structure, while the two exons are ligated to form the mature mRNA. The intron–lariat can be linearized in vitro by debranching enzyme, which specifically cleaves the branched 2′-5′ phosphodiester bond (Ruskin and Green, 1985; Arenas and Hurwitz, submitted).

The sequence immediately adjacent to the SL in the SL RNA (see Figure 1) conforms to the consensus sequence proposed for higher eukaryotic 5′ splice sites (Mount, 1982; Milhausen et al., 1984; De Lange et al., 1983). Similarly, analysis of genomic clones of trypanosome variant surface glycoprotein genes, as well as other structural genes, indicates that the point of attachment of the SL in mature mRNAs closely resembles the 3′ consensus splice sequence (Van der Ploeg et al., 1982; Boothroyd and Cross, 1982; Michiels et al., 1983; Murphy et al., 1984; Bernards et al., 1984; Cully et al., 1985; Sather and Agabian, 1985; Tschudi et al., 1985; Swinkels et al., 1986). The conservation of these consensus splice junctions in trypanosome genes would therefore argue that SL joining to structural gene transcripts in these protozoa shares some features of pre-mRNA splicing in higher eukaryotes.

One can envision several mechanisms for the generation of mature mRNAs in trypanosomes, some of which are illustrated in Figure 2. These models can be placed into two general categories: primed transcription, whereby transcription of target genes is dependent upon the SL RNA (Figures 2A, 2B, and 2C) or independent transcription, whereby the SL and target genes are independently transcribed and joined subsequently (Figures 2D and 2E). An examination of the possible intermediates and end products shown in Figure 2 indicates structures that, if detected in vivo, would differentiate between these models.

We describe here experiments designed to detect the presence of various splicing intermediates and end prod-

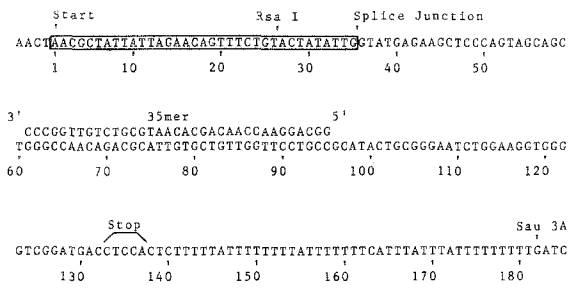

Figure 1. Sequence of the SL RNA Gene and Flanking Region

The sequence of the SL RNA repeat unit shown is taken from Milhausen et al., 1984. Nucleotide number 1 indicates the first position of the SL (enclosed in box). The 35 base oligonucleotide primer described in the text is aligned with its complementary sequence in the SL RNA gene. The SL splice junction, as well as the putative transcription initiation and termination sites (start and stop, respectively), is also indicated.

ucts. Our results reveal the presence of a 100 nucleotide fragment of the SL RNA whose 5′ end maps precisely to the 5′ splice junction. This 100-mer is readily released from poly(A)$^+$ RNA by debranching enzyme, suggesting that it is covalently attached to RNA through a 2′-5′ phosphodiester bond. Release of a free 100-mer species indicates the presence of an intermediate structure in the form of a Y rather than a lariat, as one would expect of a discontinuous intron that is spliced in *trans*.

Results

A 100 Nucleotide Fragment of the SL RNA Is Associated with High Molecular Weight RNA

Throughout this manuscript we will consider the 135 nucleotide SL RNA as having two functional domains, the 35 nucleotide SL and the 100 nucleotide downstream sequence, which is not found in mature mRNAs. We refer to this 100 nucleotide species as the 100-mer fragment, although alternative cleavages could result in an SL RNA–derived fragment of the same length. Of the pathways outlined in Figure 2, four require an endonucleolytic cleavage that would generate a 3′ 100-mer fragment of the SL RNA (Figures 2A, 2C, 2D, and 2E); only two of these would result in the production of free 100-mer RNA (Figures 2C and 2E). The remaining pathway (Figure 2B) would require an additional cleavage within the 100-mer fragment. Detection of the free 100-mer species would depend on the rate of degradation of these end products and, as outlined in Figure 2E, the presence of a trypanosome debranching enzyme. Each of the other three models (Figures 2A, 2B, and 2D) predicts that the 100-mer fragment would be covalently associated with high molecular weight RNA. We therefore assayed for the presence of free versus complexed 100-mer in different preparations of T. brucei RNA.

In Figure 3 a comparison of four different preparations of T. brucei RNA analyzed by either Northern blot (Figure 3A) or S1 protection (Figure 3B) shows that preparative and analytical methods affect the detection of the SL RNA and related molecules. Northern blot analysis reveals the SL RNA (135 nucleotides) in total and poly(A)$^-$ RNAs prepared by LiCl–urea precipitation (Figure 3A, lanes 1 and 3, respectively), whereas little, if any, of this RNA is detected in RNA prepared by the guanidine thiocyanate–CsCl method (lane 2). Small RNA molecules sediment more slowly through the CsCl shelf used in this method, thereby affecting recovery in this size range. Figure 3A, lane 4, shows that in oligo(dT)-selected, poly(A)$^+$ RNA little or no SL RNA can be detected with the oligonucleotide probe. The total and poly(A)$^-$ RNA preparations also contain trace amounts of RNA (approximately 95–103 and 120–125 nucleotides in length), which result from degradation of the SL RNA (Michaeli et al., submitted). However, as shown in Figure 3A, lane 4, these smaller RNA products are removed during the purification of poly(A)$^+$ RNA.

Families of SL RNA–derived fragments clustered around 95–103 and 120–125 nucleotides are common in preparations such as those shown in Figure 3A, as well as in preparations of ribonucleoprotein particles (RNPs) containing the SL RNA (SL RNPs; Michaeli et al., submitted). However, primer extension analysis of RNA molecules of ~100 nucleotides from these and other RNA preparations reveals essentially no authentic 100-mer RNA (unpublished observation). To avoid confusion we therefore used only RNA preparations (such as those shown in Figure 3A) that were depleted or devoid of these molecules for subsequent S1 protection experiments. S1 protection analysis was important in the detection of the 100-mer sequences because one would expect that any 100-mer RNA covalently linked to pre-mRNAs would be dispersed throughout the entire size range of pre-mRNA molecules and thus would not be detected by Northern analysis.

The S1 protection, shown in Figure 3B, was performed on the same RNA preparations as those used in Figure 3A. The experiment demonstrates that a prominent 100 nucleotide species is indeed revealed in each RNA preparation after S1 treatment. This is most clearly observed by comparing the poly(A)$^+$ RNA lanes in each blot. In spite of the fact that neither SL RNA nor 100-mer bands are detected in the Northern blot shown in Figure 3A, lane 4, a 100 nucleotide RNA is protected from digestion by nuclease S1 (Figure 3B, lane 4). This indicates that the 100-mer RNA species is tightly associated with a heterogeneous population of poly(A)$^+$ RNA. While the S1 protection in this experiment was performed using the entire 1.35 kb SL reiteration unit (to provide an internal control for genomic DNA contamination), the blot of protected fragments was probed with labeled DNA prepared from an M13 clone containing only the SL RNA sequences between the RsaI site and the Sau3A site of the SL RNA gene (see Figure 1 and Experimental Procedures). The use of this probe eliminates the possibility that this 100 nucleotide band is derived from elsewhere within the SL RNA repeat.

