MOLECULAR AND CELLULAR BIOLOGY, June 1988, p. 2597–2603
0270-7306/88/062597-07$02.00/0

Vol. 8, No. 6

trans Splicing in *Leishmania enriettii* and Identification of Ribonucleoprotein Complexes Containing the Spliced Leader and U2 Equivalent RNAs

SAMUEL I. MILLER† AND DYANN F. WIRTH*

Department of Tropical Public Health, Harvard School of Public Health, Boston, Massachusetts 02115

Received 5 October 1987/Accepted 17 March 1988

The 5′ ends of *Leishmania* mRNAs contain an identical 35-nucleotide sequence termed the spliced leader (SL) or 5′ mini-exon. The SL sequence is at the 5′ end of an 85-nucleotide primary transcript that contains a consensus eucaryotic 5′ intron-exon splice junction immediately 3′ to the SL. The SL is added to protein-coding genes immediately 3′ to a consensus eucaryotic 3′ intron-exon splice junction. Our previous work demonstrated possible intermediates in discontinuous mRNA processing that contain the 50 nucleotides of the SL primary transcript 3′ to the SL, the SL intron sequence (SLIS). These RNAs have a 5′ terminus at the splice junction of the SL and the SLIS. We examined a *Leishmania* nuclear extract for these RNAs in ribonucleoprotein (RNP) particles. Density centrifugation analysis showed that the SL RNA is predominately in RNP complexes at 60S, while the SLIS-containing RNAs are in complexes at 40S. We also demonstrated that the SLIS can be released from polyadenylated RNA by incubation with a HeLa cell extract containing debranching enzymatic activity. These data suggested that *Leishmania enriettii* mRNAs are assembled by bimolecular or *trans* splicing as has been recently demonstrated for *Trypanosoma brucei*. Furthermore, we determined the partial sequence of the *Leishmania* U2 equivalent RNA and demonstrated that it cosediments with the SL RNA at 60S in a nuclear extract. These RNP particles may be analogous to so-called spliceosomes that have been demonstrated in other systems.

The assembly of RNAs in *Leishmania enriettii* (22, 25), as in African trypanosomes, is unique in that it involves discontinuous transcription (3, 5, 7, 18, 24). All RNAs examined in this pathogen, as in all kinetoplastida, contain at their 5′ end a conserved 35-nucleotide sequence termed the spliced leader (SL) or 5′ mini-exon (4, 9, 16, 22, 25, 28, 32, 33). The SL genomic sequence is transcribed at the 5′ end of an 85-nucleotide RNA. The presence of consensus splice junctions in the SL RNA and genomic sequences 5′ to SL acceptor sites in protein-coding genes suggests that splicing of the products of two transcription events is involved in the generation of mRNA.

Different models have been proposed to explain how this occurs in vivo (3, 5, 7, 24). They can be placed into two major groups based on whether the presumed splicing machinery of the cell utilizes one or two molecular species as a substrate. These models have been termed *cis* and *trans* splicing. *cis* splicing, which could occur by transcription reinitiation (priming) or ligation of two mRNA precursors, is analogous to conventional pre-mRNA splicing, as has been well documented for mammalian cells and yeasts (3, 5) (Fig. 1A). *trans* splicing implies that the substrate for the splicing machinery involves an intermolecular splice of two molecular species. In this model, branch formation between the 5′ guanine of the SL intron sequence (SLIS) and pre-mRNA would generate a forked rather than a lariat intermediate (Fig. 1B).

Previous work in our laboratory defined putative intermediates in discontinuous transcription that contained the 3′ end of the SL transcript or SLIS at their 5′ ends (25). These RNA species were enriched relative to the SL 85-nucleotide transcript in guanidinium-extracted RNA and contained a 5′ terminus at the consensus splice junction located between the SL and the SLIS. We proposed that these SLIS-containing RNAs might be intermediaries in RNA processing. Therefore, we investigated the sedimentation of these RNAs in a nuclear extract to show their association in ribonucleoprotein (RNP) complexes. We now report the identification of RNP complexes in *L. enriettii*, a kinetoplastida that has discontinuous transcription. Complexes at 40S contained the SLIS-containing RNAs, and other complexes at 60S contained the SL RNA. The U2 equivalent RNA cosedimented in complexes at 60S with the SL RNA.

We also demonstrated that, as in African trypanosomes (26, 31), the SLIS is released from *Leishmania* RNA by a HeLa cell protein extract rich in debranching enzyme. This suggested that the SLIS is covalently linked to the protein-coding portion of pre-mRNA through a branch structure (Fig. 1B). This is what would be predicted by the *trans* splicing model of SL addition and extends support to the *trans* splicing model for *Leishmania* species (26, 31).

We also report the identification of the U2 equivalent transcript in *L. enriettii*, the sequence of its 5′ terminus, and complementarity with the SL. Additionally, by density gradient centrifugation of nuclear extracts, we showed that this molecule cosediments at 60S with the SL RNA. This 60S fraction contains RNA and protein and could contain a structure analogous to the spliceosomes (6, 12) described in well-characterized in vitro systems.

MATERIALS AND METHODS

***L. enriettii* cells and RNA purification.** Cells of *L. enriettii* were grown in Schneider drosophila medium as previously described (25). RNA was purified by either phenol-chloroform extraction or guanidinium isothiocyanate extraction as

* Corresponding author.
† Present address: Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115.

