The EMBO Journal vol.9 no.11 pp.3667-3673, 1990

The nematode spliced leader RNA participates in *trans*-splicing as an Sm snRNP

Patricia A.Maroney, Gregory J.Hannon, John A.Denker and Timothy W.Nilsen

Department of Molecular Biology and Microbiology, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA .

Communicated by J.A.Steitz

The *trans*-spliced leader RNA (SL RNA) of nematodes resembles U snRNAs both in cap structure and in the presence of a consensus Sm binding site. We show here that synthetic SL RNA, synthesized by *in vitro* transcription, is efficiently used as a spliced leader donor in *trans*-splicing reactions catalyzed by a cell free extract prepared from developing embryos of the parasitic nematode, *Ascaris lumbricoides*. Efficient utilization of synthetic SL RNA requires a functional Sm binding site. Mutations within the Sm binding sequence that prevent immunoprecipitation by Sm antisera and prevent cap trimethylation abolish *trans*-splicing. The effect on *trans*-splicing is not due to undermethylation of the cap structure.

Key words: Ascaris lumbricoides/*trans*-splicing/Sm antigens/snRNP/spliced leader RNA

Introduction

Trans-splicing in nematodes involves the utilization of a spliced leader transcript (SL RNA) which shares several features in common with the U snRNAs necessary for *cis*-splicing (see Blumenthal and Thomas, 1988; Nilsen, 1989 for review). In particular, nematode SL RNAs possess trimethylguanosine cap structures and consensus recognition signals for proteins with Sm antigenic determinants. The Sm recognition sequence in the *Caenorhabditis elegans* SL RNA has been shown to be functional, since this SL RNA is precipitable from *C.elegans* extracts with Sm antisera and becomes Sm precipitable upon incubation in extracts derived from human cells (Bruzik *et al.*, 1988; Thomas *et al.*, 1988; Van Doren and Hirsh, 1988).

Extensive studies in a variety of systems have shown that the Sm consensus sequence serves as binding site for a core group of proteins common to Sm snRNPs. Binding of these proteins appears to be a prerequisite for both cap trimethylation and nuclear accumulation of Sm snRNPs and may be required for the efficient binding of proteins specific to individual snRNPs (reviewed in Lührmann, 1988; Mattaj, 1988). Little is known about the function, if any, of core Sm proteins in formation of, or catalysis by, spliceosomes. It is clear, however, that at least some snRNAs must be assembled into snRNPs in order to carry out their functions in splicing. For example, U1 recognizes splice donor sites only as an snRNP (Mount *et al.*, 1983) and this recognition is facilitated by the C protein which is specific to U1 snRNPs (Heinrichs *et al.*, 1990). Similarly, recognition of splice acceptor sites is mediated by a protein component of the U5 snRNP (Chabot *et al.*, 1985; Gerke and Steitz, 1986; Tazi *et al.*, 1986; Anderson *et al.*, 1989). Recent successes in reconstitution of functional snRNPs *in vitro* (Fabrizio *et al.*, 1989; McPheeters *et al.*, 1989; Bindereif *et al.*, 1990) and *in vivo* (Pan and Prives, 1988; Hamm *et al.*, 1989; You and Weiner, 1989) promise to shed more light on the role of snRNP proteins in the splicing process.

The striking similarity of nematode SL RNAs to U snRNAs has led several groups to propose that SL RNAs may represent a novel class of snRNP in which an exon (the SL) has been fused to an snRNA-like sequence (Sharp, 1987; Bruzik *et al.*, 1988; Thomas *et al.*, 1988; Van Doren and Hirsch, 1988). Further, it has been suggested that SL snRNPs may have a dual role in the *trans*-splicing process i.e. donating an exon and simultaneously functioning as a U1 snRNA (Bruzik *et al.*, 1988).

We have previously shown that cell-free extracts derived from *Ascaris lumbricoides* embryos catalyze accurate and efficient *trans*-splicing of the *A.lumbricoides* SL RNA to an authentic *trans*-spliced pre-mRNA (Hannon *et al.*, 1990). In these experiments the SL was donated from SL RNA present in the extract. Here, we show that exogenous SL RNA synthesized by *in vitro* transcription can serve as an SL donor in the *trans*-splicing reaction. *Trans*-splicing is dependent upon the addition of a synthetic RNA containing a competent *trans*-splice acceptor site. To begin to assess the sequences within the SL RNA which are required for *trans*-splicing we used site directed mutagenesis to alter the SL RNA Sm binding site. Point mutations within this site abolish *trans*-splicing. This effect was not due to the absence of a trimethylguanosine cap structure. We infer from these results that the SL RNA must be assembled into an Sm snRNP in order to participate in *trans*-splicing.

Results

Synthetic SL RNA can serve as a donor for trans-splicing

To determine if synthetic SL RNA could participate in *trans*-splicing reactions, we assessed the ability of this RNA to donate the 22 nucleotide SL sequence to a 345 nt unlabeled synthetic pre-mRNA containing a competent *trans*-splice acceptor site. We have previously shown that the 345 nt RNA can participate in *trans*-splicing reactions with endogenous SL RNA (Hannon *et al.*, 1990). When a fixed amount of synthetic SL RNA and varying amounts of unlabeled pre-mRNA were co-incubated in embryo extract, three specific labeled RNAs accumulated in levels proportional to the amount of acceptor RNA included in each reaction (Figure 1B, lanes 2-5). No reaction products were evident either when acceptor RNA (Figure 1B, lane 1) or ATP were omitted from incubations (Figure 1C, lane 7).