While the experiments shown above demonstrate that the S1-protected 100-mer RNA copurifies with poly(A)$^+$ RNA, we also detected the 100-mer in poly(A)$^-$ RNA. There are two possible explanations for this observation. The incomplete removal of polyadenylated mRNA and

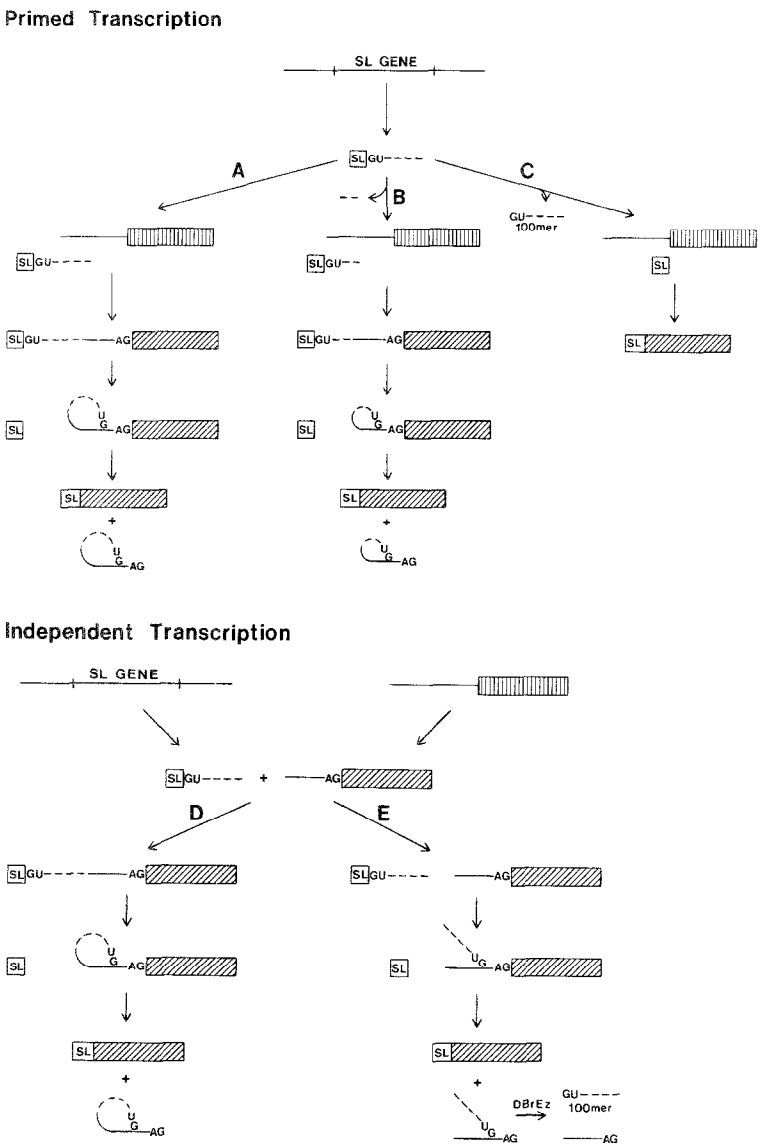

Figure 2. Models of SL Addition to Target mRNAs

Two alternative models of SL RNA and structural gene transcription (primed vs. independent) are suggested. In each case, possible pathways for the formation of mature mRNAs are schematized. The symbols are as follows: (—|—|—), SL RNA gene (i.e., the reiteration unit); SL GU - - - -, the SL RNA (the 100-mer portion includes the consensus GU dinucleotide and downstream portion); ——, structural gene; ——AG, putative pre-mRNA including consensus AG dinucleotide and upstream sequence. Primed transcription. (A) The entire SL RNA is used as a primer for transcription of structural genes resulting in a continuous pre-mRNA, which is spliced in *cis*. (B) The SL RNA is first cleaved, releasing the 3′ end of the SL RNA (- -), and a smaller fragment of the SL RNA is then used as a primer. The resulting RNA is spliced as in (A). (C) The SL RNA is cleaved at the splice junction, releasing a free 100-mer, and the SL itself primes transcription, immediately forming a mature mRNA. Independent transcription. Both the SL RNA and the structural genes are transcribed separately, and the two are ligated head-to-tail to form a continuous pre-mRNA, which is spliced in *cis* (D), or they remain as separate RNA molecules, but are spliced in *trans* (E). (A), (B), and (D) result in formation of a lariat intron structural intermediate. In (E) a discontinuous, Y intron structure is formed. DBrEz, debranching enzyme.

pre-mRNAs from the poly(A)$^-$ RNA would contribute to the 100-mer signal. However, putative introns or intron-like sequences that have been excised during the generation of mature mRNAs, and hence not polyadenylated, may also be present in this preparation. As indicated in Figure 2, these species would also contain 100-mer sequences that would be protected from S1 digestion.

It is possible that the 100-mer species observed by S1 protection may have arisen through an artifactual cleavage of the SL RNA–DNA hybrid in the vicinity of the splice junction. To examine this possibility we performed the control experiments shown in Figure 3C. High molecular weight RNA and SL RNA were isolated by hybrid selection of RNA molecules containing the 100-mer species followed by purification of the respective size classes (greater than 240 bases in length and approximately 125–145 bases) by acrylamide gel electrophoresis. Each purified fraction was subjected to S1 analysis as above. Figure 3C, lane 1, shows that only the 100-mer species is protected from S1 digestion in high molecular weight RNA. In addition, the 125–145 nucleotide RNA fraction reveals the 135 base SL RNA and virtually no 100-mer species (Figure 3C, lane 2). Figure 3C, lane 3, further shows that the 100-mer product does not result from cleavage of the M13 clone used in these protection studies. Thus the 100-mer RNA is truly a unique species associated with high molecular weight RNA.