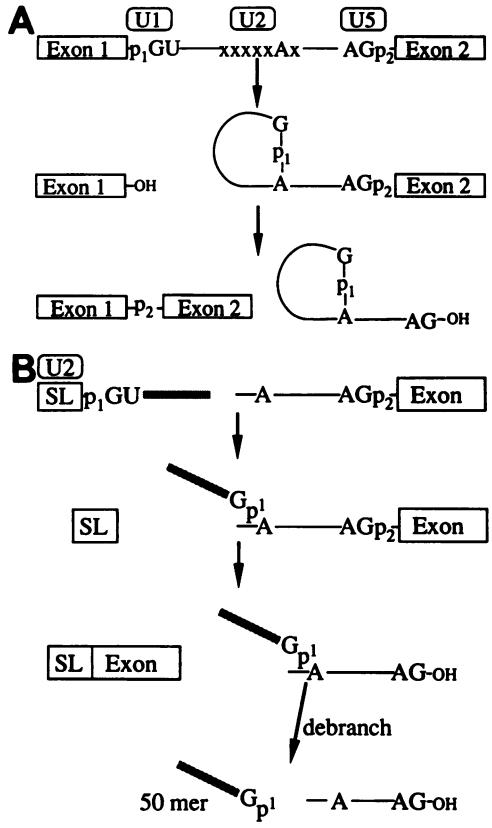

FIG. 1. (A) Conventional eucaryotic pre-mRNA splicing (*cis* splicing). A single transcription event synthesizes a pre-mRNA containing two exons and an intron. The 5′ splice site (GU), the intron branch sequence (xxxxxAx), and the 3′ splice site (AG) are indicated. snRNPs are shown interacting with the 5′ splice site (U1), the intron branch sequence (U2), and the 3′ splice site (U5). After cleavage at the 5′ splice site, a lariat intermediate is generated by formation of a 2′-5′ phosphodiester bond between the guanine at the 5′ end of the intron and an adenine in the branch sequence. After cleavage at the 3′ splice site and exon ligation, a free lariat intermediate is generated. (B) *Leishmania* mRNA *trans* splicing. Two transcription events generate the SL 85-nucleotide transcript and protein-coding pre-mRNA. The SL transcript is shown with the 35-nucleotide SL (in box) as well as the 50-nucleotide SLIS (cross hatched). The U2 equivalent snRNP is shown interacting with the SL exon. Cleavage of the SL transcript at the SL-SLIS junction generates the free SL and a forked structure with the SLIS linked by a branched covalent linkage to the protein-coding component of pre-mRNA. Cleavage at the SL consensus acceptor site and ligation generates mRNA and a free forked structure. Treatment of the forked or Y structures with debranching enzyme releases the free 50-nucleotide SLIS.

detailed elsewhere (25). Polyadenylated RNA was purified on oligo(dT)-cellulose type 3 (Collaborative Research, Inc., Waltham, Mass.) as previously described (1, 25).

RNA analysis by electrophoresis in polyacrylamide gels and blot hybridization. RNA was size fractionated in 15% polyacrylamide–8 M urea denaturing gels, transferred to Gene Screen Plus transfer membrane (NEN Du Pont Co., Wilmington, Del.), and probed with the radiolabeled SL gene sequence as previously described (25). The U2 RNA probe was prepared by using a plasmid that contains the *Trypanosoma brucei* U2 equivalent sequence cloned into an SP6 plasmid vector such that SP6 polymerase transcription generates antisense RNA complementary to the U2 equivalent RNA. The plasmid pSP65R3 was generously provided by C. Tschudi of Yale University; it contains a 375-nucleotide DNA fragment that contains the *T. brucei* U2-coding region. Transcription reactions with SP6 polymerase (NEN Du Pont) were performed by the method of Melton et al. (23). Filters so treated with radiolabeled RNA probes were treated as previously described for DNA probes (25) except that high-temperature washes were performed at 65°C. Filters were air dried and exposed to Kodak XAR-5 film with an intensifying screen.

Primer extension analysis of *Leishmania* RNA. Primer extension of the SL gene was with a 15-nucleotide primer previously described and utilized to delineate the 5′ end of the 85-nucleotide SL transcript as well as the SLIS-containing transcription intermediates (25). The U2 primer extension analysis was done with a primer that is complementary to positions 45 to 60 of the *T. brucei* U2 equivalent. This primer was generated on a Du Pont oligonucleotide synthesizing machine. All primer extension reactions were performed with reverse transcriptase and were analyzed on denaturing polyacrylamide gels as follows. Primers and RNA were annealed at 60°C in 0.4 M NaCl–10 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] (pH 6.4)–1 mM EDTA for 2 h. The hybrids were precipitated by the addition of 2 volumes of ethanol and extended with avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) as previously described (25). Primer extension sequencing was done with a primer complementary to nucleotides 53 to 67 of the *T. brucei* U2 equivalent transcript as previously described (25). Ratios of dideoxynucleotide triphosphates to deoxynucleotide triphosphates were 1:1, and 100 ng of primer was hybridized to 12.5 μg of RNA [poly(A)$^-$ fraction] at 60°C.

Debranching reactions of *L. enriettii* RNA. The HeLa cell S100 (100,000 × *g* supernatant) extract was a generous gift of Michael Green and the members of his laboratory (29). RNA was incubated for 30 min in a 25-μl reaction mixture containing 5 mM EDTA, 20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) (pH 7.5), 20 mM NaCl, and 5 μl of S100 extract. The samples were diluted into PK buffer (1% sodium dodecyl sulfate, 12.5 mM EDTA, 150 mM NaCl, 0.1 M Tris hydrochloride [pH 7.5]) and treated with 30 μg of proteinase K for 30 min. After digestion with proteinase K, the reaction mixtures were extracted with equal volumes of phenol and chloroform, ethanol precipitated, and analyzed in a 15% polyacrylamide–8 M urea gel. RNA in gels was transferred to a Gene Screen Plus hybridization filter and probed with the SL gene as described above.

