MOLECULAR AND CELLULAR BIOLOGY, Nov. 1991, p. 5516–5526
0270-7306/91/115516-11$02.00/0

Vol. 11, No. 11

# Analysis of Small Nuclear Ribonucleoproteins (RNPs) in *Trypanosoma brucei*: Structural Organization and Protein Components of the Spliced Leader RNP

MIKE CROSS, ARTHUR GÜNZL, ZSOFIA PALFI, AND ALBRECHT BINDEREIF*

*Max-Planck-Institut für Molekulare Genetik, Otto-Warburg-Laboratorium, Ihnestrasse 73, D-1000 Berlin 33 (Dahlem), Germany*

Received 3 June 1991/Accepted 30 July 1991

***trans* splicing in *Trypanosoma brucei* involves the ligation of the 40-nucleotide spliced leader (SL) to each of the exons of large, polycistronic pre-mRNAs and requires the function of small nuclear ribonucleoproteins (snRNPs). We have identified and characterized snRNP complexes of SL, U2, U4, and U6 RNAs in *T. brucei* extracts by a combination of glycerol gradient sedimentation, CsCl density centrifugation, and anti-$m_3G$ immunoprecipitation. Both the SL RNP and the U4/U6 snRNP contain salt-stable cores; the U2 snRNP, in contrast to other eucaryotic snRNPs, is not stable under stringent ionic conditions. Two distinct complexes of U6 RNA were found, a U6 snRNP and a U4/U6 snRNP. The structure of the SL RNP was analyzed in detail by oligonucleotide-directed RNase H protection and by in vitro reconstitution. Our results indicate that the 3′ half of SL RNA constitutes the core protein-binding domain and that protein components of the SL RNP also bind to the U2 and U4 RNAs. Using antisense RNA affinity chromatography, we identified a set of low-molecular-mass proteins (14.8, 14, 12.5, and 10 kDa) as components of the core SL RNP.**

In trypanosomes, all nuclear mRNAs analyzed to date are generated through a *trans*-splicing reaction whereby a noncoding 40-nucleotide spliced leader (SL) is joined to each protein-coding exon. Both of these RNAs are independently transcribed as precursors, an SL RNA of approximately 140 nucleotides with the SL portion at the 5′ end and pre-mRNAs derived from long, polycistronic transcription units. *trans* splicing proceeds through a two-step mechanism of cleavage-ligation reactions formally analogous to conventional *cis* splicing of other eucaryotic systems: the first step results in a branched RNA intermediate corresponding to the lariat intermediate RNA of the *cis*-splicing reaction; in the second step, the SL exon is ligated to the protein-coding exon and the intron portion is released in the form of a Y-shaped RNA molecule (38, 47; for reviews, see references 1 and 7).

Through the development of in vitro processing systems derived from *Saccharomyces cerevisiae* and mammalian cells, we have gained a wealth of detailed knowledge about the *cis*-splicing machinery, its components, and the biochemical mechanism. The coordinate interactions of small nuclear ribonucleoproteins (snRNPs) with themselves, the splicing substrate, and other splicing factors are essential for the assembly of the pre-mRNA into spliceosomes and for efficient and accurate *cis* splicing (for recent reviews, see references 5, 26, 29, and 46). In contrast, we know very little about the biochemical components and mechanism of *trans* splicing in trypanosomes, partly because *trans* splicing has yet to be reproduced in a trypanosomal in vitro system. On the basis of their structural homology with higher eucaryotic snRNAs, the *Trypanosoma brucei* U2, U4, and U6 RNAs have been identified and their respective genes have been cloned (36, 50, 51). Experimental evidence was recently obtained from oligonucleotide-directed RNase H cleavage studies with permeabilized trypanosomes that, in analogy to *cis* splicing, U2, U4, and U6 RNAs are necessary *trans*-splicing factors (52). In addition, RNA B, an RNA without a strong homology to any known eucaryotic snRNA, was identified as an $m_3G$-capped RNA of *T. brucei* (36). We do not know at present whether RNA B plays a role in *trans* splicing. As in all eucaryotic systems analyzed to date (18), U4 and U6 RNAs could be coimmunoprecipitated from total RNA by anti-$m_3G$ antibodies, suggesting that also in trypanosomes these RNAs interact via base pairing within a single RNP (36). In contrast, no U1 and U5 RNA homologs are known in trypanosomes; the SL RNA, however, has been proposed to function both as a U1-like snRNP and as a donor of its own 5′-terminal 40 nucleotides (11, 54). This has recently been directly demonstrated in a mammalian in vitro system, in which the 5′-terminal domain of *Leptomonas collosoma* SL RNA could substitute for the 5′ end of U1 snRNA in pre-mRNA *cis* splicing (10). In support of such a functional role, the SL sequence as well as the overall secondary structure of the SL RNA are strongly conserved between different trypanosome and nematode species (11, 22, 32, 40; for the *T. brucei* SL RNA structure, see Fig. 3A).

The trypanosomal snRNAs deviate most significantly from the otherwise strongly conserved sequence and secondary structure of eucaryotic snRNAs (18). For example, none of the known *T. brucei* snRNAs carry an Sm binding site, a sequence motif with the consensus $PuAU_{3-4}NUGPu$, which is present in U1, U2, U4, and U5 RNAs of all other eucaryotic species and usually located in a single-stranded region between two stem-loops (18, 28). In the mammalian system, the Sm binding site is essential for the trimethyl capping of snRNAs and for the cytoplasmic-nuclear transport of snRNPs (reviewed in reference 31). Furthermore, human anti-Sm antibodies do not cross-react with *T. brucei* snRNPs (33). There are other deviations of the *T. brucei* snRNAs from the eucaryotic consensus sequences and secondary structures, such as the lack of stem-loop III and of a branch point interaction region in *T. brucei* U2 RNA (20, 36, 53). Finally, the cap structure of the SL RNA is exceptional

* Corresponding author.

in that it carries an unusual, so-called cap 4 structure with a 7-methylguanosine cap and additional, internal ribose modifications (14, 15, 42, 48).

The requirement of snRNPs in *trans* splicing and the conservation of certain sequence and structural elements in the trypanosomal snRNAs and in the *trans*-splicing substrate suggest that, in analogy to *cis* splicing, snRNPs play essential roles during the assembly of *trans* spliceosomes. In *trans* splicing, the U2 and U4/U6 snRNPs may also be involved in mediating the interactions between the two precursor RNAs. It is the underlying rationale of our study that structural deviations of trypanosomal snRNPs from the otherwise highly conserved eucaryotic snRNPs reflect important functional differences between the *cis*- and *trans*-splicing mechanisms.

In this report, we describe the identification and initial characterization of snRNP complexes containing SL, U2, U4, U6, and RNA B from *T. brucei* extracts, with particular focus on the domain structure of the SL RNP, which we have analyzed by RNase H protection and protein binding assays. The core domain was mapped to the 3′ half of the SL RNA; significantly, SL RNP reconstitution experiments provided evidence for common proteins binding to the SL, U2, and U4 RNAs. Through affinity chromatography with biotinylated, antisense 2′-*O*-methyl RNA (2′-OMe) oligonucleotides, we succeeded in purifying the core SL RNP and identifying at least some of its protein components.

## MATERIALS AND METHODS

**Templates for T7 transcription.** *T. brucei* SL, U2, U4, and U6 (34, 36, 51) and *Crithidia fasciculata* SL (37) sequences coupled to an upstream T7 promoter were amplified from genomic DNA by polymerase chain reaction methods and cloned into pEMBL8 (*T. brucei* SL and U4), pUC19 (U2 and U6), or pEMBL19 (*C. fasciculata* SL) to give recombinant plasmids pT7-TbSL, pT7-TbU2, pT7-TbU4, pT7-TbU6, and pT7-CfSL, respectively. The sequence of each construct was confirmed by dideoxynucleotide sequencing with Sequenase (U.S. Biochemical). In vitro transcription was done as described previously (4, 35). The sequences of the synthetic snRNAs deduced from the DNA sequence are (the extra nucleotides of the synthetic snRNAs are separated from the natural snRNA sequence by the slash): T7-TbSL, GGC/AACU–SL-coding sequences–ACUC/CCGG (*Ava*I runoff); T7-TbU2, GGG/AUAU–U2-coding sequences–CG GU/CUAG (*Xba*I runoff); T7-TbU4, GGGAG/AAGC–U4-coding sequences–GGGA/UC (*Bam*HI runoff); T7-TbU6, G/GGAG–U6-coding sequences–UUUU/CUAG (*Xba*I runoff); and T7-CfSL, GGC/AACU–SL-coding sequence–GCUG/AAUU (*Eco*RI runoff).