Besides the S1-protected fragments described above, another product of approximately 139 nucleotides is routinely revealed by S1 protection analysis. The latter forms a doublet with the 135 base SL RNA in total and poly(A)$^-$ RNAs and is also seen in the poly(A)$^+$ RNA from which the 135 base species is absent (Figure 3B, lane 4). One can find similar evidence for the appearance of this doublet in S1 experiments reported in the literature (Campbell et al., 1984). Although we have no clear explanation for the variable ratio of the 135 and 139 base species (compare Figure 3B, lanes 1, 3, and 4), we believe that this doublet

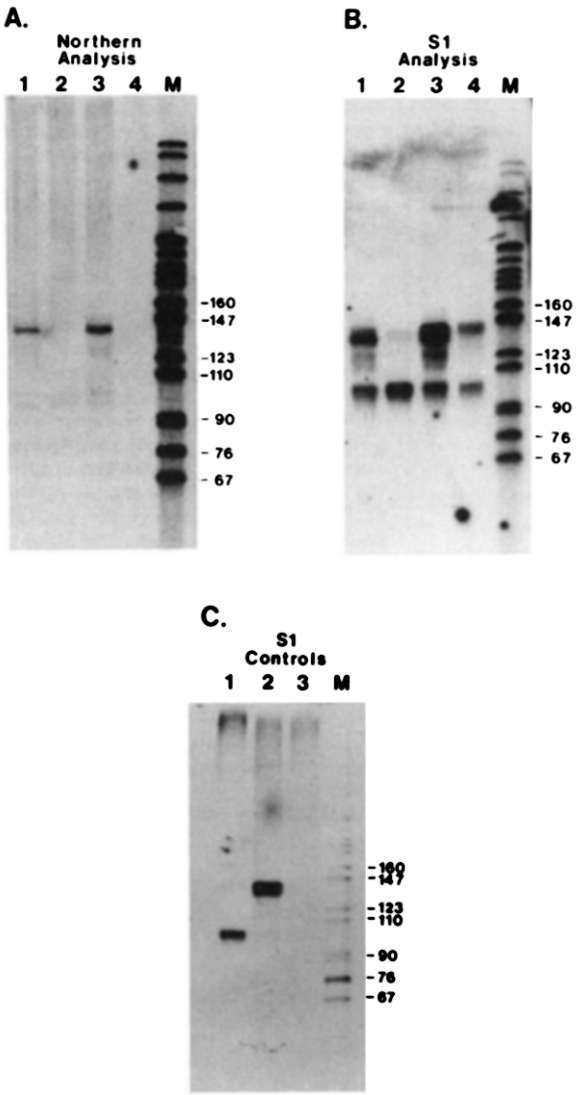

Figure 3. Northern Hybridization and S1 Nuclease Analysis of T. brucei RNA

(A) Northern hybridization analysis. RNA was electrophoresed in 5% acrylamide, 7 M urea gels, blotted, and probed with the 35 base oligonucleotide as described in Experimental Procedures. Markers (M) are 3′-end-labeled, HpaII-digested pBR322 DNA. Lane 1, RNA prepared with LiCl–urea (20 μg, prep #1); lane 2, RNA prepared with guanidine thiocyanate–CsCl (20 μg); lane 3, RNA prepared with LiCl–urea, poly(A)$^-$ fraction (20 μg, prep #2); lane 4, RNA prepared with LiCl–urea, poly(A)$^+$ fraction (6 μg, prep #2). The filter was autoradiographed for 12 days.

(B) S1 nuclease analysis. Equivalent aliquots of the same RNA preparations as used in lanes 1–4 in (A) were hybridized with the complementary, unlabeled M13 SL repeat unit clone, S1 nuclease–treated, blotted, and probed as described in Experimental Procedures.

(C) S1 nuclease controls. RNA containing the 100-mer sequence was hybrid-selected using a single-stranded M13 clone containing the complementary sequence. RNA was electrophoresed on a 5% acrylamide gel and the >240 base and ~135 base RNAs were eluted. These preparations were then analyzed with nuclease S1 as in (B). Lane 1, >240 base RNA containing the 100-mer; lane 2, 135 base RNA; lane 3, no RNA (M13 SL repeat unit DNA only).

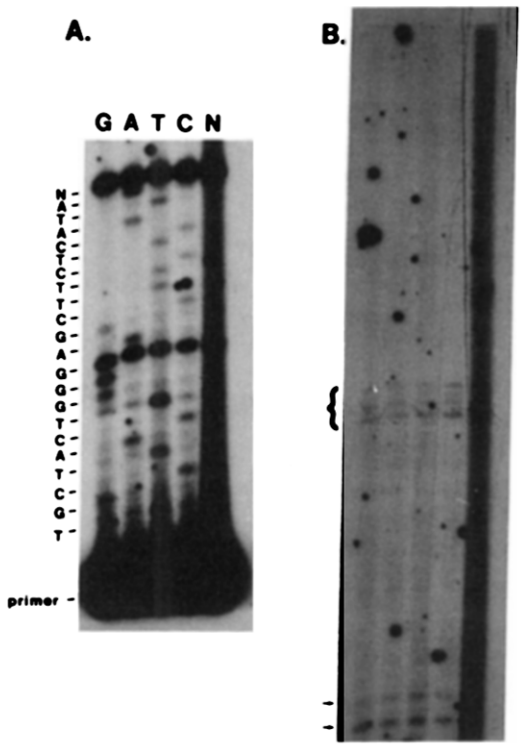

Figure 4. Primer Extension Analysis of 100-mer RNA

Poly(A)$^+$ RNA was hybridized with 5′-end-labeled 35 base oligonucleotide (see Figure 1), divided into five equal aliquots, and extended with reverse transcriptase in the presence or absence of dideoxynucleoside triphosphates as described in Experimental Procedures. (A) Lower portion of autoradiogram. Lane N, primer extension in the absence of dideoxynucleoside triphosphates. Lanes G, A, T, and C correspond to the dideoxynucleoside triphosphate used in that reaction. (B) Upper portion of autoradiogram. The bracket denotes a pentad of bands corresponding to the reported 5′ end of the SL RNA plus 4 bases beyond. The arrows point to two minor extension products 4 and 6 bases upstream from the splice junction. These additional products are described in the text. This portion of the gel has been exposed longer to show these additional bands more clearly.

is an S1 artifact created via the unusual nature of the 5′ cap structure of the SL. We have recently shown that the SL sequence in the SL RNA and mRNAs actually extends an additional 4 bases upstream from the reported 5′ end and that these extra nucleotides are modified components of the SL cap structure (Perry et al., submitted). Reports in the literature of primer extension products of the SL that extend 3 to 5 bases upstream from the accepted 5′ end of the SL (Boothroyd and Cross, 1982; Milhausen et al., 1984; De Lange et al., 1984a; Sather and Agabian, 1985; Tschudi et al., 1985; Gonzalez et al., 1985; also see Figure 4) are consistent with this cap structure. A likely explanation of the proposed SL artifact is that the base modifications of the cap interfere with base-pairing at these nucleotides and can therefore generate S1 cleavage products that are 4 bases shorter. These modifications also appear to inhibit the progress of reverse transcriptase across this portion of the SL (see below).