Density gradient centrifugation of soluble nuclear extract and RNA analysis of gradient fractions. Crude *Leishmania* nuclei were prepared as follows. *L. enriettii* cells were grown to 10^6 cells per ml, and 500 ml of cells was harvested for the preparation of crude nuclei. These cells were chilled to 4°C and then pelleted in a Beckman JA10 rotor at 2,500 rpm for 10 min. Cells were suspended and washed twice in phosphate-buffered saline with 10 mg of glucose per ml. The cells were pelleted in a Sorvall RT6000 clinical centrifuge at 1,500 rpm. They were suspended and swollen in 10 ml of Dignam buffer A (10) and kept at 4°C for 10 min. Nonidet P-40 detergent was added to a concentration of 0.5%, and the cells were incubated at 4°C for 10 min. The crude nuclei were pelleted at 3,000 rpm in a Sorvall RT6000 centrifuge and suspended in 1 ml of a buffer containing 20 mM HEPES (pH 7.5), 20 mM KCl, and 1 mM $MgCl_2$ as well as 0.5% Nonidet P-40. They were solubilized by using a Dounce homogenizer

with 15 strokes each of a type A and B pestle. The resulting soluble material was loaded on an 11-ml 10 to 30% glycerol gradient in 20 mM HEPES (pH 7.5)–20 mM KCl–1 mM $MgCl_2$. All solutions were made in 0.1% diethylpyrocarbonate-treated H_2O. Control gradient nuclei were suspended in the same buffer except for the addition of 0.5% sodium dodecyl sulfate and were exposed to proteinase K at a concentration of 50 μg/ml for 15 min at 4°C. The gradients were centrifuged in a Beckman SW41 rotor for 3 h at 38,000 rpm. Equal gradient fractions (0.3 ml) were collected, and fractions of three were pooled for analysis of RNA. Each fraction was diluted twofold with 2× PK buffer, digested with proteinase K (30 μg/ml) for 30 min at 50°C, extracted with phenol-chloroform, and ethanol precipitated. RNA was analyzed in polyacrylamide gels and by the primer extension technique as described above. Size markers were run in parallel gradients and included HeLa cell 40S and 60S ribosomes as present in an S100 extract, and the proteins horse spleen apoferritin (17.6S), bovine heart cytochrome *c* (12.5S), and chicken egg white conalbumin (5.0S) purchased from Sigma Chemical Co., St. Louis, Mo.

RESULTS

Fractionation of SL RNA in RNP complexes at 60S and SLIS-containing RNAs in complexes at 40S after density gradient centrifugation of nuclear extracts. Having defined two presumed components of RNA *trans* splicing, the SL transcript and the SLIS-containing RNAs, we investigated the formation of splicing complexes containing these structures by sedimentation of nuclear extracts in glycerol gradients and analysis of the RNA purified from gradient fractions of specific sedimentation value. The results in Fig. 2 show the analysis of nuclear extract RNA from different glycerol gradient fractions by ethidium bromide staining (Fig. 2A) and specific primer extension for the SL and SLIS-containing RNAs (Fig. 2B). The ethidium bromide-stained pattern of the RNA from different gradient fractions resolved in a polyacrylamide gel demonstrated that the majority of RNA that entered the gradient was concentrated in the 40S region (Fig. 2A). This area corresponds to the expected sedimentation value of heterogeneous RNA in mammalian and yeast cell extracts (12, 13). Parallel gradient analysis of nuclear RNA alone and of detergent (sodium dodecyl sulfate)- and proteinase K-treated samples showed that very little RNA entered the gradient under these conditions (Fig. 2), as RNA was concentrated in fractions 1 to 3 at the top of the gradient. Primer extension analysis of RNA from these parallel gradients showed that the SL, SLIS, and U2 equivalent RNAs were also at the top of the gradient under those conditions (data not shown). The sedimentation value of these RNAs in the gradient appeared to be protein dependent and was disrupted by ionic detergent and protease treatment in a manner similar to that of splicing complexes (12).

The results in Fig. 2B show the analysis of different gradient fractions for the SL transcript and SLIS by primer extension. The SL RNA appeared to be in a RNA-protein complex that sedimented at 60S. The SLIS-containing RNAs fractionated in the 40S region. We believe that we can rule out an effect of nonspecific binding of the SLIS to the large amount of RNA and protein in this region as the SLIS structures were most concentrated in a fraction without the most nucleic acid. In addition, the signal from the SLIS-containing RNAs was much stronger than the signal for the SL transcript at 40S despite the fact that the SL RNA could be detected in greater abundance in fractions which did not