**Growth of trypanosomes and preparation of cell extracts.** The procyclic form of *T. brucei brucei* 427 (12) was cultivated at 28°C in SDM-79 medium (9) supplemented with 5% (vol/vol) fetal calf serum and 5 mg of hemin per liter. For cell extract preparation, trypanosomes were harvested, washed with phosphate-buffered saline and resuspended in cold nuclei preparation buffer [20 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 7.5), 15 mM NaCl, 60 mM KCl, 0.5 mM EGTA, 4 mM EDTA, 14 mM β-mercaptoethanol, 0.15 mM spermine, 0.5 mM spermidine, 0.125 mM phenylmethylsulfonyl fluoride (PMSF)] at a density of $10^9$ cells per ml. Cells were lysed by passage through a Stansted Cell Disruptor at 1,000 lb/in$^2$, and the disrupted material was collected by centrifugation. The disrupted cell pellet was used to prepare nuclear extract essentially as described previously (13). S100 extract was prepared from the supernatant by addition of KCl to 140 mM and centrifugation at 100,000 × *g* for 60 min, the resultant supernatant being retained. All extracts were finally dialyzed against buffer D containing 100 mM KCl (13), flash-frozen, and stored at −80°C. Nuclear and S100 extracts were found to contain similar concentrations of snRNAs, as assayed by primer extension (data not shown), although the total yield of snRNAs was generally some 15- to 20-fold greater from S100 extract than from nuclear extract.

DEAE chromatography of *T. brucei* extract was used to prepare a fraction enriched in snRNPs (33). Undialyzed S100 extract (50 ml) was adjusted to a $MgCl_2$ concentration of 3 mM and applied to a 5-ml DEAE-Sepharose column (Pharmacia LKB) equilibrated in buffer $D_{100}$ (100 mM KCl, 3 mM $MgCl_2$, 20 mM N-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES; pH 8.0], 0.5 mM PMSF, 0.5 mM dithiothreitol, 5% glycerol). After washing with buffer $D_{100}$, bound material was eluted with buffer $D_{400}$ (as $D_{100}$ but with 400 mM KCl), and fractions were assayed for SL RNA by primer extension and for total RNA content by denaturing polyacrylamide gel electrophoresis and silver staining (Bio-Rad).

**Sedimentation, CsCl gradient, and immunoprecipitation analyses.** *T. brucei* S100 or nuclear extract (200 μl) was mixed with 250 μl of buffer G (20 mM KCl, 20 mM HEPES [pH 8.0], 1 mM $MgCl_2$) and fractionated through a 3.8-ml 10 to 30% glycerol gradient in buffer G (SW60, 6 h, 35,000 rpm, 4°C). RNA was prepared from each fraction and assayed for SL, U2, U4, U6, and RNA B by primer extension with DNA oligonucleotides complementary to nucleotides 110 to 121 of SL RNA, 83 to 102 of U2, 53 to 72 of U4, 40 to 59 of U6, and 48 to 64 of RNA B. Immunoprecipitation analysis with anti-$m_3G$ antibodies (R. Lührmann) was carried out as described previously (3), followed by primer extension analysis of RNA prepared from the immunoprecipitates. S100 or nuclear extract (200 μl) was subjected to CsCl density centrifugation as described previously (43). RNA was prepared from gradient fractions and assayed by primer extension.

**RNase H protection analysis of the SL RNP.** Oligonucleotides a to f were complementary to the following SL RNA sequences: a, nucleotides 1 to 15; b, nucleotides 16 to 27; c, nucleotides 31 to 45; d, nucleotides 40 to 54; e, nucleotides 66 to 85; and f, nucleotides 110 to 121 (see also Fig. 3A). The nonspecific control oligonucleotide was complementary to nucleotides 40 to 59 of *T. brucei* U6 RNA. The standard RNase H reaction mix (25 μl) contained 15 μl of extract (or deproteinized RNA prepared from the same extract), 2 U of *Escherichia coli* RNase H (Boehringer), 40 μg of oligonucleotide per ml, and 3 mM $MgCl_2$. Reaction mixes were incubated at 30°C for 60 min; RNA was prepared and analyzed by primer extension with an oligonucleotide complementary to nucleotides 125 to 144 of SL RNA.

**Reconstitution and tagging of the core SL RNP.** The standard 25-μl reconstitution reaction mix contained 1 ng of $^{32}$P-labeled synthetic *T. brucei* SL RNA, 10 mM HEPES (pH 8.0), 2 mM $MgCl_2$, and 5 μl of the DEAE fraction. Incubation was for 30 min at 37°C; during the last 5 min, heparin (1 mg/ml) was added. Complexes were resolved by native RNP gel electrophoresis as described previously (39). For the competition experiments, unlabeled competitor RNA was added at the start of the reconstitution reaction.

SL RNA or RNP present in the DEAE fraction was labeled by tagging with 1 ng of $^{32}$P-end-labeled 2′-OMe RNA oligonucleotide SL-5′, complementary to nucleotides 1 to 15

of SL RNA. Incubation was as described for the reconstitution reaction.

**Purification of the *T. brucei* SL RNP.** *T. brucei* S100 extract (50 ml; $10^{11}$ cell equivalents) was subjected to DEAE chromatography as described above. The SL RNP-containing fractions were pooled (2 ml) and loaded onto an 11-ml 10 to 30% glycerol gradient prepared with buffer $D_{400}$. After centrifugation (SW40, 24 h, 29,000 rpm, 4°C), the gradient was fractionated and assayed for total RNA content. Fractions containing the SL RNP were pooled (5 ml), adjusted to 800 mM KCl, and used for affinity purification of the SL RNP with a biotinylated 2'-OMe RNA oligonucleotide complementary to nucleotides 40 to 54 of SL RNA (5'-GGGA GCUUCUCAUACXXXXA-3', X denoting a biotinylated 2'-deoxycytidine [45]) applied at a concentration of 8 μg/ml. Incubation was for 30 min at 30°C. As controls, affinity selections were carried out either in the absence of biotinylated oligonucleotide or in the presence of an unrelated, biotinylated 2'-OMe RNA oligonucleotide (5'-UXXXXU ACUGCCACUGCGCAAAGCU-3'; complementary to the 5' end of human U4 RNA). A one-fifth volume of a 1:1 suspension of preblocked streptavidin-agarose beads (2) in washing buffer ($WB_{800}$; 20 mM Tris-HCl [pH 7.5], 0.01% Nonidet P-40, 0.5 mM PMSF, 0.5 mM dithiothreitol, 800 mM KCl) was then added, and the mixture incubated at 4°C for 60 min. The streptavidin-agarose beads were then collected by centrifugation and washed five times with $WB_{800}$. First, selected proteins were eluted by incubation with 9 M urea for 30 min at room temperature and precipitated by acetone. RNA was subsequently released by the addition of PK buffer (100 mM Tris-HCl [pH 7.5], 12.5 mM EDTA, 150 mM NaCl, 1% sodium dodecyl sulfate [SDS]) and incubation at 80°C for 10 min. Affinity-selected RNA was analyzed after phenol extraction on 8% denaturing polyacrylamide-urea gels and detected by silver staining. The efficiency of SL RNP affinity selection was determined by primer extension of SL RNA and scanning the subsequent autoradiograph with an LKB Ultroscan XL. Affinity-selected proteins were analyzed by SDS-polyacrylamide gel electrophoresis and visualized by staining with Coomassie brilliant blue G-250.

## RESULTS

***T. brucei* snRNAs are in the form of distinct snRNP complexes.** To identify and characterize RNP complexes of the trypanosomal snRNAs, we used a combination of sedimentation through glycerol gradients, isopycnic centrifugation in cesium chloride, and anti-$m_3G$ immunoprecipitation. Since snRNAs occur in *T. brucei* extracts in concentrations of only between 1 and 10% of those found in mammalian extracts (data not shown), a sensitive primer extension assay was developed for their detection. Several different oligonucleotides against the known *T. brucei* snRNAs were chosen so that in a single assay the extension products of SL, U2, U4, U6, and RNA B could be detected (Fig. 1 and 2). Figure 1 shows the fractionation of snRNA complexes from *T. brucei* nuclear extract by glycerol gradient sedimentation. SL and U2 RNAs both sediment at approximately 8S, and U4 RNA and RNA B sediment at approximately 12S. The SL RNA complexes have a size much more distinct than those of the other snRNAs, which distribute more heterogeneously. In contrast, U6 RNA complexes peak at around 7S, with some also spreading into the 12S region. Under these sedimentation conditions, free snRNAs fractionate below 5S (data not shown). Fractionation of S100 extract produced an snRNP distribution very similar to that obtained with nuclear extract (data not shown).