The appearance of the 139 base S1 protection fragment in Figure 3B, lane 4, and the absence of this band in the Northern blot of this RNA (Figure 3A, lane 4) is somewhat puzzling. It is possible that, like the S1-protected 100-mer, some intact SL RNA is also covalently attached to high molecular weight RNA. This would be expected if there

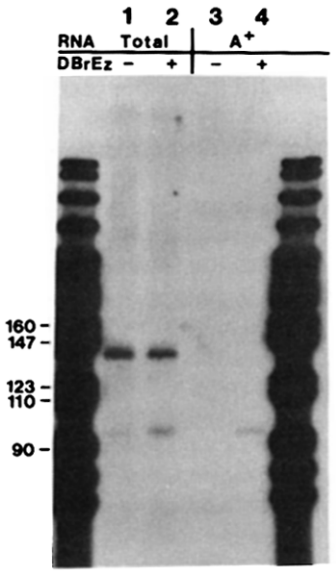

Figure 5. Debranching Analysis of T. brucei RNA

Total LiCl–urea RNA (20 μg, lanes 1 and 2) or poly(A)$^+$ RNA (6 μg, lanes 3 and 4) was incubated in the presence (lanes 2 and 4) or absence (lanes 1 and 3) of debranching enzyme. The resultant products were electrophoresed, blotted, and probed with the 100-mer oligonucleotide (Figure 1) as described in Experimental Procedures.

were some priming of transcription or ligation of SL RNA to structural gene transcripts as illustrated in Figures 2A and 2D, respectively. However, this possibility is unlikely because no 135 or 139 base product is observed in the S1 analysis of purified high molecular weight RNA (Figure 3C, lane 1). Thus, we believe that this product is the result of either an S1 artifact (i.e., a "tribrid" formed between the SL-containing mRNA, which is abundant in the poly(A)$^+$ RNA, the 100-mer RNA, and the DNA probe; see Lopata et al., 1985) or the presence of residual amounts of SL RNA purifying with the poly(A)$^+$ RNA, which is detected with the uniformly labeled S1 probe but not with the oligonucleotide probe (the specific activity of the former is approximately 17-fold greater than that of the later).

An Authentic 100-mer Is Covalently Attached to Trypanosome Poly(A)$^+$ RNA

In concurrent studies of the SL RNA structure, we have characterized SL RNA products that appear to reflect nuclease-sensitive regions of the SL RNA in the vicinity of the conserved splice junction sequence (Michaeli et al., submitted). As noted above, these cleavage products do not precisely terminate at the location of the splice junction predicted from analysis of mature mRNAs and their cognate genes. Therefore, it was crucial to demonstrate that the 100-mer species detected in the experiments described above is indeed derived from cleavage at the authentic SL RNA 5′ splice junction and that this 100-mer is covalently linked to a population of poly(A)-containing RNA.

Primer extension sequencing of poly(A)$^+$ RNA was performed using a 5′-end-labeled, 35 base oligonucleotide that, as shown in Figure 1, is complementary to the 100-mer. In the presence of the four deoxynucleoside triphosphates, an authentic 100-mer should be extended 24 nucleotides and then show a strong stop in all four sequencing reactions. The sequence generated would be complementary to the SL RNA. As shown in Figure 4A, primer extension analysis provides sequence verification that the 100-mer associated with poly(A)$^+$ RNA does, in fact, terminate at the G residue 25 nucleotides from the primer, at the position of the splice junction.

In addition to the strong stop described above, a second strong stop is seen 12 nucleotides earlier (see Figure 4A). This stop occurs at the end of a run of 3 guanosine residues and is probably a reverse transcriptase artifact. Similar misincorporation of nucleoside triphosphates by reverse transcriptase, particularly after short runs of homo-oligonucleotides, has been described (Gopinathan et al., 1979; Hartley et al., 1982; Murphy, 1984). We have also noticed that this second stop is dependent on reverse transcriptase concentration, being reduced at higher enzyme levels (data not shown).

Figure 4B shows the upper portion of the gel depicted in 4A. Two minor extension products of 104 and 106 bases can also be seen in this experiment. These sizes are typical of the ~100 base degradation products of the SL RNA described above. In addition there is a pentad of bands further up on the gel. When similar primer extension products are electrophoresed alongside the sequence of the SL RNA gene one can see that these products map to the 5′ end of the SL RNA and 4 bases beyond (not shown). These bands most likely correspond to the 139 base S1 protection product observed in Figure 3B, lane 4. As pointed out above, reverse transcriptase does not efficiently polymerize DNA across these modified nucleotides, thus the appearance of five discrete bands.

The S1 and primer extension experiments indicate the existence of an RNA structure analogous to the splicing intermediates of other eukaryotes. If the mechanism of RNA joining in trypanosomes is similar to *cis* splicing, one would predict that the 100-mer would be covalently attached to high molecular weight RNA via a 2′-5′ phosphodiester bond, which is readily cleaved by debranching enzyme. Furthermore, the intermediate Y structure, indicated in Figure 2E, would be easily discerned, since the action of debranching enzyme would release free 100-mers, which would be detected by Northern hybridization analysis.

To specifically test for the existence of a Y structure, total and poly(A)$^+$ RNA samples were incubated in the presence and absence of purified debranching enzyme and analyzed by Northern hybridization. The results of this experiment are shown in Figure 5. In the absence of debranching enzyme, total RNA contains the 135 nucleotide SL RNA plus a small amount of 95 nucleotide RNA (Figure 5, lane 1). However, when the total RNA is treated with debranching enzyme, a more intense band of roughly 95 bases (or slightly larger) is generated (Figure 5, lane 2). This increase in intensity is not due to unequal amounts of RNA in the two lanes; the intensity of the SL RNA hybridization is equivalent in the two RNA samples (if not slightly less in lane 2).

Even more striking is the appearance of the 95 nucleotide RNA (Figure 5, lane 4) after digestion of poly(A)$^+$ RNA with debranching enzyme. As shown in Figure 5, lane 3, no hybridization is detected in the absence of enzyme. We believe that this debranched RNA corresponds to the 100-mer species detected by S1 and primer extension analysis. The gel system used in these experiments exhibits an increase in mobility of RNA as compared with DNA fragments of the same size. In support of this is the observation that the DNA fragment that is completely protected from S1 digestion by the SL RNA migrates as a 139 base species, whereas the SL RNA itself migrates as a 135 base species (compare Figures 3A and 3B). Since the purified debranching enzyme used in this experiment specifically cleaves branched 2′-5′ phosphodiester linkages, we propose that the 100-mer portion of the SL RNA is covalently attached to the pre-mRNA via this same linkage. Furthermore, the release of a free 100-mer is consistent with the presence of a novel discontinuous intron or Y branch (Figure 2E), rather than a lariat structure.

Discussion

We have shown that the 100 nucleotide 3′ intron fragment of the SL RNA is associated with a heterogenous population of RNA and is attached to this RNA through a branched 2′-5′ phosphodiester linkage, as judged by release of the 100-mer upon treatment with debranching enzyme. These results demonstrate that the 135 base SL RNA is the donor of the spliced leader sequence to trypanosome mRNAs and support the hypothesis that joining of the SL occurs primarily by a unique intermolecular splicing process. The intermediate structures described in these studies are compatible with a *trans* splicing model (Figure 2E) and imply that the SL RNA and structural genes are transcribed independently and spliced to form mature mRNAs without prior ligation. As a working model, we suggest that *trans* addition of the SL differs mechanistically from splicing of nuclear mRNA introns in other eukaryotic systems primarily in that the introns are discontinuous. Accordingly, the *trans* addition of the SL would proceed via an endonucleolytic cleavage of the SL RNA that precisely separates the SL from its downstream 100-mer. This 5′ cleavage would coincide with the esterification of the 5′ phosphate of the 100-mer to a 2′ hydroxyl of a structural gene pre-mRNA. Resolution of the resultant Y intermediate would occur by joining the free SL to the structural gene transcript at the 3′ splice junction with concomitant release of the discontinuous intron as a Y branch structure (Figure 2E).