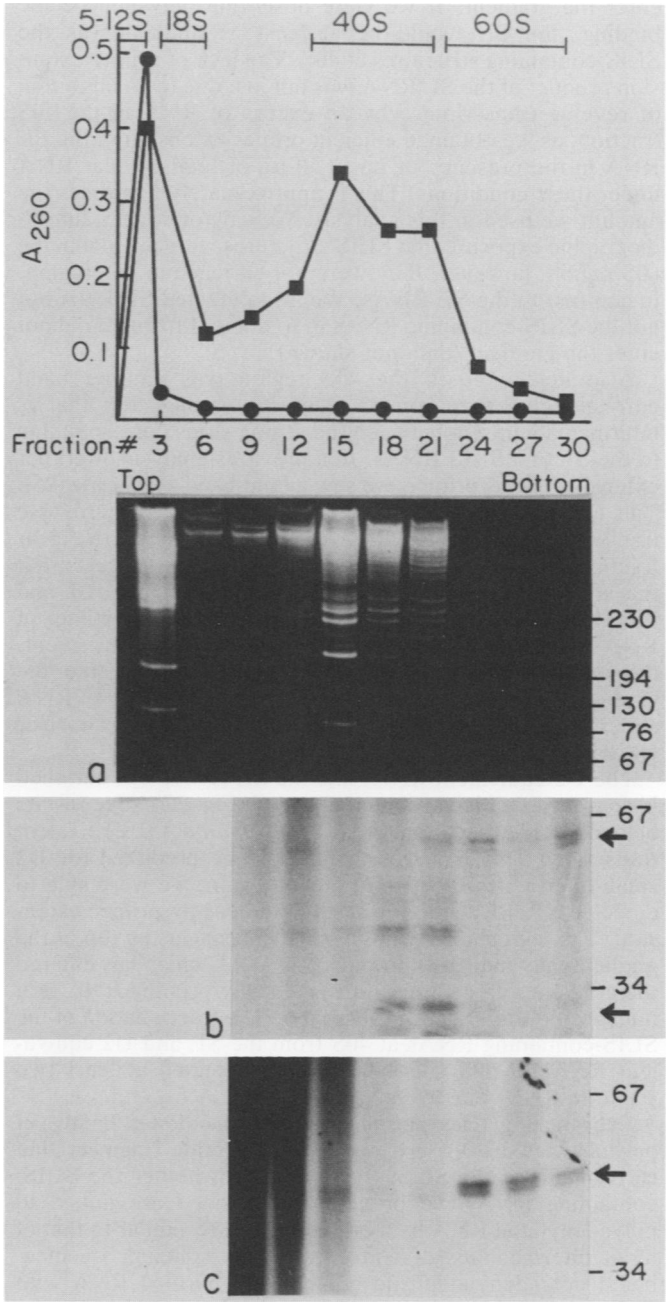

FIG. 2. Density centrifugation analysis of RNAs from *Leishmania* nuclear extract. Shown at the top is the RNA concentration in various gradient fractions as measured by A_{260}. Optical density units are shown on the ordinate, and the fractions are shown on the abscissa. Symbols: ●, profile of RNA concentration in the presence of ionic detergent and proteinase K; ■, analyzed gradient. At the top of the graph are shown the sedimentation values of gradient fractions as determined by parallel gradients with known sedimentation value. (a) Ethidium bromide-stained pattern of fractions analyzed in a 10% polyacrylamide–urea gel. Numbers on the right are in nucleotides. (b) Analysis of gradient fractions by primer extension with an SL transcript-specific oligonucleotide primer. The primer is complementary to a sequence within the SLIS. The arrow at 26 nucleotides indicates the product that corresponds to the SLIS-containing structures, and the arrow at 64 nucleotides indicates the complete SL transcript. (c) Analysis of fractions by primer extension with a U2 equivalent specific primer. The arrow at 60 nucleotides indicates the expected size corresponding to the 5′ end of U2 RNA.

enter the gradient. If we were observing only nonspecific binding, the SL would be at least as abundant as the SLIS-containing structures at 40S. The lack of primer extension product of the SL RNA was not just due to an inhibition of reverse transcriptase by an excess of RNA in the 40S fraction, as we obtained efficient primer extension of the SL RNA in the presence of up to 50 μg of total cellular RNA under these conditions. This is approximately five times the amount we used in this analysis. We saw on longer autoradiographic exposure that SLIS structures were present in the 60S region; however, they were not enriched in this region, in contrast to the SL RNAs. We also detected SL RNA but not the SLIS-containing RNAs in fractions 1 to 3 that did not enter the gradient (data not shown).

Also detectable in the 40S region was another band corresponding to a primer extension strong stop that is intermediate in size between the strong stops corresponding to the SL and SLIS RNAs. In a previous analysis of primer extension of this primer, we saw a faint band at this site (25). This may just represent incomplete transcription by reverse transcriptase off the SL RNA template; however, the clear fractionation of this "species" in the gradient suggests that this is a separate molecule containing a portion of the SL and the SLIS RNAs. Though this may represent an artifact of RNA preparation, it is a reproducible result. In any event, the majority of the SL and U2 RNAs that entered the gradient fractionated at 60S apart from the majority of RNA and the SLIS-containing RNA intermediates in nuclear extract.

The U2 equivalent RNA, like the SL RNA, was enriched in complexes which sedimented at 60S (Fig. 2C). (See results below for identification of the *Leishmania* U2 equivalent transcript.) This is larger than would be predicted for U2 small nuclear RNP (snRNP) alone. Again, we were able to detect U2 RNA at the top of the gradient by primer extension. The high background in some lanes near the top of the gradient may indicate chromosomal DNA which has entered the gradient and is either nonspecifically priming RNA or a nonspecific substrate for the primer. The fractionation of the SLIS-containing RNAs at 40S from the SL and U2 equivalent RNAs at 60S suggests that we detected at least two separate RNP complexes.

Debranching (cleavage of 2′-5′ phosphodiester bonds) of polyadenylated RNA releases a 50-nucleotide fragment that corresponds to the SLIS. To determine whether the SLIS-containing RNAs we previously described are linked to polyadenylated RNA by a branch structure similar to that of lariat intermediates of conventional *cis* splicing, we incubated polyadenylated and guanidinium-purified RNA with HeLa cell S100 extract and examined the RNA for the release of the free 50-nucleotide SLIS. The results shown in Fig. 3 document that the treatment of either guanidinium-extracted RNA (Fig. 3B) or polyadenylated phenol-chloroform-extracted RNA (Fig. 3a, lanes A+) with HeLa cell S100 extract released a 50-nucleotide RNA fragment which hybridizes to the SL gene. HeLa cell S100 extract contains enzymatic debranching activity which cleaves 2′-5′ phosphodiester bonds of lariat mRNA splicing intermediates (29). This hybridization is consistent with the release of free SLIS by the debranching enzymatic activity. The most likely interpretation of these results is that these species are intermediates in *trans* splicing that are forked or Y structures generated after cleavage of the SL from the SL transcript and attack of the 5′ guanine of the SLIS on the protein-coding substrate of *trans* splicing.