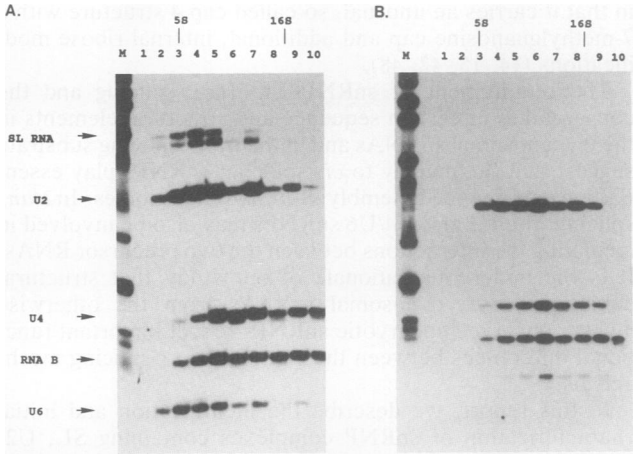


FIG. 1. Glycerol gradient sedimentation of *T. brucei* snRNPs. Nuclear extract was fractionated through a 10 to 30% glycerol gradient (fractions 1 to 10, from top to bottom), and the distribution of SL, U2, U4, U6, and RNA B was detected by primer extension. (A) Analysis of total fractions (the apparent reduction of extension products in lanes 6 and 8 was not reproducible); (B) analysis after anti-$m_3G$ immunoprecipitation. The positions of size markers (5S and 16S) are indicated at the top. Lane M, pBR322 *Hpa*II-digested marker fragments.

To determine whether there are different U6 RNA complexes, we immunoprecipitated fractions of an identical gradient with anti-$m_3G$ antibodies prior to the primer extension assay (Fig. 1B). U6 RNA, which has no $m_3G$ cap structure, is immunoprecipitated only when associated with an $m_3G$-capped RNA, most likely U4 (36). It is clear that only the 12S fraction of U6 RNA is immunoprecipitated (Fig. 1B). Immunoprecipitated U4 RNA complexes peak in the same region of the gradient, consistent with their representing the U4/U6 snRNP, although we cannot rule out the possibility that an $m_3G$-capped RNA other than U4 is associated with U6 RNA. In sum, two different U6 RNA complexes could be distinguished from each other, a U6 snRNP of approximately 7S and a 12S RNP complex, most likely the U4/U6 snRNP (see also below and Fig. 2). U2 RNA and RNA B complexes of heterogeneous size were also immunoprecipitated; as expected, complexes of SL RNA, which has a 7-methylguanosine cap, did not react with anti-$m_3G$ antibodies.

Besides sedimentation, isopycnic centrifugation in cesium chloride provides a useful approach to characterize snRNP complexes because it allows the separation of free proteins, RNA-protein complexes, and free RNA according to density. In the presence of $MgCl_2$, snRNP core complexes are characteristically stable under these stringent conditions (27). Figure 2 displays the distribution of snRNP complexes from *T. brucei* nuclear extract after fractionation in a CsCl gradient. The SL RNA distributed in fractions 2 to 4, and RNA B, U4, and U6 RNAs distributed in fractions 3 to 5, corresponding to buoyant densities of 1.43 and 1.45 g/ml, respectively. Less than 20% of RNA B and of SL and U4 RNAs, but about 50% of U6 RNA, behaved as free RNA (Fig. 2, fraction 10). The cofractionation of U4 and U6 RNAs in fractions 3 to 5 supports the notion that they interact in the form of a U4/U6 snRNP (compare with Fig. 1). Significantly and in contrast to SL, U4, U6, and RNA B, we found that

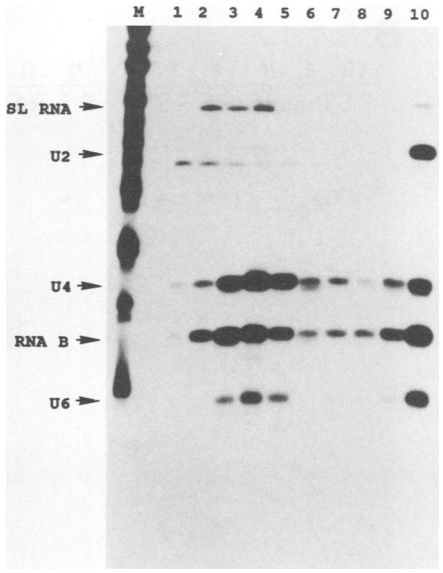


FIG. 2. Isopycnic CsCl gradient centrifugation of *T. brucei* snRNPs. Nuclear extract was subjected to CsCl density centrifugation, and RNAs from gradient fractions 1 to 10 (from top to bottom) were assayed by primer extension. Lane M, *Hpa*II-digested pBR322 marker fragments.

U2 fractionated almost completely as free RNA (Fig. 2, fraction 10), indicating that the *T. brucei* U2 snRNP is not stable under these stringent conditions. The same result was obtained when S100 extract was fractionated (data not shown). In sum, with the exception of the U2 snRNP, complexes of all RNAs tested (SL, U4/U6, and RNA B) possess salt-stable cores. We therefore conclude that not all *T. brucei* snRNPs appear to share the same core domain, in contrast to the U1, U2, U4/U6, and U5 snRNPs of all other eucaryotic species examined, which have the Sm domain as a common core structure.

**Domain structure of the SL RNP: the 3′ half of SL RNA constitutes a core protein-binding domain.** The foregoing data show that in *T. brucei* extracts, SL RNA exists in the form of an RNP particle. To investigate which regions of SL RNA are sites for protein binding, we analyzed the SL RNP by oligonucleotide-directed RNase H cleavage and protection. Six oligonucleotides were synthesized, each complementary to a specific region of the proposed SL RNA secondary structure (Fig. 3A), and tested for their ability to direct cleavage of SL RNA. S100 extract (Fig. 3B, lanes S) was incubated with each oligonucleotide and RNase H at 30°C for 60 min. In parallel, deproteinized RNA (lanes R) prepared from the same extract was used to determine whether protection against cleavage was due to either protein binding or SL RNA secondary structure. Specific cleavage of SL RNA was assayed by primer extension. The cleavage-protection patterns produced by each oligonucleotide were similar under both low- and high-salt conditions (60 and 300 mM KCl, respectively), with different S100 extracts, and also with nuclear extract (data not shown).

Incubation of extracts with a nonspecific control oligonucleotide did not result in degradation of SL RNA (Fig. 3B, lanes ctr; the primer extension product of ca. 55 nucleotides was not SL RNA dependent; data not shown). This was also the case with oligonucleotide e, which is complementary to part of the predicted large, central stem-loop of SL RNA (lanes e), consistent with the proposed secondary structure; therefore, we cannot draw any conclusions about protein binding in this region. In contrast, most of the SL RNA-specific oligonucleotides were able to induce cleavage of SL RNA in S100 extract or in RNA prepared from extract (Fig. 3B). Oligonucleotides c and d, complementary to nucleotides 31 to 54 (Fig. 3A), induced complete cleavage of SL RNA in both extract and deproteinized RNA, yielding the expected shorter cleavage products (Fig. 3B, lanes c and d). The small primer extension products of between 30 to 40 nucleotides seen in lane d/R are probably due to cross-hybridization of oligonucleotide d with nucleotides 109 to 116 (GGGAAUCU, six matches from eight) and 118 to 123 (GAAGGU, five matches from six). We conclude that the region of SL RNA spanning the 5′ splice site is sufficiently free of protein to allow oligonucleotide-RNA hybridization and subsequent cleavage by RNase H. Oligonucleotide a, complementary to the 5′-terminal 15 nucleotides of the spliced leader, induced partial cleavage of SL RNA in deproteinized extract but almost complete cleavage in untreated extract, indicating that the end of the spliced leader is exposed in the SL RNP complex (lanes a).

The proposed secondary structure of the 3′ half of the trypanosomatid SL RNAs consists of a phylogenetically conserved, single-stranded region flanked by two stem-loops, reminiscent of the eucaryotic Sm binding site (11). To test whether this single-stranded region of *T. brucei* SL RNA is protected in an extract-dependent manner, we used oligonucleotide f against nucleotides 110 to 121 (Fig. 3A). Deproteinized RNA could be completely degraded by oligonucleotide f and RNase H, whereas in S100 extract, full protection against cleavage was observed (Fig. 3B, lanes f). This striking extract-dependent protection strongly suggests protein binding within the single-stranded region between the two 3′-terminal stem-loops. This conclusion was further supported by the previous result obtained with oligonucleotide d, showing that secondary cleavage sites within the single-stranded region were protected in S100 extract but not in free RNA (see above and Fig. 3B, lanes d). Extract-dependent protection against cleavage was also reproducibly observed between nucleotides 16 and 27 of the SL, albeit to a lesser extent (Fig. 3B, lanes b).