As illustrated in Figure 1, panel A, a *trans*-splicing reaction between labeled SL RNA and the acceptor pre-

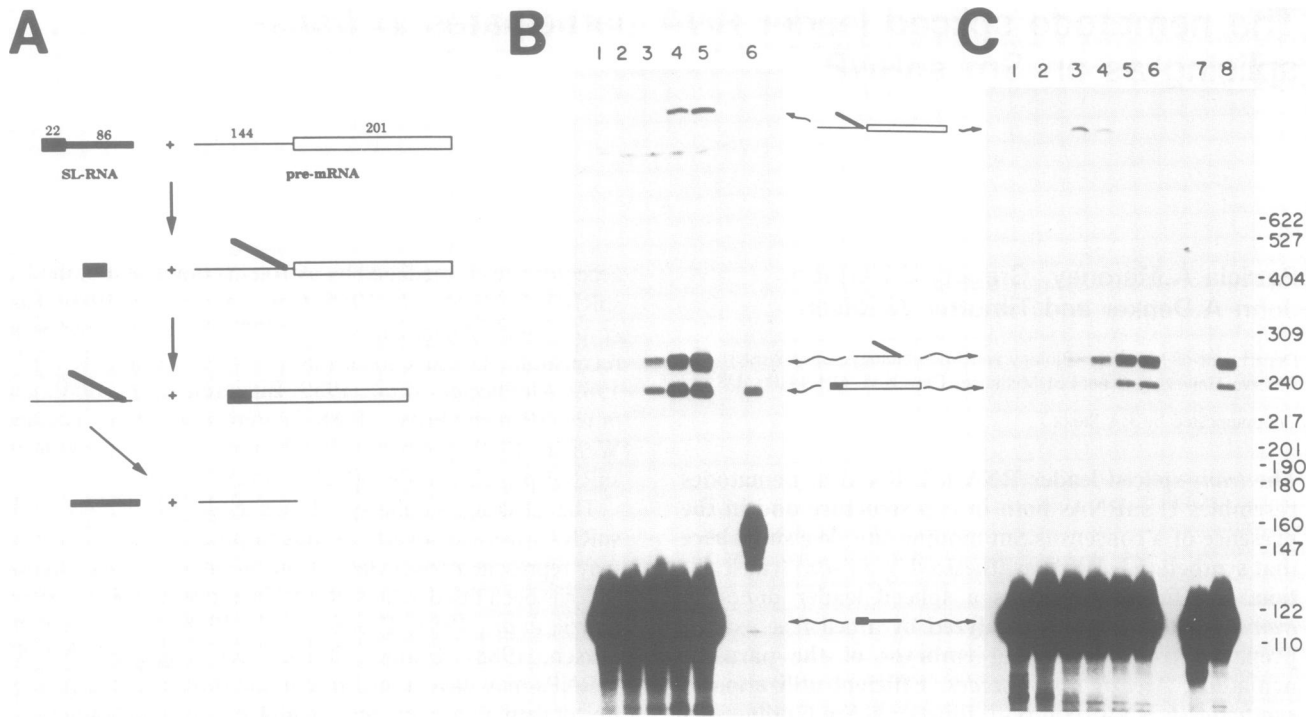

Fig. 1. Synthetic SL RNA functions as an SL donor in *trans*-splicing. **A**. Schematic representation of a *trans*-splicing reaction between SL RNA and a synthetic 345 nt pre-mRNA containing a *trans*-splice acceptor site (see text). **B**. Dependence of *trans*-splicing upon inclusion of acceptor molecules. Splicing reactions contained 200 000 c.p.m. (5 ng) of labeled SL RNA and 0 ng, lane 1; 10 ng, lane 2; 20 ng, lane 3; 50 ng, lane 4 or 100 ng, lane 5 of unlabeled 345 nt acceptor RNA. In lane 6, reactions contained 20 000 c.p.m. (3 ng) of an SL RNA which had been labeled in the cap structure using vaccinia guanyl transferase. This RNA contained a 34 base 3′ extension (see Materials and methods). Reactions were assembled, incubated for 60 min at 30°C and analyzed as described in Materials and methods. **C**. Time course of *trans*-splicing reactions containing labeled SL RNA. Each incubation contained 200 000 c.p.m. (5 ng) of labeled SL RNA and 50 ng of unlabeled acceptor RNA. Reactions were incubated at 30°C for 0 min, lane 1; 10 min, lane 2; 20 min, lane 3; 30 min, lane 4; 60 min, lane 5; or 90 min, lane 6, before being analyzed as described in Materials and methods. Lanes 7 and 8 are reactions identical to those in lane 5 with the exception that ATP was omitted from the reaction in lane 7. The identities of the individual reaction products are indicated schematically. The characterization of RNAs is described in the text. Indicated sizes correspond to the mobilities of labeled restriction fragments electrophoresed in parallel lanes to the reactions in C.

mRNA would yield two labeled intermediates (Y-intron + exon and Y-intron) and two labeled end-products (a 223 nt *trans*-spliced RNA and an 86 nt 'debranched' intron derived from the 'intron' portion of the SL RNA). The three RNAs produced during processing reactions had apparent sizes of ~230 nt, ~270 nt and >1 kb, (Figure 1). A time course of incubation showed that the ~230 nt and the ~270 nt RNA accumulated throughout the processing reaction while levels of the >1 kb RNA peaked early in the reaction and then declined (Figure 1C, lanes 1–6). The ~230 nt RNA was tentatively identified as *trans*-spliced product because of its mobility, kinetics of accumulation and the fact that just this RNA appeared in reactions containing SL RNA labeled only in the cap structure (Figure 1B, lane 6). Based upon similar considerations, it seemed likely that the >1 kb RNA represented the Y-intron + exon intermediate, and possible that the ~270 nt RNA could represent the Y-intron intermediate.