We detected in total RNA prepared by phenol-chloroform

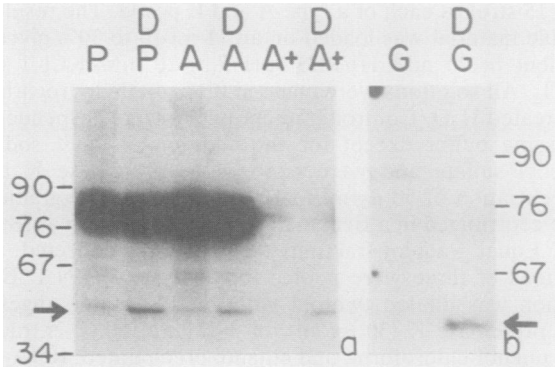

FIG. 3. Exposure of *Leishmania* RNA to debranching enzymatic activity releases the free SLIS. The figure shows the autoradiograms resulting from treatment of *L. enriettii* RNA with HeLa cell S100 extract containing debranching enzymatic activity, size fractionization in 15% polyacrylamide denaturing gels, transfer of the RNA to a hybridization membrane, and probing with the radiolabeled SL gene. (a) Phenol-purified RNA samples. Lanes: P, 15 μg of promastigote RNA; A, 15 μg of amastigote RNA; A+, 1 μg of promastigote polyadenylated RNA. The D designation indicates identical samples that have been treated with HeLa cell S100 extract. (b) Guanidinium-purified RNA. Lanes: G, 15 μg of guanidinium-purified RNA; D, identical sample treated with HeLa cell S100 extract. The size markers refer to nucleotides and are derived from lanes containing pBR322 *Msp*I and φX *Hae*III DNA digests. The arrows indicate the hybridization to the 50-nucleotide SLIS sequence. The dense hybridization band at 85 nucleotides is the hybridization of the full-length SL transcript.

extraction as well as in nonpolyadenylated RNA (data not shown) a free 50-nucleotide RNA which corresponded to the SLIS by hybridization (Fig. 3a). However, we cannot be sure that this product occurs in vivo since a recent kinetic study of the SL transcript in *T. brucei* has documented increased nuclease sensitivity of that site (21). The data in Fig. 3a demonstrate that exposure of total RNA prepared by phenol extraction to HeLa cell S100 extract rich in debranching enzymatic activity increased the intensity of hybridization at 50 nucleotides and demonstrate that RNAs from both developmental stages of *L. enriettii* contained branched structures containing the SLIS as well as the free SLIS RNA.

Homology of *Leishmania* U2 equivalent to the 5′ end of the *T. brucei* U2 equivalent and its potential to form Watson-Crick base pairs with the SL. Numerous studies of conventional *cis* splicing have documented the importance of the U2 snRNP for proper processing of pre-mRNA in vitro and in vivo (13). Mammalian splicing complexes contain the U2 RNA. One goal of this work was to identify other molecules that might be involved in *trans* splicing and could be present in RNP complexes containing mRNA intermediates. Therefore, we wished to determine whether the *Leishmania* U2 equivalent RNA was present in large RNP complexes and whether, as in *T. brucei*, there is potential for the formation of Watson-Crick base pairs between the SL and U2 RNAs. To detect the *L. enriettii* U2 equivalent RNA, we used heterologous hybridization with the cloned *T. brucei* U2 equivalent (kindly provided by C. Tschudi, Yale University) and extension of an oligonucleotide primer complementary to a conserved sequence between *Leishmania mexicana* and *T. brucei* with the enzyme reverse transcriptase. This technique was communicated to us by C. Tschudi as successful in identifying the U2 equivalent RNA of *L. mexicana*. Using an in vitro-synthesized RNA transcript complementary to

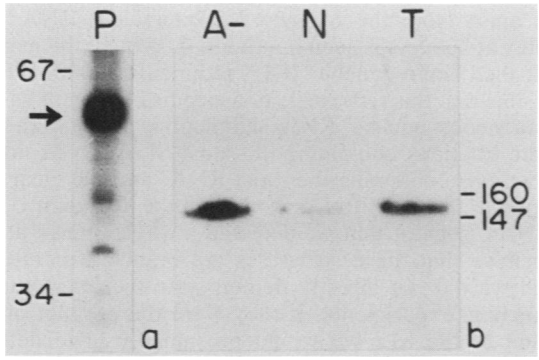

FIG. 4. Identification of the *L. enriettii* U2 equivalent. (a) Identification of a primer extension cDNA by using a primer complementary to nucleotides 44 to 60 of the *T. brucei* RNA with homology to mammalian U2. The primer extension product terminates at 60 nucleotides. The P designation indicates promastigote RNA. The arrow indicates the primer extension band. (b) Northern (RNA) blotting analysis of *Leishmania* RNA probed with a radiolabeled probe complementary to the *T. brucei* U2 RNA. Lanes: A−, 10 μg of nonpolyadenylated RNA; N, 10 μg of nuclear RNA; T, 10 μg of total RNA. All RNAs were prepared from promastigotes by phenol extraction. The indicated size markers refer to nucleotides and are from pBR322 *Msp*I digests.