On the basis of the evidence of efficient RNase H protection in the 3′-terminal half of SL RNA (Fig. 3), we next investigated whether this SL RNA region is sufficient for core complex stability. First, SL RNA present in *T. brucei* S100 extract was cleaved by oligonucleotide d (Fig. 3) and RNase H. Second, the resulting fragments of the SL RNP were fractionated by CsCl density gradient centrifugation and analyzed by primer extension. Two primers were used to assay for the 3′ and 5′ halves of the SL RNA (complementary to nucleotides 125 to 144 and 31 to 45, respectively). Figure 4A displays the fractionation of the SL RNP after control RNase H cleavage with a nonspecific oligonucleotide; intact SL RNA migrated as a core complex with a buoyant density of approximately 1.43 g/ml (fractions 3 and 4; compare with Fig. 2). Significantly, after RNase H cleavage with the SL RNA-specific oligonucleotide d, the 3′ half of SL RNA remained in the form of an RNP complex resistant to the stringent conditions of CsCl density centrifugation (Fig. 4B). The mobility of this residual core complex was shifted to fractions 1 and 2, corresponding to a buoyant density of 1.39 g/ml or less. In contrast to the 3′ half, the 5′ half of the SL RNA was found at the bottom of the CsCl gradient, behaving as free RNA (Fig. 4C). We conclude therefore that the 3′-terminal half of SL RNA constitutes a

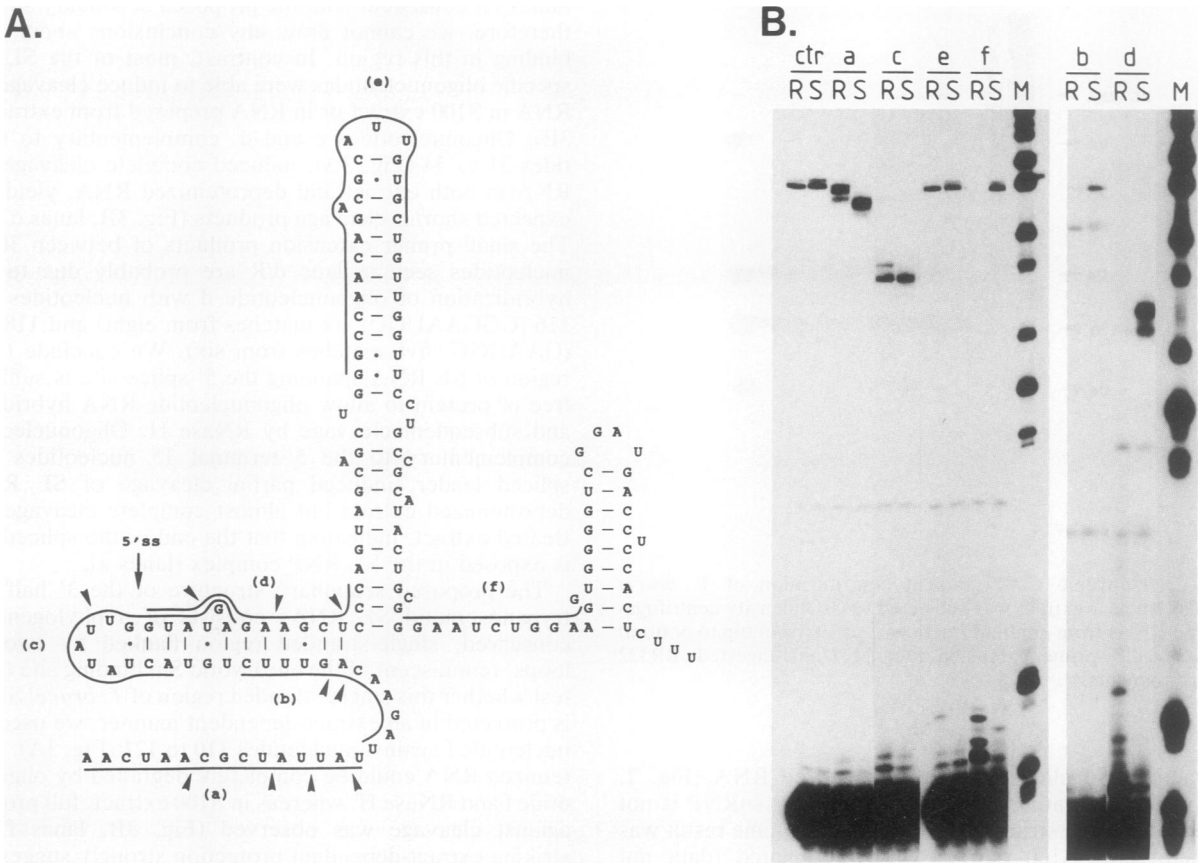


FIG. 3. RNase H protection analysis of the *T. brucei* SL RNP. (A) Proposed secondary structure of *T. brucei* SL RNA (11). Oligonucleotides a to f complementary to SL RNA are shown as solid lines. Arrowheads indicate the approximate positions of oligonucleotide-directed RNase H cleavage of SL RNA in S100 extract. The 5′ splice site (5′ss) is shown by an arrow. (B) Primer extension analysis of the SL RNA products after RNase H cleavage in deproteinized extract (lanes R) and S100 extract (lanes S). The oligonucleotides are specified above the lanes (a to f; ctr, a nonspecific oligonucleotide). Lane M, *Hpa*II-digested pBR322 marker fragments.

domain sufficient for stability of the salt-resistant SL RNP core.

**Reconstitution of the core SL RNP: evidence for common core protein components of trypanosomal snRNPs.** We used a reconstitution approach to investigate in more detail RNA-protein interactions in the SL RNP. RNA-protein binding reactions in crude S100 extract gave rise to several heterogeneously migrating complexes when analyzed by native gel electrophoresis (data not shown); therefore, reconstitution was performed in a DEAE fraction enriched in snRNPs (33; see Materials and Methods). Incubation of $^{32}$P-labeled *T. brucei* SL RNA in the DEAE fraction followed by heparin treatment resulted in the formation of one complex (Fig. 5, lane R). This complex was resistant to increased ionic strength (400 mM KCl), and its formation depended on protein since no complex was present after incubation with deproteinized extract (data not shown).

To assess the authenticity of the reconstituted SL RNP complex, we compared its electrophoretic mobility with that of the endogenous core SL RNP, which was detected by tagging with a $^{32}$P-labeled 2′-OMe RNA oligonucleotide (SL-5′) complementary to the 5′ end of SL RNA. Oligonucleotide tagging did not alter the mobility of the SL RNP complex, as demonstrated by comparison of tagging and Northern (RNA) analysis (data not shown). As shown in Fig. 5, the reconstituted and the endogenous SL RNP core complexes were found to comigrate during native gel electrophoresis (compare lanes R and T). A further characteristic of the endogenous core SL RNP is its stability during CsCl density centrifugation (Fig. 4). We confirmed that after reconstitution, a significant portion of the $^{32}$P-labeled SL RNA fractionated in the form of an RNA-protein complex with a buoyant density of 1.43 g/ml (data not shown). This value compares favorably with the estimated density of the endogenous SL RNP (Fig. 2). In conclusion, we have reconstituted an RNA-protein complex which resembles the endogenous core SL RNP by the criteria of heparin resistance, mobility on native gels, and buoyant density.

Next, we performed reconstitution reactions under competitive conditions to determine the specificity of RNA-protein binding in the core SL RNP. Addition of a 50-fold molar excess of unlabeled SL RNA was found to dramatically reduce the amount of specific complex formed; higher concentrations of competitor almost completely eliminated complex formation, whereas lower amounts competed only partially (Fig. 5, panel TbSL, and data not shown). Yeast tRNA was used as a nonspecific competitor and was found to have only a very limited effect on reconstitution, even at a 1,000-fold molar excess (panel tRNA). These results confirmed further the specific nature of the reconstituted SL RNA-protein complex. Since SL RNAs of trypanosomatids share a common secondary structure element, we also

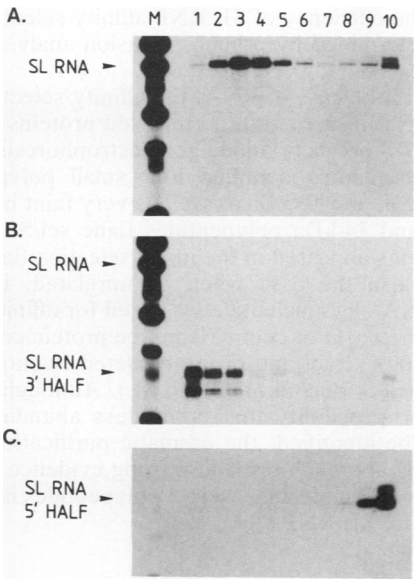


FIG. 4. Mapping of the salt-resistant core domain of the *T. brucei* SL RNP. S100 extract was first incubated with a nonspecific control oligonucleotide (A) or with the SL RNA-specific oligonucleotide d (B and C). The resulting fragments of the SL RNP were then fractionated by CsCl density gradient centrifugation. Full-length SL RNA (A) and its 3'- and 5'-terminal cleavage fragments (B and C, respectively) were detected by primer extension of RNA prepared from gradient fractions 1 to 10 (from top to bottom). The expected positions of full-length and cleaved SL RNA are indicated. Lane M, *Hpa*II-digested pBR322 marker fragments.