To confirm these interpretations, we performed a series of analyses similar to those used previously to establish the identities of *trans*-spliced products and intermediates generated with labeled acceptor molecules (Hannon *et al.*, 1990). The three RNAs were excised individually from preparative gels and digested with RNase H in the presence of each of four oligodeoxynucleotides complementary to either the 22 nt SL sequence, the 'intron' portion of the SL RNA, the exon portion of the acceptor RNA, or the intron portion of the acceptor molecule (Hannon *et al.*, 1990). The digestion patterns obtained were those expected for the *trans*-spliced product and the two Y-branched intermediates (data not shown). In addition, each RNA was also incubated with HeLa cell S100 extract containing 'debranching' enzyme. Whereas these incubations did not alter the ~230 nt RNA, treatment of the suspected intermediates with 'debranching' extract produced RNAs whose size was ~86 nt, (data not shown).

To characterize the RNAs in more detail, we performed two dimensional RNase T1 fingerprint analysis of the *trans*-spliced product RNA as well as the Y-intron intermediate both before and after 'debranching'. The ~230 nt RNA contained only four labeled RNase T1 oligonucleotides (a capped oligonucleotide and the oligonucleotides UUUAAU-UACCCAAGp, UUUGp, and AGp) (Figure 2C). These four oligonucleotides were exactly those expected from an RNA containing the labeled 22 nt SL sequence at its 5′ end.

The fingerprint of the presumptive Y-intron intermediate before 'debranching' revealed that this RNA contained the balance of the RNase T1 oligonucleotides expected to arise from the SL RNA as well as two novel oligonucleotides (see spots marked by arrows in Figure 2D). The migration of these two oligonucleotides was coincident with the migration of oligonucleotides previously identified as branched oligonucleotides when labeled acceptor molecules were analyzed (Hannon *et al.*, 1990). Upon debranching, the two oligonucleotides disappeared and were replaced by a single oligonucleotide which did not have a counterpart in the

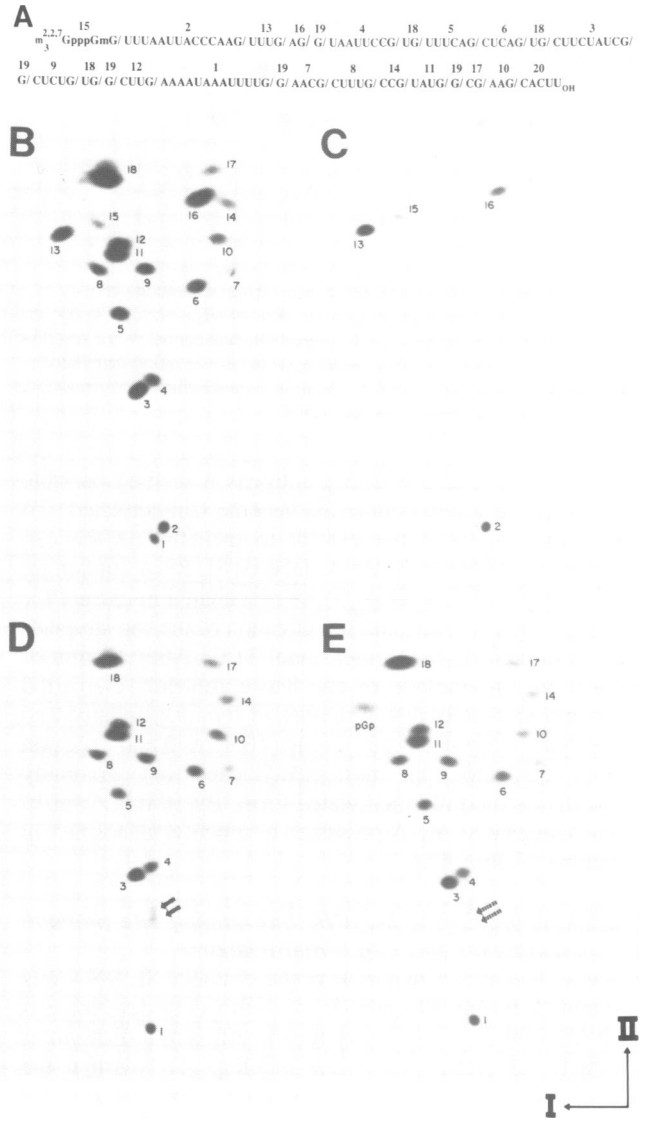

Fig. 2. RNase T1 fingerprint analysis of *trans*-spliced product and Y-intron intermediate. **A**. The sequence of the labeled SL RNA used in these experiments (Nilsen *et al.*, 1989) with slashes separating oligonucleotides expected to arise following digestion with RNase T1. **B**. RNase T1 fingerprint of [α-^{32}P]GTP-labeled SL RNA with oligonucleotides numbered with reference to A. Since this RNA was synthesized in the presence of [^{32}P]GTP (see Materials and methods) the oligonucleotides numbered 19 and 20 in A are not present. **C**. RNase T1 fingerprint of the *trans*-spliced product RNA. **D**. RNase T1 fingerprint of the Y-intron intermediate prior to debranching. Arrows indicate two novel oligonucleotides not present in the substrate SL RNA. **E**. RNase T1 fingerprint of the Y-intron RNA following debranching. Dashed arrows indicate the position of the two oligonucleotides in (D) sensitive to debranching enzyme. Disappearance of these oligonucleotides upon debranching is accompanied by the appearance of pGp (indicated) (see text). Identification of individual oligonucleotides has been previously described (Maroney *et al.*, 1990). RNAs were eluted from preparative gels and fingerprinted exactly as described (Hannon *et al.*, 1989). The first (I) and second (II) dimensions (electrophoresis at pH 3.5 and homochromatography respectively) are indicated.