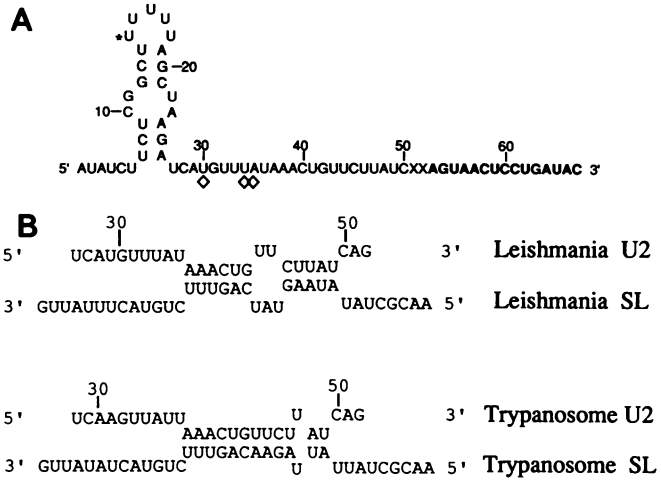

FIG. 5. (A) Sequence of the *L. enriettii* U2 equivalent RNA. The figure shows the sequence of the 67 nucleotides at the 5′ end of the U2 equivalent RNA as obtained by primer extension sequencing of a cDNA with dideoxynucleotides. The primer used was complementary to nucleotides 53 to 67 of the *T. brucei* U2 equivalent. The asterisk indicates the only difference between our results and those obtained by C. Tschudi for *L. mexicana* (personal communication). The diamonds indicates the only difference between the *L. enriettii* sequence and that of *T. brucei*. An X indicates nucleotides too near the primer to be determined. (B) Complementarity of the SL and U2 equivalent RNAs in kinetoplastida. The 35 nucleotides of the two SL sequences (the 5′ ends of the SL RNAs) are shown. The SL splice site is immediately 3′ to the sequence shown. The sequence of the U2 RNAs 30 to 50 nucleotides from the 5′ end is shown. The complementarity and the possible formation of Watson-Crick base pairs are indicated.

the *T. brucei* U2 equivalent as a hybridization probe, we detected hybridization to an approximately 150-nucleotide *Leishmania* transcript (Fig. 4B). The results also document that as expected, the U2 equivalent RNA is in nonpolyadenylated RNA. In preparations of RNA from nuclei, the U2 transcript did not appear to be enriched; this result was unexpected and may be related to the method of nucleus preparation. Perhaps the use of Nonidet P-40 in the preparation of nuclei allowed leakage of these RNPs when they were not in large complexes.

U2 RNAs are most conserved at their 5′ end. Therefore, to define the 5′ end of the U2 transcript and to obtain the direct sequence of cDNA made from the transcript, we attempted to primer extend nonpolyadenylated RNA with a primer that is complementary to nucleotides 45 to 60 of the *T. brucei* U2 equivalent. Assuming that the region of sequence complementary to the primer is the same, we should see a 60-nucleotide primer extension product. The results in Fig. 4A document that this is indeed the case. We also sequenced a cDNA primer extension product from *L. enriettii* utilizing a primer complementary to nucleotides 53 to 67 of *T. brucei* (Fig. 5). These results are identical to that obtained by C. Tschudi (personal communication) for the *L. mexicana* U2 equivalent obtained by primer extension sequencing except for the change of one uridine to an adenine at position 15 in the transcript (Fig. 5A). These data demonstrate that the *L. enriettii* U2 equivalent RNA could form Watson-Crick base pairs with the SL RNA, though fewer base pairs are theoretically possible than between the same RNAs of *T. brucei* (Fig. 5B). Additionally the changes in the SL sequence between *T. brucei* and *L. enriettii* do not involve compensatory base changes, i.e., changes in the SL sequence that could be base pairs do not change in the U2 RNA of *L. enriettii*. In *L. enriettii*, 11 base pairs are possible, 6 in one stretch, as opposed to *T. brucei*, in which 12 base pairs could form and 10 could occur in one stretch (33).

DISCUSSION

We describe here RNP complexes that contain intermediates in kinetoplastida *trans* splicing. Complexes sedimenting at 40S contain the SLIS RNAs, and others sedimenting at 60S contain the SL and U2 equivalent RNAs. We investigated RNP complexes to define physical characteristics of RNA splicing intermediates we previously identified. We reasoned that we would detect RNP complexes despite the novel nature of discontinuous transcription, because many elements common to conventional eucaryotic mRNA processing, such as intron-exon splice junctions and an RNA homologous to mammalian U2, are present in kinetoplastida (3, 5, 32).

We also demonstrated that exposure of *Leishmania* polyadenylated RNA to a HeLa cell S100 extract containing debranching enzymatic activity released a 50-nucleotide RNA species which corresponded to the SLIS. Our previous work, using primer extension and S1 nuclease mapping techniques, delineated RNA species containing the SLIS with a 5′ terminus at the consensus 5′ exon-intron splice junction within the SL transcript (25). The result we demonstrated here is most consistent with the existence of numerous intermediates of mRNA processing in which the SLISs are attached to high-molecular-weight RNA species through a covalent branch structure. Our results characterizing the branched linkage of the SLIS to polyadenylated RNA are analagous to those recently obtained for *T. brucei* and provide further evidence of *trans* splicing in kinetoplastida (21, 26, 31).

Kinetoplastida splicing shares many similarities with nuclear pre-mRNA *cis* splicing (12) (Fig. 1A). *cis* and *trans* splicing are analogous in that they have similar splice junctions and form branch structures. The major defined differences in kinetoplastida transcription are a common 5′ exon,

two pre-mRNA substrates, and a forked rather than a lariat intermediate. We suggest here that *L. enriettii* also generates RNP complexes with a physical sedimentation similar to that of mammalian splicing complexes. These complexes differ from those of well-studied systems in that branch structures are present in a complex at 40S. In conventional *cis* splicing, lariat intermediates are found in the mature spliceosome at 60S (12, 13, 17).