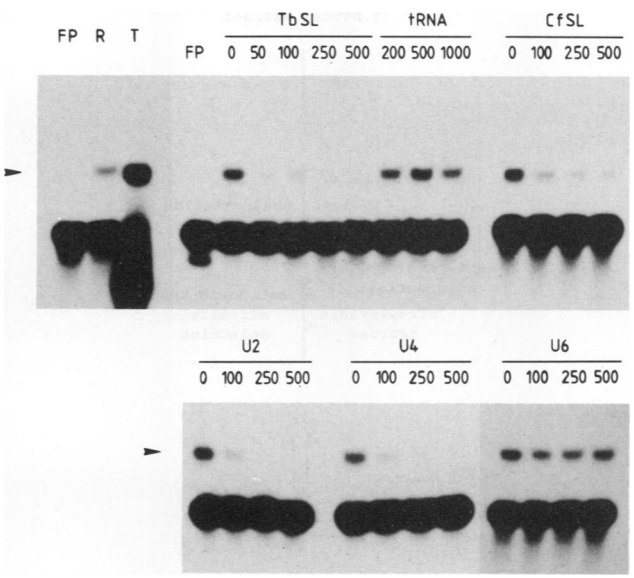


FIG. 5. Reconstitution of the core SL RNP. $^{32}$P-labeled *T. brucei* SL RNA was incubated with DEAE extract, and complex formation was analyzed by native gel electrophoresis (lane R). Endogenous SL RNP present in the DEAE extract was detected by 2'-OMe RNA oligonucleotide tagging (lane T). Lane FP, free SL RNA. The other panels show the competition of core SL RNP reconstitution by *T. brucei* SL RNA (TbSL), yeast tRNA (tRNA), *C. fasciculata* SL RNA (CfSL), and *T. brucei* U2 (U2), U4 (U4), and U6 (U6) RNAs. The molar excess of unlabeled competitor RNA is indicated. The position of the reconstituted SL RNP core complex is marked by an arrowhead.

competed SL RNA complex formation with SL RNA of *C. fasciculata*, a distant relative of *T. brucei* (37). Competition of the reconstituted SL RNP occurred, albeit at a slightly reduced efficiency compared with *T. brucei* SL RNA (panel CfSL), suggesting that *C. fasciculata* SL RNA binds the same protein(s) that interact with the *T. brucei* SL RNA.

To determine whether the protein binding observed with SL RNA is common with the other *T. brucei* snRNAs, we extended the competition experiments to *T. brucei* U2, U4, and U6 RNAs. Addition of synthetic U2 or U4 RNA to the binding reaction resulted in competition of SL RNP formation at levels comparable to those found with SL RNA (Fig. 5, panels U2 and U4). In contrast, U6 RNA was unable to compete for SL RNP formation (panel U6). These results suggest that the same protein or set of proteins bind to SL, U2, and U4 RNAs but not to U6 RNA.

**Identification of *T. brucei* SL RNP core proteins.** We next wanted to identify the protein components of the core SL RNP. UV cross-linking of $^{32}$P-labeled SL RNA to protein during in vitro reconstitution failed to detect specific label transfer to protein (data not shown); therefore, we purified the endogenous SL RNP from *T. brucei* S100 extract through a combination of DEAE chromatography, glycerol gradient sedimentation, and antisense RNA affinity chromatography. Our purification strategy is schematically presented in Fig. 6. The steps prior to RNA affinity selection were necessary to reduce the amount of contaminating RNA and protein to an acceptable level (data not shown). At each step of the procedure, total RNA content was analyzed by polyacrylamide gel electrophoresis and silver staining, and the level of SL RNA was quantitated by primer extension (Fig. 7).

S100 extract was first fractionated by anion-exchange chromatography with DEAE-Sepharose (33). The SL RNP remained bound at 100 mM KCl, whereas ca. 70% of the protein from the S100 extract was found in the flowthrough fraction. At 400 mM KCl, the SL RNP was quantitatively eluted, resulting in a ninefold purification relative to protein (data not shown). RNA analysis indicated that high-molecular-weight RNA, 7S RNA, small rRNAs, 5S rRNA, and tRNA still contaminated the SL RNA-containing fractions (Fig. 7C; compare lanes S and D). To reduce nonspecific interactions to a minimum, the subsequent purification steps were performed at increased ionic strength (800 mM KCl). Our previous experiments had demonstrated that a core SL RNP remains stable under stringent conditions (Fig. 2 and 4). This was confirmed through 2'-OMe RNA tagging and Northern hybridization, by which we could detect only one SL RNP complex in the DEAE fraction (Fig. 5, lane T, and data not shown).

The DEAE fraction was then subjected to glycerol gradient sedimentation (Fig. 7A). Primer extension assays of RNA from the glycerol gradient fractions revealed that the SL RNP complexes sedimented under these conditions at around 6S (Fig. 7B). This reduced mobility in sedimentation of the SL RNP, compared with that under low-salt conditions (Fig. 1A; SL RNA sedimented as an 8S complex), is consistent with the notion that SL RNA was at this purification step in the form of a salt-stable core SL RNP. Analysis of the total RNA content of the gradient fractions by gel electrophoresis and silver staining showed that 7S RNA, 5S rRNA, about 75% of tRNA, and a fraction of the small rRNAs still cosedimented with the SL RNA (Fig. 7A). In contrast, high-molecular-weight RNA sedimented above

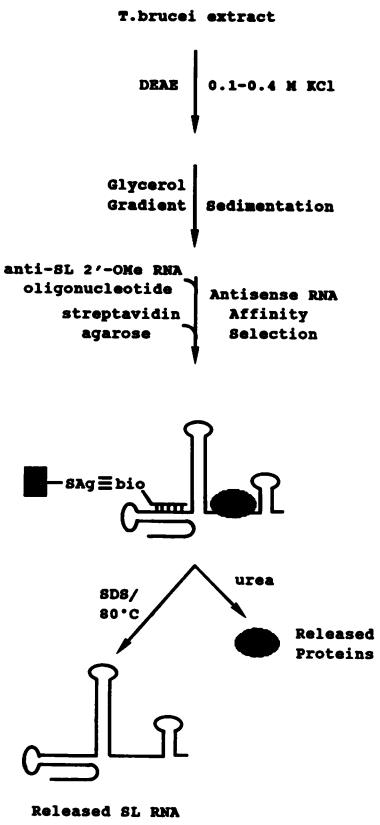


FIG. 6. Schematic summary of the purification of the *T. brucei* SL RNP.

16S, the great majority pelleting under these conditions (Fig. 7A, fractions 10 to 13 and pellet).

For further purification by antisense RNA affinity chromatography, gradient fractions 3 to 7 containing the SL RNP were pooled. On the basis of our RNase H protection experiments (Fig. 3), the nucleotide sequence immediately following the 5′ splice site of the SL RNP had been found accessible to hybridization. Therefore, a 2′-OMe RNA oligonucleotide complementary to nucleotides 40 to 54 of SL RNA (equivalent to DNA oligonucleotide d; Fig. 3A) was synthesized and biotinylated at its 3′ end. This oligonucleotide was incubated with the pooled gradient fractions, allowing hybridization with the SL RNA and subsequent recovery of the SL RNP by streptavidin-agarose affinity selection. The immobilized SL RNP was incubated with urea to elute the bound proteins and then subjected to SDS and heat treatment to release the selected RNA, all of which was done sequentially with the same sample.

Figure 7C (lane sel/SL) shows that an RNA with the apparent size of SL RNA was selected, together with only a very low level of smaller RNA species which could represent SL RNA degradation products or tRNA. As expected, selection of this RNA depended on the presence of the SL RNA-specific oligonucleotide, since no RNA could be detected from a parallel mock affinity selection without oligonucleotide (lane sel/mk). To confirm that the major RNA species was indeed SL RNA, affinity-purified RNA was cleaved with a SL-specific DNA oligonucleotide and RNase H, resulting in the loss of this RNA and the concomitant appearance of the expected cleavage products (data not shown). The efficiency of SL RNP affinity selection was ca. 33%, as determined by primer extension analysis (data not shown).