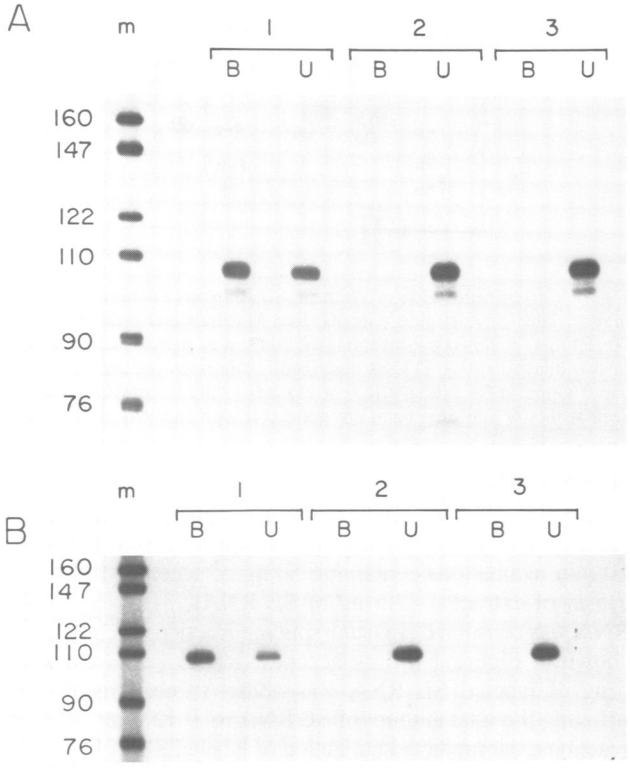

Fig. 3. Synthetic SL RNA contains a functional Sm binding site. **A**. 5 ng (200 000 c.p.m.) of labeled SL RNA (see Materials and methods) was mixed with 7.5 μl of embryo extract and incubated at 30°C as described in Materials and methods. After 60 min, equal aliquots of the incubation were immunoprecipitated with anti Sm [serum SK (1)], anti U1 [serum DO(2)] or control sera (serum Me BL) (3) as described (Materials and methods) (sera were provided by Joan Steitz, Yale University). Following immunoprecipitation, bound and unbound RNAs as indicated were deproteinized, fractionated on a 6% denaturing polyacrylamide gel and analyzed by autoradiography. **B**. Northern blot analysis: Incubations lacking labeled SL RNA (otherwise identical to A) were immunoprecipitated as in A. Bound and unbound fractions as indicated were deproteinized and fractionated as in A before transfer to Gene Screen (NEN). The blot was hybridized with a labeled oligonucleotide complementary to bases 55–73 of *A.lumbricoides* SL RNA. Lanes M: labeled restriction fragments of known size (indicated).

substrate RNA. If the branched molecule was generated by the *trans*-splicing reaction shown in Figure 1A, debranching would be expected to yield pGp. Comparison of the migration of the novel oligonucleotide with the mobility of labeled pGp analyzed in parallel indicated that the novel oligonucleotide was in fact pGp. The label in pGp is derived from the 5′ phosphate of guanosine 23 of the SL RNA and thus cleavage at the splice donor site of the SL RNA yielded 3′ hydroxyl-5′ phosphate termini with the 5′ phosphate used for the 2′5′ phosphodiester bond of the branch. We concluded from these analyses that synthetic SL RNA could serve as a SL donor to a *trans*-splice acceptor in a processing reaction identical to that previously characterized for endogenous SL RNA (Hannon *et al.*, 1990).

Given the similarity between the reactions involving SL RNA present in the extract and synthetic SL RNA, it was of interest to compare the relative efficiencies with which the two populations of SL RNAs were used. For such a comparison it was necessary to determine the level of endogenous SL RNA and its utilization as a *trans*-splice donor under defined conditions. The amount of endogenous SL RNA was determined by comparing the level of primer extension products generated using RNA prepared from a fixed amount of extract with primer extension products generated using known amounts of synthetic SL RNA. This analysis indicated that 1 μl of extract contained ~1 ng of SL RNA (data not shown). We then performed *trans*-splicing reactions with different amounts (10–100 ng) of labeled

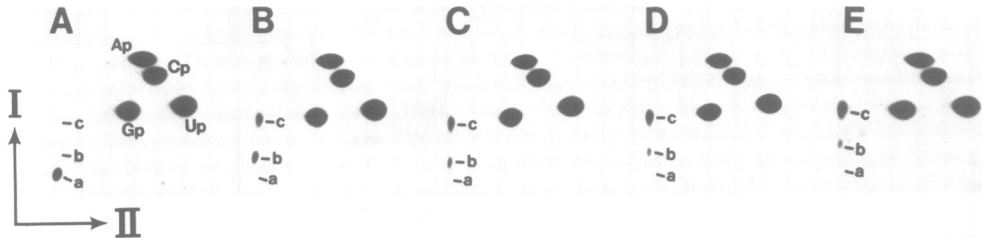

Fig. 4. Cap trimethylation of synthetic SL RNA in *A.lumbricoides* embryo extract. Synthetic SL RNA (200 000 c.p.m., 5 ng) containing the unmethylated cap structure (GpppG) (see Materials and methods) was incubated with 7.5 μl of embryo extract for 0, (A), 10 (B), 20 (C), 30 (D) or 60 (E) min. Following incubation, the RNA was deproteinized and digested to completion with RNase T2. After digestion, nucleotides were resolved in two dimensions by TLC as described in Materials and methods. Spot (a) is the unmethylated cap GpppGp while spot (b) is the monomethylated cap ^{m7}GpppGmGp and spot (c) corresponds to the trimethylated cap ^{m2,2,7}GpppGmGp (Maroney *et al.*, 1990). Additional minor digestion products seen at the 60 min time point are not cap related and may reflect uncharacterized internal modifiations of the SL RNA.