Although we have demonstrated the existence of the 100-mer in branched RNAs in trypanosomes, the location of branch points in specific pre-mRNAs remains to be determined. We presume that the 2′-5′ branch is located immediately upstream from the 3′ splice junctions conserved in all trypanosome structural genes characterized to date. Although our search of trypanosome 5′-flanking sequences has not revealed a highly conserved sequence such as the yeast TACTAAC box (Langford and Gallwitz, 1983), sequences that resemble the branch point consensus of higher eukaryotes are present a short distance upstream of the 3′ splice junctions (unpublished observations).

The hypothesis that SL addition occurs in *trans*, and yet is analogous to pre-mRNA splicing in other eukaryotes, is not without precedent. Although splicing of two separate RNAs has not been observed in vivo, the ability of the splicing machinery to carry out this type of reaction has been demonstrated, albeit at low efficiency, in vitro (Solnick, 1985; Konarska et al., 1985); however, the efficiency of in vitro *trans* splicing increases with the introduction of short regions of complementarity at the ends of the respective substrates (Konarska et al., 1985). The 5′-flanking regions of several trypanosome structural genes have been examined for sequences complementary to the SL RNA. While some short regions of SL homology exist in the gene-flanking regions, these are not always present or on the appropriate strand to provide complementarity to the SL RNA and target pre-mRNAs. However, since the 5′ limits of structural gene pre-mRNAs have not been determined, we can not rule out the possibility that complementarity exists and, in fact, contributes to the apparent high efficiency of this reaction. A more interesting possibility is that trypanosomes may have evolved specific modifications of the splicing machinery, allowing for the efficient *trans* splicing of the SL and structural gene transcripts. Evidence for one such novel aspect of the SL RNA that may be relevant to the process of *trans* splicing (i.e., the existence of a specific SL RNA–containing small ribonucleoprotein particle) will be published elsewhere (Michaeli et al., submitted).

What is the purpose of SL joining in trypanosomes? At the very least, because the SL RNA is itself capped, its addition to target mRNAs provides a cap for the mature messenger. Many tandemly organized and reiterated gene families have been described in the trypanosomatid protozoa (Thomashow et al., 1983; Tschudi et al., 1985; Clayton, 1985; Michels et al., 1986; Gonzalez et al., 1985), all of which possess homologous SL RNAs (Nelson et al., 1984; De Lange et al., 1984b) and presumably also use *trans* RNA splicing in mRNA processing. In T. cruzi, Gonzalez et al. (1985) have presented evidence that a polycistronic mRNA is derived from one such gene family, suggesting that these transcripts can be resolved into single, capped mRNAs by *trans* splicing. Because there is little evidence for the conservation of untranslated leader sequences in mRNAs of other eukaryotes, even among genes that are under similar regulatory control in the same tissue, it would be unusual if trypanosomes require the SL sequence per se for some aspect of posttranscriptional mRNA function. In fact, presently no evidence can be brought to bear on the question of any possible function of the 35 nucleotides of the SL sequence. Perhaps the real significance of SL joining lies not within the specific SL sequence, but rather in its participation in *trans* splicing itself. The entire process of splicing provides many sites for regulatory interplay between RNA synthesis and processing and nuclear export that may only be maintained if there are introns to splice, regardless of *cis* or *trans* orientation. The ubiquity of splicing in eukaryotic systems, even in those having very few genes with inter-

vening sequences, such as yeast, suggests a larger role of splicing in the molecular biology of the eukaryotic cell.

The identification of branched RNA splicing intermediates in trypanosomes is only a first step in understanding gene expression in these organisms. For example, very little is known about the process of transcription or the nature of their primary transcripts. Although there is circumstantial evidence that the variant surface glycoprotein (VSG) genes are transcribed as part of larger precursor RNAs (De Lange et al., 1985; Bernards et al., 1985), a continuous transcript spanning the VSG expression site has never been demonstrated. The covalent attachment of the 100-mer to pre-mRNAs should allow for the hybrid selection of these precursors and their subsequent characterization. Precise definition of the substrates of the *trans* splicing reaction will also eliminate some empiricism in developing an in vitro *trans* splicing system. While the demonstration of *trans* splicing itself reveals many fascinating biochemical and evolutionary problems, more importantly its description and experimental sequela finally allow us to approach the regulatory network that activates VSG gene transcription during antigenic variation and trypanosome development.

Experimental Procedures

Growth and Isolation of Trypanosomes

IsTat 1.1 trypanosomes (Milhausen et al., 1983; Parsons et al., 1983) were grown in male rats to a density of 2×10^9 to 3×10^9 cells per milliliter of blood and collected by cardiac puncture. The heparinized blood sample was either used directly (see LiCl–urea RNA preparation, below) or adjusted to 15 mM in Hepes buffer (pH 7.5) and 0.2% in glucose and chilled on ice. After centrifugation at 1,200 × g for 12 min, the buffy coat layer, containing the trypanosomes, was removed and used for guanidine thiocyanate–CsCl RNA preparation described below.

Preparation of Trypanosome RNA

LiCl–Urea RNA Preparation

Infected blood collected by cardiac puncture was immediately mixed with 10 volumes of ice cold 3 M LiCl–6 M urea, homogenized thoroughly in a blender (Auffray and Rougeon, 1980), and precipitated overnight on ice. The RNA was centrifuged at 18,000 × g at 0°C for 60 min and suspended in buffer containing 25 mM Tris–HCl (pH 8.0), 100 mM NaCl, 5 mM EDTA, 1% SDS, and 200 μg/ml proteinase K. After digestion for 1 hr at 37°C, the RNA was extracted twice with phenol–$CHCl_3$ and twice with ether. The solution was adjusted to 0.3 M in NaOAc (pH 5.6) and precipitated with 3 volumes of absolute ethanol. After centrifugation, the RNA pellet was suspended in RNAase-free H_2O, precipitated twice from 3 M NaOAc (pH 6.0) to remove contaminating oligodeoxyribonucleotides (Schrader and O'Malley, 1981), and stored as an ethanol precipitate (as above) at −70°C. Poly(A)$^+$ RNA was isolated from a preparation of LiCl–urea RNA as described by Aviv and Leder (1972), with the modifications suggested by Chirgwin et al. (1979), and stored as described above.

Guanidine Thiocyanate–CsCl RNA Preparation

RNA was isolated essentially as described by Chirgwin et al. (1979). Buffy coat blood preparations containing trypanosomes were immediately mixed with 2 volumes of 6 M guanidine thiocyanate, 75 mM Tris–HCl, 15 mM EDTA, 0.75% sarkosyl, and 1.1% beta-mercaptoethanol and vortexed to disperse the cells completely. The mixture was layered over 5 ml of 5.8 M CsCl and 0.1 M Na_2EDTA and spun at 29,000 rpm in the Beckman SW41 rotor for 24 hr. The RNA pellet was suspended in RNAase-free H_2O and precipitated by addition of 0.1 volume of 3 M NaOAc (pH 5.6) and 3 volumes of absolute ethanol. Sedimented RNA was washed with 70% ethanol, dried, suspended in buffer containing 40 mM Tris–HCl (pH 7.9), 10 mM NaCl, and 6 mM $MgCl_2$, and digested with RNAase-free DNAase I (Promega Biotech). This sample was then extracted with phenol–$CHCl_3$, washed with ether, and precipitated twice from 3 M NaOAc (pH 6.0). The RNA was stored as an ethanol precipitate at −70°C (see above).