Prior to the release of RNA, the affinity-selected material was treated with urea and the released proteins were separated by SDS-polyacrylamide gel electrophoresis (Fig. 7D). Coomassie staining identified four small polypeptides of 14.8, 14, 12.5, and 10 kDa as well as very faint bands above the 14.8- and 14-kDa polypeptides (lane sel/SL). None of these proteins appeared in the mock selection (lane sel/mk). This was also the case when an unrelated, biotinylated 2′-OMe RNA oligonucleotide was used for affinity selection (data not shown). For comparison, the protein compositions of total S100 extract and of the depleted fractions are also shown (lanes S, dep/mk, and dep/SL). Although we cannot exclude the possibility that other, less abundant proteins remain to be identified, the dramatic purification achieved through this approach provides strong evidence that only a small set of low-molecular-weight polypeptides is associated with the core SL RNP.

## DISCUSSION

We have presented results on the identification and initial characterization of snRNP complexes from *T. brucei*, in particular on the SL RNP. We found all known spliceosomal snRNAs of *T. brucei* (SL, U2, U4, and U6) in the form of RNA-protein complexes, as had been determined recently for *T. brucei* SL and U2 RNAs (33). Similarly to the mammalian snRNPs, salt-stable core complexes of the SL RNP and of the U4/U6 snRNP were detected in CsCl gradients. In contrast, the U2 snRNP behaved differently and was unstable under these stringent ionic conditions. A recent study had reported on the stability of core complexes of the *T. brucei* SL RNP and U2 snRNP in $Cs_2SO_4$ gradients (33). Apparently, only the more stringent conditions of CsCl density centrifugation (see, for example, reference 24) revealed this striking difference between the mammalian and the trypanosomal U2 snRNP. We have obtained further evidence for a special core structure of the trypanosomal U2 snRNP: only the U2 snRNP was unstable at high ionic strength, whereas the SL RNP and U4/U6 snRNP were in the form of core particles, as detected by native gel electrophoresis and Northern hybridization analysis (17). Similarly to U6 RNA from other eucaryotes (6, 8, 19, 21), the *T. brucei* U6 RNA was identified as a component of the U6 snRNP and of a larger complex, most likely the U4/U6 snRNP. A large fraction of U6, but not of U4 RNA complexes, was dissociated to free RNA under the conditions of CsCl density gradient centrifugation (Fig. 2). This result is most likely explained by the instability of the U6 snRNP under stringent conditions. We have recently identified a protein in *T. brucei* extract that binds specifically to the 5′-terminal domain of U6 RNA and that may be a component of the U6 snRNP (16). In contrast to the mammalian system, in which most of the U4 and U6 RNAs are in the form of a U4/U5/U6 multi-snRNP (25), our analysis by glycerol gradient sedimentation has so far given no indication of a larger, U4/U6-containing snRNP complex in *T. brucei*. Although no U5 RNA analog is known in *T. brucei* (36), we cannot rule out the possibility that there is an additional spliceosomal snRNA without any strong homology to U5 RNA, that a putative *T. brucei* multi-snRNP represents only a minor snRNP fraction, or that it is not stable under our fractionation conditions.

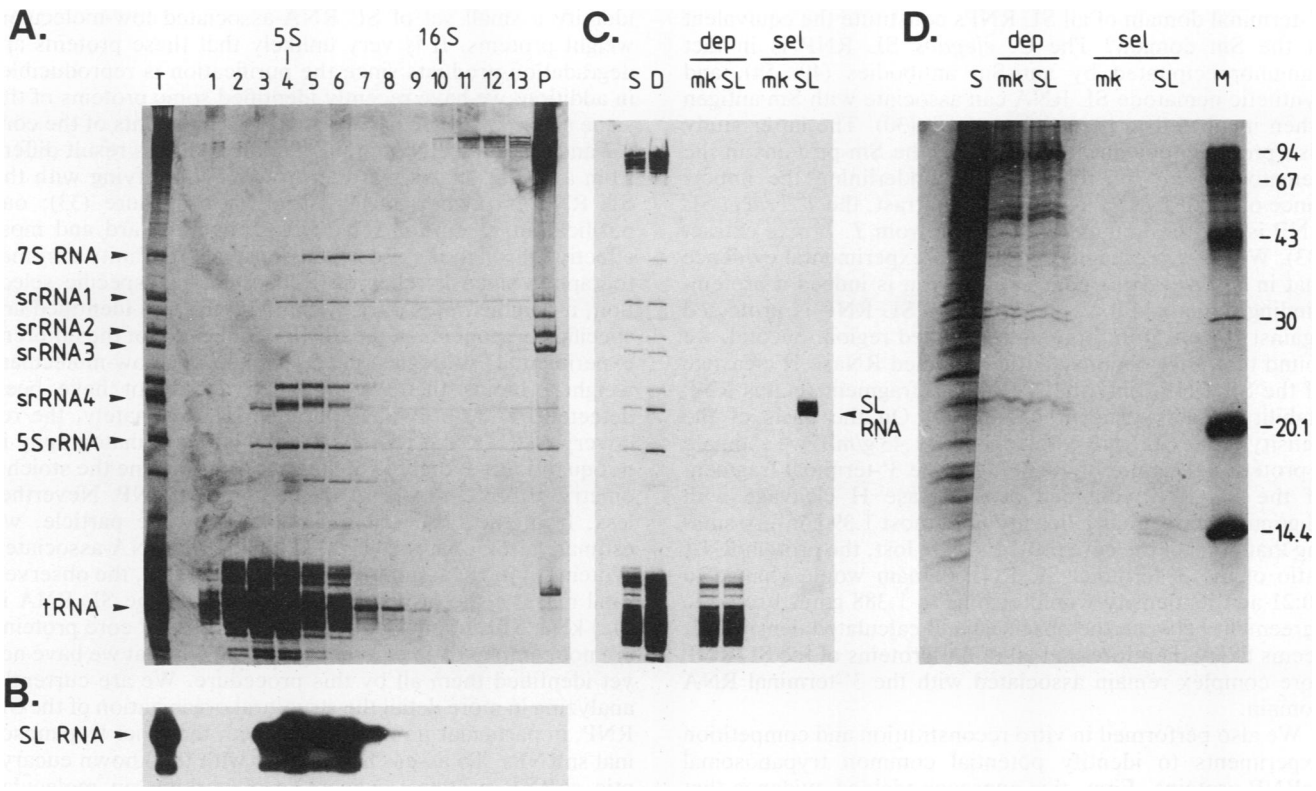


FIG. 7. Affinity purification of the *T. brucei* SL RNP and identification of its protein components. (A) Glycerol gradient sedimentation of the SL RNP-containing DEAE fraction from 50 ml of *T. brucei* S100 extract. Gradient fractions 1 to 13 (from top to bottom) were analyzed for RNA composition by denaturing gel electrophoresis and silver staining. In addition, RNA from the pelleted material (lane P) and from 10 μl of the DEAE fraction before glycerol gradient sedimentation (lane T) were analyzed. The major RNAs identified by their electrophoretic mobility are indicated at the left (small rRNAs [srRNAs]), and the positions of the size markers (5S and 16S) are shown at the top. (B) Analysis of the SL RNA distribution across the glycerol gradient by primer extension. (C) RNA analysis of SL RNP affinity purification. Gradient fractions containing the SL RNP (fractions 3 to 7) were pooled and subjected to affinity selection, using a 2′-OMe RNA oligonucleotide complementary to SL RNA (lanes SL). As a control, a mock affinity selection was done in the absence of oligonucleotide (lanes mk). RNA from 10 μl of *T. brucei* S100 extract (lane S), 10 μl of the DEAE fraction (lane D), 10 μl of the unbound fractions after affinity selection (dep), and all the affinity-selected material (sel) was analyzed by denaturing gel electrophoresis and silver staining. The affinity-selected material was equivalent to $5 \times 10^{10}$ cells. The position of the SL RNA is indicated at the right. (D) Protein analysis of affinity-purified SL RNP. Bound proteins (from the equivalent of $5 \times 10^{10}$ cells) were released from the affinity-purified SL RNP and analyzed by SDS-polyacrylamide gel electrophoresis and Coomassie blue staining (sel, lane SL). As a control, protein analysis of the mock selection ($5 \times 10^{10}$ cell equivalents) was included (sel, lane mk). For comparison, the protein composition of 4 μl of S100 extract (lane S) and of 4 μl of unbound fractions after affinity selection (dep) is shown. Lane M, protein molecular weight markers (sizes in kilodaltons are indicated on the right).