acceptor RNA and determined that the production of *trans*-spliced RNA (and thus consumption of SL RNA) varied linearly with the amount of input acceptor RNA. In 60 min incubations containing 50 ng of acceptor RNA and 7.5 μl of extract (~7.5 ng of SL RNA), ~0.4 ng of SL RNA was consumed. In *trans*-splicing reactions such as those shown in Figure 1 (50 ng acceptor RNA, 7.5 μl extract), 5 ng of labeled SL RNA was added to each incubation bringing the total amount of SL RNA to ~12.5 ng. In such reactions, the predicted utilization of labeled SL RNA would be ~0.16 ng if synthetic SL RNA and endogenous SL RNA were used equivalently. By quantitating the amount of *trans*-spliced product, we determined that ~0.11 ng of synthetic SL RNA was actually consumed, a value in reasonable agreement with the predicted value. We therefore concluded that synthetic and endogenous SL RNAs were used with comparable efficiency for *trans*-splicing.

Synthetic SL RNA becomes Sm precipitable and acquires a trimethylguanosine cap structure upon incubation in A.lumbricoides embryo extract

As noted in the Introduction, nematode SL RNAs are remarkably similar to vertebrate U snRNAs both in cap structure and in their ability to bind proteins with Sm antigenic determinants. A number of experiments were designed to ask whether either or both of these snRNA-like characteristics were important determinants in the use of synthetic SL RNA in *trans*-splicing.

To assess the ability of synthetic SL RNA to associate with Sm antigens immunoprecipitation experiments were performed using a human Sm antiserum. Upon incubation with extract, ~70% of synthetic SL RNA became Sm precipitable whereas no precipitation was observed with U1 or control antisera (Figure 3, panel A). Northern blot analysis indicated that a similar fraction of endogenous SL RNA was Sm precipitable (Figure 3, panel B) and, as with synthetic SL RNA, the SL RNA present in extract was not recognized by the U1 antiserum. The human U1 antiserum, however, did precipitate a substantive fraction of the *A.lumbricoides* U1 snRNP (data not shown). Despite increasing antibody to extract ratios we have not been able to precipitate quantitatively either endogenous or synthetic SL RNA.

Mattaj (1986) has previously shown that cap trimethylation of U2 snRNA in injected *Xenopus* oocytes requires a functional Sm binding site and we have shown (Maroney *et al.*, 1990) that the majority of SL RNA synthesized by *in vitro* transcription in *A.lumbricoides* extract receives a trimethylguanosine cap structure. Since synthetic SL RNA efficiently associated with Sm antigens it seemed possible that its cap structure could also become trimethylated. To assess the ability of the extract to trimethylate the cap structure of synthetic SL RNA, labeled SL RNA containing an unmethylated cap (GpppG) was incubated in extract for varying lengths of time and modification of the cap structure was monitored by two dimensional TLC. After 10 min of incubation, the unmodified cap was completely converted to an approximately equal mixture of monomethylated and trimethylated cap structures (Figure 4B spots b and c). Upon further incubation, the monomethyl caps were gradually converted to the fully methylated form. By 60 min, >90% of the synthetic SL RNA possessed a trimethylguanosine cap structure (Figure 4E).

Mutations within the SL RNA Sm binding site prevent Sm precipitation and cap trimethylation

Since we had shown that most of the synthetic SL RNA was assembled into an SL snRNP, we wished to determine if such assembly was necessary for the efficient utilization of synthetic SL RNA for *trans*-splicing. To disrupt the *A.lumbricoides* Sm binding site (AAUUUUGG) we used site directed mutagenesis to create two mutations; G78, in which the second U of the Sm binding site was changed to G and G79 in which the third uridine residue was changed to G. Labeled RNAs corresponding to each mutant were synthesized and incubated with embryo extract. In each case, these altered RNAs were not precipitable with a Sm antiserum (Figure 5A). Analysis of the cap structures present on the altered RNAs indicated that failure to become associated with Sm antigens was correlated with an absence of cap trimethylation (Figure 5, panel B).

SL RNAs containing non-functional Sm binding sites are not used efficiently in trans-splicing

The relative ability of the mutant and unaltered SL RNAs to be used for *trans*-splicing was compared in reactions containing equal c.p.m. (and mass) of each donor RNA. As shown in Figure 6 the SL RNAs with non-functional Sm binding sites were not used as *trans*-splice donors. It seemed possible that the observed reduction in *trans*-splicing could be due to the absence of a trimethylguanosine cap structure. To assess the importance of cap trimethylation in *trans*-splicing, processing reactions were carried out with unaltered SL RNA in the presence of the methylation inhibitor S-adenosyl homocysteine (SAH). Under these conditions cap structures remained either completely unmethylated or received a monomethyl cap (see Figure 5, panel B SAH).