S1 Nuclease Analysis

S1 nuclease analysis was performed essentially as described by Maniatis et al. (1982). A single SL RNA repeat unit, generated by cleavage with the restriction enzyme Sau3A, was subcloned from the plasmid pTb1.4-1 (Nelson et al., 1983) into the BamHI site of M13mp19, and a phage containing the complementary strand of the SL RNA sequence was isolated. RNA samples were mixed with 25 ng of this M13 clone, and, when necessary, yeast carrier transfer RNA was added to make a total of 20 μg of RNA. The nucleic acids were precipitated with ethanol, washed with 70% ethanol, dried, and suspended in 10 μl of 80% formamide, 400 mM NaCl, 40 mM PIPES buffer (pH 6.4), and 1 mM EDTA. Samples were heated to 85°C for 15 min and incubated at 56°C for 3 hr. Three hundred microliters of 280 mM NaCl, 50 mM NaOAc (pH 4.6), 4.5 mM $ZnSO_4$, 20 μg/ml single-stranded DNA, and 1,000 U/ml S1 nuclease (BRL) were added to each sample and incubated at 30°C for 30 min. Reactions were stopped by addition of 75 μl of 2.5 M NH_4OAc, 50 mM EDTA (pH 8.0), 10 μg of carrier tRNA, and an equal volume of isopropyl alcohol.

Hybridization Analysis

RNA or S1-protected DNA fragments were centrifuged, washed with 70% ethanol, and dried. Samples were suspended in a small volume of 90% formamide, 1 mM EDTA, 0.5% xylene cyanole, and 0.5% bromphenol blue, then boiled for 5 min and quickly chilled on ice. After electrophoresis in 5% acrylamide (acrylamide:bis, 20:1), 7 M urea gels, the samples were electro-blotted onto Nytran filter membranes (Schleicher and Schuell) according to the manufacturers protocols. The filters were then baked at 80°C for 2 hr in a vacuum oven.

RNA blots were probed with a 35 base oligonucleotide complementary to the region beginning 25 bases downstream from the SL (see Figure 1). The oligonucleotide was labeled with polynucleotide kinase (Pharmacia P-L) and [γ-^{32}P]ATP as described by Maniatis et al. (1982). The probe for the S1 blots was made essentially as described by Hu and Messing (1982). An M13 clone (150 ng) containing the DNA strand complementary to the SL RNA and between the RsaI and Sau3A sites (see Figure 1) was isolated (M13RS−) and hybridized with 0.5 pmol of the −20 region sequencing primer (New England Biolabs). The primer was extended with Klenow polymerase (Pharmacia P-L) in the presence of 100 mM NaCl, 10 mM Tris–HCl (pH 7.5), 7 mM $MgCl_2$, 5 mM dithiothreitol, 100 μg/ml bovine serum albumin, 0.25 mCi each of α-^{32}P-labeled dATP, dCTP, dGTP, and dTTP (dried; New England Nuclear), and 0.8 μM each of cold dATP, dCTP, dGTP, and dTTP for 1 hr at 25°C. The reaction was stopped by the addition of an equal volume of the formamide dye mixture described above, and the DNA was denatured by boiling 5 min. After electrophoresis in a 5% acrylamide, 7 M urea gel (as above), the appropriate sized DNA fragments (i.e., fragments larger than the distance from the primer to the opposite side of the insert fragment) were excised from the gel and electro-eluted.

Filters were prehybridized in an excess of 5× SSC (750 mM NaCl and 75 mM sodium citrate), 10× Denhardt's solution (0.2% each of Ficoll, polyvinylpyrrolidone, and BSA), 0.2% SDS, 0.1% sodium pyrophosphate, 100 μg/ml single-stranded, sheared salmon sperm DNA, 100 μg/ml tRNA, and 1 mM ATP (for blots probed with labeled oligonucleotide) for 4 hr or longer. The filters were hybridized with 50 μl/cm^2 of the same solution containing 5×10^6 cpm of probe for 24 hr at 42°C (oligonucleotide probe) or at 65°C (uniformly labeled probe) and washed three times, for 30 min each, in either 3× SSC and 0.1% SDS at 55°C (oligonucleotide probe) or 1× SSC, 0.1% SDS at 65°C (uniformly labeled probe). Filters were then autoradiographed.

Primer Extension Analysis

The 35 base oligonucleotide primer was labeled as above and gel-purified on a 20% sequencing gel to remove minor contaminating bands. Poly(A)$^+$ RNA (50 μg) and 1×10^7 cpm of primer were mixed, ethanol-precipitated, washed with 70% ethanol, and dried. The nucleic acid was suspended in 27.5 μl of H_2O and 5 μl of 200 mM sodium pyrophosphate; the sample was heated to 90°C for 2 min followed by 10 min at 37°C. Five microliters of RNasin (40 U/μl; Promega Biotec)

and 25 μl of 200 mM dithiothreitol were added, and the sample was divided into five equal aliquots. Ten microliters of 5× buffer (250 mM Tris–HCl [pH 8.3], 50 mM $MgCl_2$, and 700 mM KCl) was added to each tube. One tube then received 25 μl of 2 mM dATP, dCTP, dGTP, and dTTP. Each of the other four tubes received 25 μl of the appropriate dideoxy:deoxynucleoside triphosphate mixture (ddA:dA, 30 μM:8 μM; ddC:dC, 45 μM:12 μM; ddG:dG, 66 μM:16 μM; ddT:dT, 117 μM:24 μM). The reactions were started by the addition of 2.5 μl of AMV reverse transcriptase (30 U/μl; Amersham). Samples were incubated at 42°C for 30 min, then chased by adding 4 μl of 5 mM dATP, dCTP, dGTP, and dTTP and incubating for an additional 15 min at 42°C. The reactions were stopped by the addition of 13 μl of 2.5 M NH_4OAc, 50 mM EDTA (pH 8.0), and an equal volume of isopropyl alcohol. The samples were centrifuged, washed in 70% ethanol, dried, and suspended in the formamide dye buffer described above. The samples were boiled 5 min and quickly chilled on ice. Electrophoresis was carried out in 5% acrylamide (acrylamide:bis, 20:1), 7 M urea sequencing gels at 2,000 V until the bromphenol blue reached the bottom of the gel. The gel was dried and autoradiographed.