Since the SL RNP is unique to the *trans*-splicing mechanism, we concentrated our further characterization of trypanosomal snRNPs on this RNP complex, which by the criteria of sedimentation and CsCl density centrifugation also appeared most homogeneous among the *T. brucei* snRNPs. We mapped two regions in the SL RNP that are protected against RNase H in an extract-dependent manner: one in the 5′-terminal region (Fig. 3, oligonucleotide b) and one in the single-stranded region between the central and 3′-terminal stem-loops (oligonucleotide f). These regions may thus contain sites for protein binding or for interaction with another snRNP. Significantly, oligonucleotide b demonstrated protection of the SL RNP in a region previously proposed to interact with the U2 snRNP (51). Although we did not detect an SL-U2 snRNP complex in gradient-fractionated snRNPs, in crude extract the U2 snRNP or a protein factor may interact with the SL RNP. Of potential significance for the *trans*-splicing mechanism is our finding that within the SL RNP, the 5′ splice junction was accessible to RNase H probing (Fig. 3, oligonucleotides c and d). This result is in agreement with the finding that nucleotides 17 to 60 of the SL RNA are accessible in permeabilized *T. brucei* cells (52).

A comparison of the sequences from a variety of trypanosome and nematode species revealed that the 3′ half of the SL RNA can be folded into a conserved secondary structure consisting of two stem-loops separated by a single-stranded region (11). In some species, for example, *Caenorhabditis elegans* or *C. fasciculata*, this single-stranded region contains an $A(U)_nG$ sequence conforming to the Sm binding motif (28); other species, such as *T. brucei*, have a single cytidine interspersed in the U stretch of an Sm-like sequence (AUCUG, nucleotides 113 to 117; Fig. 3A). There are naturally occurring, functional Sm binding sites where the U stretch is interrupted by other nucleotides (for a review, see reference 18). In addition, an extensive genetic analysis of yeast U5 RNA demonstrated that the Sm binding site is surprisingly tolerant to mutation (23). Does therefore the

3′-terminal domain of all SL RNPs constitute the equivalent of the Sm domain? The *C. elegans* SL RNP is in fact immunoprecipitated by anti-Sm antibodies (49, 54), and synthetic nematode SL RNA can associate with Sm antigen when incubated in nematode extract (30). The latter study also provided evidence for a role of the Sm proteins in the nematode *trans*-splicing reaction, underlining the importance of this SL RNP domain. In contrast, the *T. brucei* SL RNP is not Sm immunoprecipitable from *T. brucei* extract (33). We have presented two lines of experimental evidence that in *T. brucei* this conserved region is indeed a protein-binding domain. First, the *T. brucei* SL RNP is protected against RNase H in this single-stranded region. Second, we found that after oligonucleotide-directed RNase H cleavage of the SL RNP, only the 3′-terminal fragment retains RNP stability under stringent conditions. On the basis of the density of the SL RNP core complex (1.43 g/ml), we estimate a protein/RNA ratio of 70:30 (44). The 3′-terminal fragment of the SL RNP obtained after RNase H cleavage with oligonucleotide d has a density of at most 1.39 g/ml. Assuming that none of the core proteins were lost, the protein/RNA ratio of the 3′-terminal SL RNP domain would change to 70:21 and its density would change to 1.388 g/ml. From the agreement between the observed and calculated densities, it seems likely therefore that all of the proteins of the SL RNP core complex remain associated with the 3′-terminal RNA domain.

We also performed in vitro reconstitution and competition experiments to identify potential common trypanosomal snRNP proteins. First, this approach yielded evidence that the SL RNP core structure may be conserved between *T. brucei* and *C. fasciculata*. Although the *C. fasciculata* SL RNA contains a functional Sm binding site when assayed in mammalian extract (11), the SL RNP is not immunoprecipitated from *Crithidia* extract by Sm antibodies (our unpublished observations). The trypanosome SL RNPs therefore appear to possess a core domain which, although antigenically distinct from, shows certain similarities with the Sm domain of other eucaryotic snRNPs, namely, RNA secondary structure and stable protein binding. The development of an in vitro *trans*-splicing system derived from trypanosomes will be required to evaluate the possible functional significance of this domain in *trans* splicing. Second, competition assays demonstrated that the same protein or set of proteins bind to the *T. brucei* SL, U2, and U4 RNAs, suggesting that the SL RNP and the U2 and U4/U6 snRNPs have similar core structures. Indeed, both U2 and U4 RNAs have the potential to form a 3′ domain with a secondary structure similar to that of the SL RNA, that is, a single-stranded region flanked by stem-loops, although no obvious sequence similarities can be identified between them. Significantly, despite this evidence for a common core RNA-protein domain, the U2 snRNP differs from the SL RNP and the U4/U6 snRNP with respect to its instability under conditions of high ionic strength. Whether this reflects the absence of certain core protein components in the U2 snRNP or U2 RNA-specific protein-binding determinants can be definitively decided only after the biochemical analysis of purified *T. brucei* snRNPs. It will be of particular interest to analyze in more detail the structure, protein composition, and interactions of the *T. brucei* U2 snRNP, since this particle appears to have specially adapted to the requirements of the *trans*-splicing mechanism and may thus reveal some of the specific features of this RNA processing reaction.

Finally, by purifying the salt-stable core SL RNP through antisense RNA affinity chromatography, we were able to identify a small set of SL RNA-associated low-molecular-weight proteins. It is very unlikely that these proteins are degradation products since the purification is reproducible; in addition, we have recently identified some proteins of the same apparent molecular weights as components of the core U2 and U4/U6 snRNPs (41). We note that this result differs from a recent analysis of the proteins copurifying with the SL RNP through a highly stringent procedure (33); our purification scheme is relatively straightforward and most effective through the use of antisense RNA affinity chromatography. Since it relies on RNA sequence-specific selection, it demonstrates directly that the proteins identified are specific components of the SL RNP. Because of the different experimental strategies used, the set of low-molecular-weight proteins that we identified could not have been detected by the previous approach. Unfortunately, the recovery of RNA and protein after affinity chromatography is not quantitative; therefore, we cannot determine the stoichiometry of protein to RNA in the core SL RNP. Nevertheless, from the observed density of the core particle, we estimate a total mass of 87 kDa for the SL RNA-associated proteins (44). By assuming a 1:1 stoichiometry, the observed total mass of the proteins copurifying with the SL RNA is 51.3 kDa. This implies either that the SL RNP core proteins are not complexed in equimolar amounts or that we have not yet identified them all by this procedure. We are currently analyzing in more detail the structural organization of the SL RNP, in particular its relationship with the other trypanosomal snRNPs. To assess homologies with the known eucaryotic snRNP proteins, we are also working on molecular cloning of the *T. brucei* genes for the newly identified snRNP proteins.

## ACKNOWLEDGMENTS

We acknowledge the excellent technical assistance of Björn Wieland, Stephanie Peczynski, Tse-I Lin, and Monika Witte. We thank Uli Göringer for essential help with the 2′-OMe RNA oligonucleotide synthesis, Piet Borst and Marc Ouelette for advice and trypanosome stocks, Reinhard Lührmann for anti-$m_3G$ antibodies, and Albert Cornelissen, Christian Tschudi, and members of our laboratory for critical comments and advice.

This work was supported by the Deutsche Forschungsgemeinschaft.

## REFERENCES

1. **Agabian, N.** 1990. *Trans* splicing of nuclear pre-mRNAs. Cell **61**:1157–1160.
2. **Barabino, S. M. L., B. S. Sproat, U. Ryder, B. J. Blencowe, and A. I. Lamond.** 1989. Mapping U2 snRNP-pre-mRNA interaction using biotinylated oligonucleotides made of 2′-OMe RNA. EMBO J. **8**:4171–4178.
3. **Bindereif, A., and M. R. Green.** 1986. Ribonucleoprotein complex formation during pre-mRNA splicing in vitro. Mol. Cell. Biol. **6**:2582–2592.
4. **Bindereif, A., and M. R. Green.** 1987. An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. EMBO J. **6**:2415–2424.
5. **Bindereif, A., and M. R. Green.** 1990. Identification and functional analysis of mammalian splicing factors, p. 201–224. *In* J. K. Setlow (ed.), Genetic engineering—principles and methods, vol. 12. Plenum Publishing Corp., New York.
6. **Black, D. L., and A. L. Pinto.** 1989. U5 small nuclear ribonucleoprotein: RNA structure analysis and ATP-dependent interaction with U4/U6. Mol. Cell. Biol. **9**:3350–3359.
7. **Borst, P.** 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annu. Rev. Biochem. **55**:701–732.
8. **Bringmann, P., B. Appel, J. Rinke, R. Reuter, H. Theissen, and**