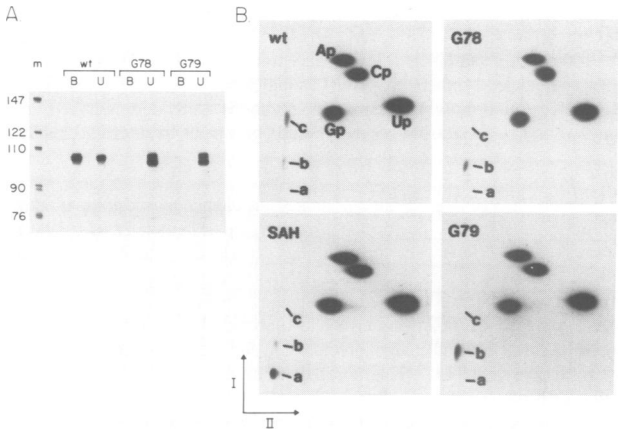

Fig. 5. SL RNA with altered Sm binding sites are not precipitable with a Sm antiserum and do not become cap trimethylated. **A.** Sm precipitation of wild-type SL RNA and SL RNAs with altered Sm binding sites as indicated. RNAs were incubated with embryo extract and precipitated with Sm antiserum as described in the legend to Figure 3. Bound and unbound RNAs as indicated were analyzed as described in Materials and methods. Lane M, restriction fragments of known length. **B.** Cap structures of wild-type and mutant SL RNAs following incubation in embryo extracts. Labeled SL RNAs as indicated were digested with RNase T2 and analyzed by two dimensional TLC after incubation with embryo extract as described in the legend to Figure 4. Panel designated SAH is a chromatogram of wild-type SL RNA which was incubated for 60 min in reactions containing 0.2 mM S-adenosyl homocysteine instead of 0.2 mM SAM. Spot a is GpppGp. Spot b is ^{m7}GpppGmGp and spot c is ^{m2,2,7}GpppGmGp. Thin layer chromatography was as described in the legend to Figure 4.

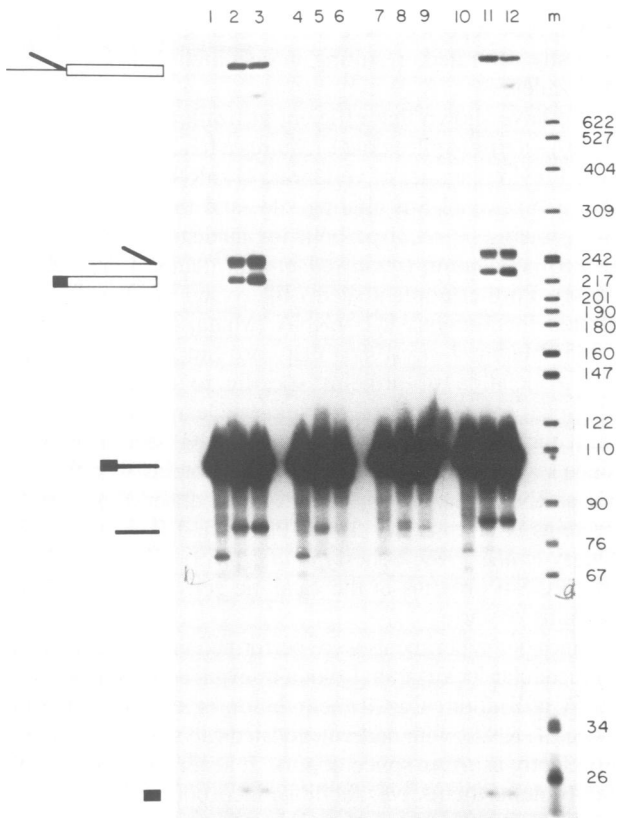

Fig. 6. SL RNAs with altered Sm binding sites are used inefficiently in *trans*-splicing. **A.** Labeled SL RNAs corresponding to wild-type SL RNA (lanes 1–3 and 10–12), G78 (lanes 4–6) or G79 (lanes 7–9) were incubated with embryo extract for 0 min (lanes 1, 4, 7 and 10), 30 min (lanes 2, 5, 8 and 11) or 60 min (lanes 3, 6, 9 and 12). Reactions in lanes 10–12 contained wild-type SL RNA and 0.2 mM S-adenosyl homocysteine. Incubations were assembled and analyzed as described in Materials and methods. Lane M, labeled restriction fragments of known size. The identities of reaction products and intermediates are indicated schematically with reference to Figure 1A. On this gel, RNAs indicated as debranched 'intron' and free exon 1 have been tentatively identified on the basis of mobility and sensitivity to digestion with RNAse H using diagnostic oligodeoxynucleotides (data not shown).

As shown in Figure 6, (lanes 11 and 12) inclusion of SAH had no effect on the efficiency of *trans*-splicing. We concluded from these experiments that the efficient utilization of SL RNA as a *trans*-splice donor required assembly of the SL RNA into an snRNP containing proteins with Sm antigenic determinants, but did not require cap trimethylation.

Discussion

We have shown that synthetic *A. lumbricoides* SL RNA serves as a donor for *trans*-splicing only if it is reconstituted into an RNP containing proteins with Sm antigenic determinants. These results provide direct support for the notion that SL RNAs represent a novel class of snRNP, in which an exon has been fused to an snRNA-like sequence.

We have previously characterized a *trans*-splicing reaction between SL RNA present in the *A. lumbricoides* extract and the same synthetic pre-mRNA used for the studies reported here (Hannon *et al.*, 1990). With endogenous SL RNA, adenosine residues 18 and 19 nucleotides upstream from the *trans*-splice acceptor site were used equally well as branch points (Hannon *et al.*, 1990). The same branch points are used when synthetic SL RNA is the splice donor. Using labeled acceptor molecules, we could not directly demonstrate the origin of the phosphate in the 2′5′ linkage within the branch. It is now clear that this phosphate is derived from the 3′5′ phosphodiester bond between the 22 nt SL sequence and the first guanosine (G23) of the 'intron' portion of the SL RNA. This result, in conjunction with our previous analyses, indicates that *trans*-splicing as catalyzed by the *A. lumbricoides* extract is mechanistically indistinguishable from *cis*-splicing, at least with respect to substrate RNAs. While the *trans*-splicing reactions using either labeled acceptor pre-mRNA or labeled donor SL RNA are qualitatively identical, there are some apparent differences in accumulation of intermediates and intron-related end products. For example, using labeled acceptor RNA, we did not detect significant accumulation of the Y-intron intermediate and observed a linear accumulation of 'debranched' intron (Hannon *et al.*, 1990). Using labeled donor molecules, we observed significant accumulation of the Y-intron and relatively little accumulation of 'debranched' intron sequence (see Figure 6). There are at least three possible explanations for this apparent discrepancy. First, in [^{32}P]G-labeled SL RNA, the SL sequence contains four labeled guanosines while the 'intron' contains 21. Thus, while the relative accumulations of Y-intron and *trans*-spliced product RNAs seem equal (see Figures 1 and 6) by autoradiography, the actual molar accumulation of product RNA is 5-fold greater than that of Y-intron. Second, it seems possible that the 'intron' portion of the SL RNA could be subject to rapid degradation in the extract following debranching. Finally, to obtain sufficient sensitivity for these experiments we routinely used 50–100 ng of unlabeled acceptor RNA. This contrasts to the 3–5 ng of labeled acceptor RNA used in