Isolation of RNA Containing the 100-mer Species

Hybridization selection of RNA containing the 100-mer species was performed essentially as described by Maniatis et al. (1982). Nitrocellulose filters containing 20 μg of M13RS− DNA (see above) were prepared according to Kafatos et al. (1979) and used in the hybridization reaction. After elution of bound RNA from filters, the RNA was precipitated by addition of 0.1 volume of 3 M NaOAc (pH 5.2) and 3 volumes of ethanol. The precipitated RNA was electrophoresed on a 5% acrylamide, 7 M urea gel alongside end-labeled, HpaII-digested pBR322 markers. Two sections of the gel containing approximately >240 base RNA and approximately 125–145 base RNA were removed, crushed, and soaked in a solution containing 0.75 M ammonium acetate, 10 mM magnesium acetate, 2% phenol, 0.1% SDS, 0.1 mM EDTA, and 25 μg/ml tRNA at 37°C overnight (Frendewey and Keller, 1985). The eluted RNA was phenol-extracted, concentrated with 2-butanol, and ethanol-precipitated. An aliquot of each eluted fraction was analyzed with nuclease S1 as described above.

Debranching Enzyme Digests

The debranching enzyme preparation used here had been purified 700-fold and has a specific activity of 9.6×10^6 U/mg protein (where 1 U is the amount of enzyme that linearizes 1 fmol of lariat RNA in 30 min at 30°C). The enzyme preparation shows no evidence of contamination by other endonucleases, exonucleases, phosphatases, or nonbranched 2′-5′ phosphodiesterase activities. RNA samples were digested in 20 μl containing 20 mM Hepes (pH 7.0), 3 mM $MgCl_2$, 0.25 mg/ml bovine serum albumin, and 240 U of enzyme for 30 min at 30°C. Samples were extracted with phenol–$CHCl_3$ first, then with ether and precipitated with 0.1 volume of 3 M NaOAc (pH 5.6) and 3 volumes of absolute ethanol. The final RNA pellets were washed with 70% ethanol and dried. After suspending in formamide dye buffer, the samples were denatured, electrophoresed, blotted, and probed with the 35 base oligonucleotide as described above.

Acknowledgments

The authors are especially grateful to Dr. Jaime Arenas and Dr. Jerard Hurwitz of the Sloan-Kettering Institute for the generous gift of debranching enzyme and to Dr. Phillip Barr of Chiron Corporation for the generous gift of the oligonucleotide. We also thank Drs. Richard Nelson, Karen Perry, Shulamit Michaeli, and Judith Levin and our other colleagues for helpful discussions and critical reading of this manuscript. This work was supported in part by NIH grant AI-21975, by the John D. and Catherine T. MacArthur Foundation, and by the Office of Naval Research. N. A. is a Burroughs Wellcome Scholar of Molecular Parasitology.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 30, 1986; revised September 9, 1986.

References

Auffray, C., and Rougeon, F. (1980). Purification of mouse immunoglobulin heavy-chain RNAs from total myeloma tumor RNA. Eur. J. Biochem. *107*, 303–314.

Aviv, H., and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA *69*, 1408–1412.

Bernards, A., De Lange, T., Michels, P. A. M., Liu, A. Y. C., Huisman, M. J., and Borst, P. (1984). Two modes of activation of a single surface antigen gene of Trypanosoma brucei. Cell *36*, 163–170.

Bernards, A., Kooter, J. M., and Borst, P. (1985). Structure and transcription of a telomeric surface antigen gene of *Trypanosoma brucei*. Mol. Cell. Biol. *5*, 545–553.

Boothroyd, J. C., and Cross, G. A. M. (1982). Transcripts coding for the variant surface glycoproteins of *Trypanosoma brucei* have a short, identical exon at their 5′ end. Gene *20*, 281–289.

Campbell, D. A., Thornton, D. A., and Boothroyd, J. C. (1984). Apparent discontinuous transcription of *Trypanosoma brucei* variant surface antigen genes. Nature *311*, 350–355.

Cheng, S.-C., and Abelson, J. (1986). Fractionation and characterization of a yeast mRNA splicing extract. Proc. Natl. Acad. Sci. USA *83*, 2387–2391.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry *18*, 5294–5299.

Clayton, C. E. (1985). Structure and regulated expression of genes ecoding fructose biphosphate aldolase in *Trypanosoma brucei*. EMBO J. *4*, 2997–3003.

Cully, D. F., Ip, H. S., and Cross, G. A. M. (1985). Coordinate transcription of variant surface glycoprotein genes and expression site associated gene family in Trypanosoma brucei. Cell *42*, 173–182.

De Lange, T., Liu, A. Y. C., Van der Ploeg, L. H. T., Borst, P., Tromp, M. C., and Van Boom, J. H. (1983). Tandem repetition of the 5′ miniexon of variant surface glycoprotein genes: a multiple promoter for VSG gene transcription? Cell *34*, 891–900.

De Lange, T., Michels, P. A. M., Veerman, H. J. G., Cornelissen, A. W. C. A., and Borst, P. (1984a). Many trypanosome mRNAs share a common 5′ terminal sequence. Nucl. Acids Res. *12*, 3777–3790.

De Lange, T., Berkvens, T. M., Veerman, H. J. G., Frasch, A. C. A., Barry, J. D., and Borst, P. (1948b). Comparison of the genes coding for the common 5′ terminal sequence of messenger RNAs in three trypanosome species. Nucl. Acids Res. *12*, 4431–4443.

De Lange, T., Kooter, J. M., Luirink, J., and Borst, P. (1985). Transcription of a transposed trypanosome surface antigen gene starts upstream of the transposed segment. EMBO J. *4*, 3299–3306.

Frendewey, D., and Keller, W. (1985). Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. Cell *42*, 355–367.

Gonzalez, A., Lerner, T. J., Huecas, M., Sosa-Pineda, B., Nogueira, N., and Lizardi, P. M. (1985). Apparent generation of a segmented mRNA from two separate tandem gene families in *Trypanosoma cruzi*. Nucl. Acids Res. *13*, 5789–5804.

Gopinathan, K. P., Weynouth, L. A., Kunkel, T. A., and Loev, L. A. (1979). Mutagenesis in vitro by DNA polymerase from an RNA tumor virus. Nature *278*, 857–859.

Hardy, S. F., Grabowski, P. J., Padgett, R. A., and Sharp, P. A. (1984). Cofactor requirements of splicing of purified messenger RNA precursors. Nature *308*, 375–377.

Hartley, J. L., Chen, K. K., and Donelson, J. E. (1982). A mercury-thiol affinity system for rapid generation of overlapping labelled DNA fragments for sequence analysis. Nucl. Acids Res. *10*, 4009–4025.

Hu, N.-t., and Messing, J. (1982). The making of strand-specific M13 probes. Gene *17*, 271–277.

Kafatos, F. C., Jones, C. W., and Efstratiadis, A. (1979). Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucl. Acids Res. *7*, 1541–1552.

Keller, W., and Noon, W. A. (1984). Intron splicing: a conserved internal

signal in introns of animal pre-mRNAs. Proc. Natl. Acad. Sci. USA *81*, 7417–7420.

Konarska, M. M., Padgett, R. A., and Sharp, P. A. (1985). *Trans* splicing of mRNA precursors in vitro. Cell *42*, 165–171.

Kooter, J. M., De Lange, T., and Borst, P. (1984). Discontinuous synthesis of mRNA in trypanosomes. EMBO J. *3*, 2387–2392.