The precursor SL RNAs were enriched in the RNP complexes sedimenting at 60S, and the partially processed SLIS-containing RNAs were enriched in 40S complexes. This is in contrast to the situation in *cis* splicing in which the pre-mRNA is found in an RNP complex at 40S, while the lariat intermediates are found in a complex at 60S (12, 13, 17). Recently, several groups have shown evidence of large transcription units of the protein-coding component of mature mRNA in kinetoplastida (11, 14, 15). Two of these groups demonstrated large transcription units of tandemly repeated genes (11, 14). In *L. enriettii*, both the alpha- and beta-tubulin genes exist as tandem repeats within the genome (22). The tubulin and calmodulin genes are organized as tandem repeats within the genome of African trypanosomes (30, 33). It is possible that the protein-coding RNA substrate in *Leishmania trans* splicing is a multimeric transcription unit of tandem repeats. One model to explain these results is that the 60S complexes contain large polycistronic RNAs and multiple U2 and SL RNAs with their associated proteins. Once RNA processing has been initiated by the formation of the branched intermediate and polyadenylation, the partially processed mRNA is cleaved from the polycistronic transcript and the complex is reduced to 40S. Future studies of these complexes may demonstrate that they contain multimeric protein-coding pre-mRNAs, that they are truly related to RNA processing, and that they contain more RNA and protein components analagous to those present in yeast and mammalian cell splicing complexes.

The 60S RNP complexes we demonstrated here contained an *L. enriettii* RNA that is homologous to the U2 equivalent RNA of *T. brucei*. As in *T. brucei*, the U2 equivalent RNA is 150 nucleotides long and is likely to have a functional role in RNA processing of mRNAs given the requirement for snRNPs in other eucaryotic splicing systems. The *L. enriettii* U2 RNA shares 56 of 60 nucleotides at its 5′ end with the *T. brucei* U2 RNA. The *L. mexicana* and *L. enriettii* U2 equivalent RNAs are identical over the same region of sequence except for one nucleotide (C. Tschudi, personal communication). There is potential for the formation of Watson-Crick base pairs between the SL and the U2 equivalent RNAs in *L. enriettii*, as in *T. brucei* (Fig. 5B) (32).

Our demonstration of the SL and U2 equivalent RNA complementarity deserves comment. While it is not as impressive as the degree of complementarity observed in *T. brucei* (32), it still may be significant. In other systems, U1 snRNPs form base pairs with only two nucleotides at the 5′ exon-intron junction (13). The U2 equivalent in yeasts forms Watson-Crick base pairs at the consensus intron branch sequence with seven nucleotides. It seems possible that the sequence complementarity we observed has significance to mRNA processing despite the lack of compensatory nucleotide change across species of kinetoplastida.

In this work, we used density centrifugation and analysis of RNA to demonstrate that the SL and the U2 equivalent RNAs cosediment at 60S in an *L. enriettii* nuclear extract. These RNAs could be associated with protein in separate or the same RNP complexes. We also demonstrated that in the *L. enriettii* nuclear extract, the SLIS-containing RNAs sediment apart from the SL and U2 equivalent RNAs in a complex at 40S. As is seen in other eucaryotes, the majority of so-called heterogeneous RNA sediments at 40S. While it is possible that this represents nonspecific binding of RNA to heterogeneous nuclear RNA sedimenting at 40S, the fact that the fractions containing the most SLIS RNAs are not the fractions containing the most RNA, and the clear fractionation of the SLIS RNAs from both the abundant U2 and SL RNAs, suggest that we detected a true complex at 40S. We believe that these structures are branch structures although we did not directly demonstrate their structure by debranching experiments. If they were the product of degradation of the SL, we would presumably have detected them at 60S with the SL. In our previous work, we demonstrated that the primer extension signal we obtained is directly related to the RNA that sediments in cesium chloride gradients and that does not contain small RNAs, i.e., 50-nucleotide SLIS RNAs that could be formed by RNA degradation or processing at a nuclease-sensitive site.

A remarkable finding is that the majority of the U2 snRNPs sediment at 60S apart from the majority of the SLIS-containing intermediates at 40S. In the present models of conventional eucaryotic splicing, U2 is required for successful intron removal and interacts at the site of intron branch formation. In the yeast *Saccharomyces cerevisiae*, a U2 equivalent RNA forms Watson-Crick base pairs with the single consensus branch sequence as a requirement for splicing (27). In mammalian cells, U2 interacts (if it does not form base pairs) with the variable intron branch sequences (13). Given the possibility for formation of Watson-Crick base pairs between the SL and U2, it suggests that the SL and U2 are interacting and might be in physical association.

Although the RNP complexes we described here were not proved to be involved in RNA splicing of discontinuous transcription precursors, we believe we demonstrated a clear separation of the SL and U2 from SLIS-containing RNAs and that complexes of 40 and 60S contain protein- and RNA-processing intermediates. Understanding the kinetics of *trans* splicing and the significance of our results will require the development of an in vitro RNA-processing system, but we made a beginning in the understanding of this process with the detection of native complexes containing the U2 and SL RNAs.

The possible physical association of U2 and SL RNAs in an RNP complex leads us to speculate on the role of a U2-like snRNP in RNA *trans* splicing. If the SL and U2 RNAs are physically associated through the formation of Watson-Crick base pairs in a splicing complex, this might indicate that the U2 snRNP has a unique or dual function in this system. It could be that in conventional *cis* splicing the U2 snRNP physically interacts with the 5′ exon but that in *Leishmania trans* splicing, because the 5′ exon is always the same, this sequence complementarity evolved to maximize efficiency. Perhaps in *trans* splicing, U2 snRNP does not interact with the branch site as it does in *cis* splicing. The recent delineation of *trans* splicing with a common 5′ spliced leader sequence in some but not all of the nematode *C. elegans* genes (20) should allow another comparison of U2 and SL RNA complementarity in a more evolutionarily divergent multicellular organism.

ACKNOWLEDGMENTS

We thank Michael Green for seminal discussions about this work as well as his encouragement. We specifically thank Claudio Pickelney as well as all the members of Michael Green's laboratory for

helpful discussions and S100 extracts. We thank Chris Tschudi for the personal communication of results prior to publication.