previous experiments. It seems possible that under the conditions reported here the 'debranching' enzyme is partially saturated, resulting in accumulation of branched molecules.

We have shown that efficient utilization of synthetic SL RNA as a *trans*-splice donor requires a functional Sm binding site. Previous studies of vertebrate U snRNAs have revealed that a functional Sm binding site, and thus association of Sm proteins, is required both for nuclear localization and cap trimethylation (reviewed in Lührmann, 1988; Mattaj, 1988). Furthermore, it seems likely and has been demonstrated for *Xenopus* U2 snRNA (Mattaj and De Robertis, 1985; Mattaj *et al.*, 1986) that core Sm proteins facilitate the association of proteins specific to each U snRNA. However, binding of core proteins is clearly not an absolute prerequisite for association of some snRNP specific proteins, since several groups have shown that proteins specific to the U1 snRNP can bind U1 snRNA in the absence of core Sm proteins (Hamm *et al.*, 1988; Patton and Pederson, 1988; Query *et al.*, 1989; Scherly *et al.*, 1989). The role, if any, of core Sm proteins in the catalytic activity of U snRNPs remains obscure.

Since our experiments were performed in cell free extracts, intracellular localization is not a relevant determinant for SL RNA utilization. Cap trimethylation is also not important since SL RNAs with undermethylated cap structures are used efficiently as *trans*-splice donors. It seems possible that the Sm proteins themselves facilitate incorporation of SL RNA into spliceosomes or that they function indirectly by facilitating the binding of a protein(s) to the SL snRNP which is necessary for *trans*-spliceosome assembly. Support for either alternative awaits the characterization and functional analysis of the individual proteins present in SL snRNPs.

Materials and methods

Preparation of synthetic SL RNAs

To construct a phage T3 transcription template encoding the *A.lumbricoides* SL RNA, the T3 pomoter sequence TCGGAATTAACCCTCACTAAA (Stratagene) was fused to the coding sequence of the SL RNA such that transcription by T3 RNA polymerase would initiate synthesis with the first guanosine residue of the 22 nt SL sequence. This was accomplished by using the polymerase chain reaction (PCR) technique with two oligodeoxynucleotides, one containing the T3 promoter and the first 15 nt of the SL sequence, the other complementary to the 3′ end of the SL sequence. For some experiments, (Figures 1, 2 and 4) the PCR product was excised from preparative gels and transcribed directly with T3 RNA polymerase. In this case, transcripts were 109 nt in length and terminated with a uridine residue 1 base downstream of the authentic 3′ terminus of the *A.lumbricoides* SL RNA. For other experiments, (Figures 3, 5 and 6) the PCR fragment was first subcloned into the *Sma*I site of pSP65 (Promega) and then recloned into PBS M13$^+$ (Stratagene) lacking the T3 promoter. This plasmid was transformed into dut$^-$, ung$^-$ *Escherichia coli* (CJ236) (Kunkel *et al.*, 1987). And the plus strand of this plasmid was rescued and prepared as described (Vieira and Messing, 1987) before being subjected to site directed mutagenesis as described (Kunkel *et al.*, 1987). Initial mutagenesis generated a *Sma*I site at the 3′ terminus of the SL RNA sequence. When cleaved with *Sma*I, run off transcription of this plasmid generated a 107 nt transcript which terminated in CCC. Authentic SL RNA terminates in CACU (Maroney *et al.*, 1990). To create altered Sm binding sites the '*Sma*I' plasmid was subjected to further mutagenesis as described. In all cases transcripts were initiated with the unmethylated cap structure GpppG and synthesized in the presence of [α-^{32}P]GTP. The cap-labeled SL RNA used in Figure 1 was prepared using vaccinia guanyl transferase (BRL) and [α-^{32}P]GTP according to the supplier's instructions. This RNA contained a 34 nt 3′ extension consisting of pSP65 polylinker sequence.

Preparation of extracts and trans-splicing reactions

A.lumbricoides whole cell extracts were prepared exactly as described (Hannon *et al.*, 1990). *In vitro trans*-splicing reactions in 12.5 μl contained 7.5 μl extract, 2 mM ATP, 20 mM creatine phosphate, 2 mM DTT, 4.2 mM $MgCl_2$, 60 mM KCl, 0.2 mM SAM, 12 mM Tris−HCl, pH 7.9, 20−100 ng of an unlabeled 345 nt acceptor pre-mRNA (Hannon *et al.*, 1990) and labeled synthetic SL RNAs as indicated in the text and individual figure legends. Following incubation, reactions were deproteinized and analyzed on 6% denaturing polyacrylamide gels prior to autoradiography.