**R. Lührmann.** 1984. Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing. EMBO J. **3**:1357–1363.

9. **Brun, R., and M. Schönenberger.** 1979. Cultivation and *in vitro* cloning of procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Acta Trop. **36**:289–292.
10. **Bruzik, J. P., and J. A. Steitz.** 1990. Spliced leader RNA sequences can substitute for the essential 5′ end of U1 RNA during splicing in a mammalian *in vitro* system. Cell **62**:889–899.
11. **Bruzik, J. P., K. Van Doren, D. Hirsh, and J. A. Steitz.** 1988. *Trans* splicing involves a novel form of small nuclear ribonucleoprotein particles. Nature (London) **335**:559–562.
12. **Cross, G. A. M.** 1975. Identification, purification and properties of clone specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. Parasitology **71**:393–417.
13. **Dignam, D. L., 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.
14. **Freistadt, M. S., G. A. M. Cross, A. D. Branch, and H. D. Robertson.** 1987. Direct analysis of the mini-exon donor RNA of *Trypanosoma brucei*: detection of a novel cap structure also present in messenger RNA. Nucleic Acids Res. **15**:9861–9879.
15. **Freistadt, M. S., G. A. M. Cross, and H. D. Robertson.** 1988. Discontinuously synthesized mRNA from *Trypanosoma brucei* contains the highly methylated 5′ cap structure, $m^7$GpppA* A*C(2′-O)mU*A. J. Biol. Chem. **263**:15071–15075.
16. **Gröning, K., Z. Palfi, S. Gupta, M. Cross, T. Wolff, and A. Bindereif.** 1991. A new U6 small nuclear ribonucleoprotein-specific protein conserved between *cis*- and *trans*-splicing systems. Mol. Cell. Biol. **11**:2026–2034.
17. **Günzl, A., M. Cross, and A. Bindereif.** Submitted for publication.
18. **Guthrie, C., and B. Patterson.** 1988. Spliceosomal snRNAs. Annu. Rev. Genet. **22**:387–419.
19. **Hamm, J., and I. W. Mattaj.** 1989. An abundant U6 snRNP found in germ cells and embryos of *Xenopus laevis*. EMBO J. **8**:4179–4187.
20. **Hartshorne, T., and N. Agabian.** 1990. A new U2 RNA secondary structure provided by phylogenetic analysis of trypanosomatid U2 RNAs. Genes Dev. **4**:2121–2131.
21. **Hashimoto, C., and J. A. Steitz.** 1984. U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle. Nucleic Acids Res. **12**:3283–3293.
22. **Huang, X.-Y., and D. Hirsh.** 1989. A second *trans*-spliced RNA leader in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **81**:2713–2717.
23. **Jones, M. H., and C. Guthrie.** 1990. Unexpected flexibility in an evolutionarily conserved protein:RNA interaction: genetic analysis of the Sm binding site. EMBO J. **9**:2555–2561.
24. **Karn, J., G. Vidali, L. C. Bofa, and V. G. Allfrey.** 1977. Characterization of the non-histone nuclear protein associated with rapidly labeled heterogeneous nuclear RNA. J. Biol. Chem. **252**:7307–7322.
25. **Konarska, M. M., and P. A. Sharp.** 1987. Interactions between small ribonucleoprotein particles in formation of spliceosomes. Cell **49**:763–774.
26. **Krainer, A. R., and T. Maniatis.** 1988. RNA splicing, p. 131–206. *In* B. D. Hames and D. M. Glover (ed.), Transcription and splicing. IRC Press, Oxford.
27. **Lelay-Taha, M.-N., I. Reveillaud, D. M. Sri-Widada, C. Brunel, and P. Jeanteur.** 1986. RNA-protein organization of U1, U5 and U4-U6 small nuclear ribonucleoproteins in HeLa cells. J. Mol. Biol. **189**:519–532.
28. **Liautard, J., D. M. Sri-Widada, C. Brunel, and P. Jeanteur.** 1982. Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. J. Mol. Biol. **162**:623–643.
29. **Lührmann, R.** 1988. snRNP proteins, p. 71–99. *In* M. L. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Berlin.
30. **Maroney, P. A., G. J. Hannon, J. A. Denker, and T. W. Nilsen.** 1990. The nematode spliced leader RNA participates in *trans*-splicing as an Sm snRNP. EMBO J. **9**:3667–3673.
31. **Mattaj, I. W.** 1988. UsnRNP assembly and transport, p. 100–114. *In* M. L. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Berlin.
32. **McCarthy-Burke, C., Z. A. Taylor, and G. A. Buck.** 1989. Characterization of the spliced leader genes and transcripts in *Trypanosoma cruzi*. Gene **82**:177–189.
33. **Michaeli, S., T. G. Roberts, K. P. Watkins, and N. Agabian.** 1990. Isolation of distinct small ribonucleoprotein particles containing the spliced leader and U2 snRNAs of *Trypanosoma brucei*. J. Biol. Chem. **265**:10582–10588.
34. **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.
35. **Milligan, J. F., and O. C. Uhlenbeck.** 1989. Synthesis of small RNAs using T7 RNA polymerase. Methods Enzymol. **180**:51–62.
36. **Mottram, J., K. L. Perry, P. M. Lizardi, R. Lührmann, N. Agabian, and R. G. Nelson.** 1989. Isolation and sequence of four small nuclear U RNA genes of *Trypanosoma brucei* subsp. *brucei*: identification of the U2, U4 and U6 RNA analogs. Mol. Cell. Biol. **9**:1212–1223.
37. **Muhich, M. L., D. E. Hughes, A. M. Simpson, and L. Simpson.** 1987. The monogenetic kinetoplastid protozoan, *Crithidia fasciculata*, contains a transcriptionally active, multicopy mini-exon sequence. Nucleic Acids Res. **15**:3141–3153.
38. **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.
39. **Nelson, K. K., and M. R. Green.** 1989. Mammalian U2 snRNP has a sequence-specific RNA-binding activity. Genes Dev. **3**:1562–1571.
40. **Nilsen, T. W., J. Shambaugh, J. Denker, G. Chubb, C. Faser, L. Putnam, and K. Bennett.** 1989. Characterization and expression of a spliced leader RNA in the parasitic nematode *Ascaris lumbricoides* var *suum*. Mol. Cell. Biol. **9**:3543–3547.
41. **Palfi, Z., A. Günzl, M. Cross, and A. Bindereif.** 1991. Affinity purification of *Trypanosoma brucei* snRNPs reveals common and specific protein components. Proc. Natl. Acad. Sci. USA, in press.
42. **Perry, K. L., K. P. Watkins, and N. Agabian.** 1987. Trypanosome mRNAs have unusual "cap 4" structures acquired by addition of a spliced leader. Proc. Natl. Acad. Sci. USA **84**:8190–8194.
43. **Pikielny, C. W., A. Bindereif, and M. R. Green.** 1989. *In vitro* reconstitution of snRNPs: a reconstituted U4/U6 snRNP participates in splicing complex formation. Genes Dev. **3**:479–487.
44. **Spirin, A. S.** 1969. Informosomes. Eur. J. Biochem. **10**:20–35.
45. **Sproat, B. S., A. I. Lamond, B. Beijer, P. Neuner, and U. Ryder.** 1989. Highly efficient chemical synthesis of 2′-*O*-methyloligoribonucleotides and tetrabiotinylated derivatives; novel probes that are resistant to degradation by RNA or DNA specific nucleases. Nucleic Acids Res. **17**:3373–3384.
46. **Steitz, J. A., D. L. Black, V. Gerke, K. A. Parker, A. Krämer, D. Frendewey, and W. Keller.** 1988. Functions of the abundant U-snRNPs, p. 115–154. *In* M. L. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Berlin.
47. **Sutton, R. E., and J. C. Boothroyd.** 1986. Evidence for *trans* splicing in trypanosomes. Cell **47**:527–535.
48. **Sutton, R. E., and J. C. Boothroyd.** 1988. The cap of both miniexon-derived RNA and mRNA of trypanosomes is 7-methylguanosine. Mol. Cell. Biol. **8**:494–496.
49. **Thomas, J. D., R. C. Conrad, and T. Blumenthal.** 1988. The *C. elegans trans*-spliced leader RNA is bound to Sm and has a trimethylguanosine cap. Cell **54**:533–539.
50. **Tschudi, C., A. R. Krainer, and E. Ullu.** 1988. A U6 small nuclear RNA from *Trypanosoma brucei*. Nucleic Acids Res. **16**:11375.

51. **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.
52. **Tschudi, C., and E. Ullu.** 1990. Destruction of U2, U4, or U6 small nuclear RNA blocks *trans* splicing in trypanosome cells. Cell **61:**459–466.
53. **Tschudi, C., S. P. Williams, and E. Ullu.** 1990. Conserved sequences in the U2 snRNA-encoding genes of Kinetoplastida do not include the putative branchpoint recognition region. Gene **91:**71–77.
54. **Van Doren, K., and D. Hirsh.** 1988. *Trans*-spliced leader RNA exists as small nuclear ribonucleoprotein particles in *Caenorhabditis elegans*. Nature (London) **335:**556–559.