This work was supported by Public Health Service grant AI 21365 from the National Institutes of Health and a grant from the Catherine and John D. MacArthur Foundation. S.I.M. was supported by NRSA grant AI 07179. D.F.W. is a Burroughs Wellcome scholar in molecular parasitology.

LITERATURE CITED

1. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
2. **Black, D. L., B. Chabot, and J. A. Steitz.** 1985. U2 as well as U1 small nuclear riboproteins are involved in pre-messenger RNA splicing. Cell **42**:737–750.
3. **Boothroyd, J. C., D. A. Campbell, and R. E. Sutton.** 1985. Expression of surface antigen genes in Trypanosoma brucei involves a novel system of discontinuous transcription, p. 61–66. *In* R. A. Lerner, R. M. Channock, and F. Brown (ed.), Modern approaches to vaccines. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
4. **Boothroyd, J. C., and G. A. M. Cross.** 1982. Transcripts coding for different variant surface glycoproteins in *Trypanosoma brucei* have a short identical exon at their 5'-end. Gene **20**:279–287.
5. **Borst, P.** 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annu. Rev. Biochem. **55**:701–732.
6. **Brody, E., and J. Abelson.** 1985. "The spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. Science **228**:963–967.
7. **Campbell, D. A., D. A. Thornton, and J. C. Boothroyd.** 1984. Apparent discontinuous transcription of *Trypanosoma brucei* surface antigen genes. Nature (London) **311**:350–355.
8. **Comeau, A. M., S. I. Miller, and D. F. Wirth.** 1986. Chromosome location of four genes in *Leishmania*. Mol. Biochem. Parasitol. **21**:161–169.
9. **De Lange, T., P. A. M. Michels, H. J. G. Veerman, A. W. C. A. Cornelissen, and P. Borst.** 1984. Many trypanosome messenger RNAs share a common 5' terminal sequence. Nucleic Acids Res. **12**:3777–3789.
10. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
11. **Gonzalez, A., T. J. Lerner, M. Huecas, B. Sosa-Pineda, N. Nogueria, and P. M. Lizardi.** 1985. Apparent generation of a segmented mRNA from two separate tandem gene families in *Trypanosoma cruzi*. Nucleic Acids Res. **13**:5789–5804.
12. **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.
13. **Green, M. R.** 1986. Pre-mRNA splicing. Annu. Rev. Genet. **20**:671–708.
14. **Imboden, M. A., P. W. Laird, M. Affolter, and T. Seebeck.** 1987. Transcription of the intergenic regions of the tubulin gene cluster of Trypanosoma brucei: evidence for a polycistronic transcription unit in a eukaryote. Nucleic Acids Res. **15**: 7357–7368.
15. **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.
16. **Kaplan, G. M., K. Zhang, and S. M. Beverly.** 1987. Sequence and S1 nuclease mapping of the 5' region of the dihydrofolate reductase thymidylate synthetase gene of *Leishmania major*. Nucleic Acids Res. **15**:3369–3383.
17. **Konsarka, M. M., and P. A. Sharp.** 1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell **49**:763–774.
18. **Kooter, J. M., T. De Lange, and P. Borst.** 1984. Discontinuous synthesis of mRNA in trypanosomes. EMBO J. **3**:2387–2392.
19. **Krainer, A. R., and T. Maniatis.** 1985. Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for pre-mRNA splicing in vitro. Cell **42**:725–736.
20. **Krause, M., and D. Hirsh.** 1987. A trans-spliced leader sequence on actin mRNA in C. elegans. Cell **49**:753–761.
21. **Laird, P. W., C. B. M. Zomerdijk, D. Korte, and P. Borst.** 1987. In vivo labeling of intermediates in the discontinuous synthesis of mRNAs in *Trypanosoma brucei*. EMBO J. **6**:1055–1062.
22. **Landfear, S. M., S. I. Miller, and D. F. Wirth.** 1986. Transcriptional mapping of *Leishmania enrietti* tubulin mRNAs. Mol. Biochem. Parasitol. **21**:235–245.
23. **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.
24. **Milhausen, M., R. G. Nelson, S. Sather, M. Selkirk, and N. Agabian.** 1984. Identification of a small RNA containing the trypanosome spliced leader: a donor of shared 5' sequences of trypanosomatid mRNAs? Cell **38**:721–729.
25. **Miller, S. I., S. M. Landfear, and D. F. Wirth.** 1986. Cloning and characterization of a *Leishmania enriettii* gene encoding a RNA spliced leader sequence. Nucleic Acids Res. **14**:7341–7360.
26. **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.
27. **Parker, R., P. G. Siliciano, and C. Guthrie.** 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2 like snRNA. Cell **49**:229–239.
28. **Parsons, M., R. G. Nelson, K. P. Watkins, and N. Agabian.** 1984. Trypanosome mRNAs share a common 5' spliced leader sequence. Cell **38**:309–316.
29. **Ruskin, B., and M. R. Green.** 1985. An RNA processing activity that debranches RNA lariats. Science **229**:135–140.
30. **Sather, S., and N. Agabian.** 1985. A 5' spliced leader is added in trans to both alpha and beta tubulin transcripts in *Trypanosoma brucei*. Proc. Natl. Acad. Sci. USA **82**:5695–5699.
31. **Sutton, R. E., and J. C. Boothroyd.** 1986. Evidence for trans splicing in trypanosomes. Cell **47**:525–535.
32. **Tschudi, C., F. F. Richards, and E. Ullu.** 1986. The U2 RNA analogue of *Trypanosoma brucei* gambiense: implications for a splicing mechanism in trypanosomes. Nucleic Acids Res. **14**:8893–8903.
33. **Tschudi, C., A. S. Young, L. Ruben, C. L. Patton, and F. F. Richards.** 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.