RNA analysis

Two dimensional RNase T1 fingerprint analysis of splicing substrate, product and intermediates was performed as described (Hannon *et al.*, 1989). Debranching reactions were carried out as described by Ruskin and Green (1985). For cap analyses, labeled SL RNAs as indicated in the text were incubated with whole cell extract under splicing conditions for times indicated in the figure legends. Following incubation, labeled SL RNAs were recovered by phenol extraction and ethanol precipitation and the cap structure identified by two dimensional TLC (Nishimura, 1972) exactly as described (Maroney *et al.*, 1990).

Immunoprecipitation experiments were carried out essentially as described by Bruzik *et al.* (1988). Labeled SL RNAs (5 ng) were incubated with extract under splicing conditions for 60 min before being diluted to 200 μl with 150 mM NaCl, 0.05% NP40, 50 mM Tris pH 7.5 (NET 2) and 1 μl was added to 250 μl NET 2 which contained 2.5 mg protein A sepharose−antibody complex. Following a 60 min incubation at 4°C, the sepharose beads were washed five times with NET2 and bound RNAs were recovered by phenol extraction and ethanol precipitation. Bound and unbound RNAs were analyzed on 6% polyacrylamide−8 M urea gels. Analysis of endogenous RNA was identical, except that bound and unbound RNAs were identified on RNA blots using an oligodeoxynucleotide complementary to the SL RNA as a probe.

Acknowledgements

We are grateful to Dr Joan Steitz (Yale University) for providing antisera and to Jim Bruzik (Yale University) for advice on immunoprecipitation. Supported by grants from the John D. and Catherine T.MacArthur Foundation and the PHS (GM31528).

References

Anderson,G.J., Bach,M., Luhrman,R. and Beggs,J.D. (1989) *Nature*, **342**, 819−821.
Bindereif,A., Wolff,T. and Green,M.R. (1990) *EMBO J.*, **9**, 251−255.
Blumenthal,T. and Thomas,J. (1988) *Trends Genet.*, **4**, 305−308.
Bruzik,J.P., van Doren,K., Hirsh,D. and Steitz,J.A. (1988) *Nature*, **335**, 559−562.
Chabot,B., Black,D.L., Lemaster,D.M. and Steitz,J.A. (1985) *Science*, **230**, 1344−1349.
Fabrizio,P., McPheeters,D.S. and Abelson,J. (1989) *Genes Dev.*, **3**, 2137−2150.
Gerke,V. and Steitz,J.A. (1986) *Cell*, **42**, 345−353.
Hamm,J., van Santen,V.L., Spritz,R.A. and Mattaj,I.W. (1988) *Mol. Cell. Biol.*, **8**, 4787−4791.
Hamm,J., Dathan,N.A. and Mattaj,I.W. (1989) *Cell*, **59**, 159−169.
Hannon,G.J., Maroney,P.A., Branch,A., Benefeld,B.J., Robertson,H.D. and Nilsen,T.W. (1989) *Mol. Cell. Biol.*, **9**, 4422−4431.
Hannon,G.J., Maroney,P.A., Denker,J.A. and Nilsen,T.W. (1990) *Cell*, **61**, 1247−1255.
Heinrichs,V., Vach,M., Winkelmann,G. and Lührmann,R. (1990) *Science*, **247**, 69−72.
Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) *Methods Enzymol.*, **154**, 367−382.
Lührmann,R. (1988) In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles.* Springer-Verlag, Heidelberg, pp. 71−99.
Maroney,P.A., Hannon,G.J. and Nilsen,T.W. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 709−713.
Mattaj,I.W. (1986) *Cell*, **46**, 905−911.
Mattaj,I. (1988) In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles.* Springer-Verlag, Heidelberg, pp. 100−114.
Mattaj,I.W. and DeRobertis,E. (1985) *Cell*, **40**, 111−118.
Mattaj,I.W., Habets,W.J. and van Venrooij,W.J. (1986) *EMBO J.*, **5**, 997−1002.
McPheeters,D.S., Fabrizio,P. and Abelson,J. (1989) *Genes Dev.*, **3**, 2124−2136.
Mount,S.M., Pettersson,I., Hinterberger,J., Karmas,A. and Steitz,J.A. (1983) *Cell*, **33**, 509−518.

Nishimura,S. (1972) *Prog. Nucleic Acid Res. Mol. Biol.*, **12**, 48–85.
Nilsen,T.W. (1989) *Exp. Parasitol.*, **69**, 413–416.
Nilsen,T.W., Shambaugh,J., Denker,J.A., Chubb,G., Faser,C., Putnam,L. and Bennett,K. (1989) *Mol. Cell. Biol.*, **9**, 3543–3547.
Pan,Z.Q. and Prives,C. (1988) *Science*, **241**, 1328–1331.
Patton,J.R. and Pederson,T. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 747–751.
Query,C.C., Bentley,R.C. and Keene,J.D. (1989) *Mol. Cell. Biol.*, **9**, 4872–4881.
Ruskin,B. and Green,M. (1985) *Science*, **229**, 135–140.
Scherly,D., Boelens,W., van Venrooij,W.J., Dathan,N.A., Hamm,J. and Mattaj,I.W. (1989) *EMBO J.*, **13**, 4163–4170.
Sharp,P.A. (1987) *Cell*, **50**, 147–148.
Tazi,J., Alibert,C., Temsamani,J., Reveillaud,I., Cathala,G., Brunel,C. and Jeanteur,P. (1986) *Cell*, **47**, 755–766.
Thomas,J.D., Conrad,R.C. and Blumenthal,T. (1988) *Cell*, **54**, 533–539.
Van Doren,K. and Hirsh,D. (1988) *Nature*, **335**, 556–558.
Vieira,J. and Messing,J. (1987) *Methods Enzymol.*, **153**, 3–11.
You,C.Y. and Weiner,A.M. (1989) *Genes Dev.*, **3**, 697–707.

Received on July 17, 1990