Krainer, A. R., Maniatis, T., Ruskin, B., and Green, M. R. (1984). Normal and mutant human β-globin pre-mRNAs are faithfully and efficiently spliced in vitro. Cell *36*, 993–1005.

Langford, C. J., and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell *33*, 519–527.

Lin, R.-J., Newman, A. J., Cheng, S.-C., and Abelson, J. (1985). Yeast mRNA splicing in vitro. J. Biol. Chem. *260*, 14780–14792.

Lopata, M. A., Sollner-Webb, B., and Cleveland, D. W. (1985). Surprising S1-resistant trimolecular hybrids: potential complication in interpretation of S1 mapping analyses. Mol. Cell. Biol. *5*, 2842–2846.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Michels, P. A. M., Poliszczak, A., Osinga, K. A., Misset, O., Van Beeumen, J., Wierenga, R. K., Borst, P., and Opperdoes, F. R. (1986). Two tandemly linked genes code for the glycosomal glyceraldehyde-phosphate dehydrogenase in *Trypanosoma brucei*. EMBO J. *5*, 1049–1056.

Michiels, F., Matthyssens, G., Kronenberger, P., Pays, E., Dero, B., Van Assel, S., Darvill, M., Cravador, A., Steinert, M., and Hamers, R. (1983). Gene activation and re-expression of a *Trypanosoma brucei* variant surface glycoprotein. EMBO J. *2*, 1185–1192.

Milhausen, M., Nelson, R. G., Parsons, M., Newport, G., Stuart, K., and Agabian, N. (1983). Molecular characterization of initial variants from the IsTat I serodeme of *Trypanosoma brucei*. Mol. Biochem. Parasitol. *9*, 241–254.

Milhausen, M., Nelson, R. G., Sather, S., Selkirk, M., and Agabian, N. (1984). Identification of a small RNA containing the trypanosome spliced leader: a donor of shared 5′ sequences of trypanosomatid mRNAs? Cell *38*, 721–729.

Mount, S. M. (1982). A catalogue of splice junction sequences. Nucl. Acids Res. *10*, 459–472.

Murphy, W. J. (1984). Analysis of the duplication and transposition of a variant surface glycoprotein gene of trypanosomes. Ph.D. thesis, University of Iowa, Iowa City, Iowa.

Murphy, W. J., Brentano, S. T., Rice-Ficht, A. C., Dorfman, D. M., and Donelson, J. E. (1984). DNA rearrangements of the variable surface antigen genes of the trypanosomes. J. Protozool. *31*, 65–73.

Nelson, R. G., Parsons, M., Barr, P. J., Stuart, K., Selkirk, M., and Agabian, N. (1983). Sequences homologous to the variant antigen mRNA spliced leader are located in tandem repeats and variable orphons in Trypanosoma brucei. Cell *34*, 901–909.

Nelson, R. G., Parsons, M., Selkirk, M., Newport, G., Barr, P. J., and Agabian, N. (1984). Sequences homologous to variant antigen mRNA spliced leader in Trypanosomatidae which do not undergo antigenic variation. Nature *308*, 665–667.

Newman, A. J., Lin, R., Cheng, S., and Abelson, J. (1985). Molecular consequences of specific intron mutations on yeast mRNA splicing in vivo and in vitro. Cell *42*, 335–344.

Padgett, R. A., Konarska, M. M., Grabowski, P. J., Hardy, S. F., and Sharp, P. A. (1984). Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. Science *225*, 898–903.

Parsons, M., Nelson, R. G., Newport, G., Milhausen, M., Stuart, K., and Agabian, N. (1983). Genomic organization of *Trypanosoma brucei* variant antigen gene families in sequential parasitemias. Mol. Biochem. Parasitol. *9*, 255–269.

Parsons, M., Nelson, R. G., Watkins, K. P., and Agabian, N. (1984). Trypanosome mRNAs share a common 5′ spliced leader sequence. Cell *38*, 309–316.

Perkins, K. K., Furneaux, H. M., and Hurwitz, J. (1986). RNA splicing products formed with isolated fractions from HeLa cells are associated with fast-sedimenting complexes. Proc. Natl. Acad. Sci. USA *83*, 887–891.

Pikielny, C. W., Teem, J. L., and Rosbash, M. (1983). Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implications for U1 RNA and metazoan mRNA splicing. Cell *34*, 395–403.

Reed, R., and Maniatis, T. (1985). Intron sequences involved in lariat formation during pre-mRNA splicing. Cell *41*, 95–105.

Ruskin, B., and Green, M. R. (1985). An RNA processing activity that debranches RNA lariats. Science *229*, 135–140.

Ruskin, B., Krainer, A. R., Maniatis, T., and Green, M. R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell *38*, 317–331.

Ruskin, B., Greene, J. M., and Green, M. R. (1985). Cryptic branch point activation allows accurate in vitro splicing of human β-globin intron mutants. Cell *41*, 833–844.

Sather, S., and Agabian, N. (1985). A 5′ spliced leader is added in trans to both α- and β-tubulin transcripts in *Trypanosoma brucei*. Proc. Natl. Acad. Sci. USA *82*, 5695–5699.

Schrader, W. T., and O'Malley, B. T. (1981). Laboratory Methods Manual for Hormone Action and Molecular Endocrinology, fifth edition. (Houston, Texas: Biological Associates).

Solnick, D. (1985). *Trans* splicing of mRNA precursors. Cell *42*, 157–164.

Swinkels, B. W., Gibson, W. C., Osinga, K. A., Kramer, R., Veeneman, G. H., Van Boom, J. H., and Borst, P. (1986). Characterization of the gene for the microbody (glycosomal) triosephosphate isomerase of *Trypanosoma brucei*. EMBO J. *5*, 1291–1298.

Teem, J. L., Abovich, N., Kaufer, N. F., Schwindinger, W. F., Warner, J. R., Levy, A., Woolford, J., Leer, R. J., van Raamsdonk-Duin, M. M. C., Mager, W. H., Planta, R. J., Schultz, L., Friesen, J. D., Fried, H., and Rosbash, M. (1984). A comparison of yeast ribosomal protein gene DNA sequences. Nucl. Acids Res. *12*, 8295–8312.

Thomashow, L. S., Milhausen, M., Rutter, W. J., and Agabian, N. (1983). Tubulin genes are tandemly linked and clustered in the genome of Trypanosoma brucei. Cell *32*, 35–43.

Tschudi, C., Young, A. S., Ruben, L., Patton, C. L., and Richards, F. F. (1985). Calmodulin genes in trypanosomes are tandemly repeated and produce multiple mRNAs with a common 5′ leader sequence. Proc. Natl. Acad. Sci. USA *82*, 3998–4002.

Van der Ploeg, L. H. T., Liu, A. Y. C., Michels, P. A. M., De Lange, T., Borst, P., Majumder, H. K., Weber, H., Veeneman, G. H., and Van Boom, J. (1982). RNA splicing is required to make the messenger RNA for variant surface antigen in trypanosomes. Nucl. Acids Res. *10*, 3591–3604.

Zeitlin, S., and Efstratiadis, A. (1984). In vivo splicing products of the rabbit β-globin pre-mRNA. Cell *39*, 589–602.