The EMBO Journal vol.10 no.12 pp.3869–3875, 1991

Intramolecular base pairing between the nematode spliced leader and its 5′ splice site is not essential for *trans*-splicing *in vitro*

Patricia A.Maroney, Gregory J.Hannon, Joseph D.Shambaugh and Timothy W.Nilsen

Department of Molecular Biology and Microbiology, Case Western Reserve University School of Medicine, Cleveland, OH 44106-4690, USA

Communicated by J.A.Steitz

The spliced leader RNAs of both trypanosomes and nematodes can form similar secondary structures where the *trans*-splice donor site is involved in intramolecular base pairing with the spliced leader sequence. It has been proposed that this base pairing could serve to activate autonomously the SL RNA splice donor site. Here, we have examined exon requirements for *trans*-splicing in a nematode cell free system. Complete disruption of secondary structure interactions at and around the *trans*-splice donor site did not affect the ability of the SL RNA to function in *trans*-splicing. In addition, the highly conserved 22 nt sequence could be productively replaced by artificial exons ranging in size from 2 to 246 nucleotides. These results reinforce the view that the 'intron' portion of the SL RNA functions as an independent Sm snRNP whose role is to deliver exon sequences to the *trans*-spliceosome.

Key words: *Ascaris lumbricoides*/nematode/RNA processing/Sm snRNP/*trans*-splicing

Introduction

Maturation of a fraction of mRNAs in nematodes and all mRNAs in trypanosomatid protozoans involves the acquisition of a 5′ terminal spliced leader exon (SL) from an SL RNA via a *trans*-splicing reaction (see Agabian, 1990; Nilsen, 1989; Blumenthal and Thomas, 1988; Borst, 1986 for review). *Trans*-splicing *in vitro* proceeds through a two-step reaction which generates intermediates and products analogous to those found in snRNP mediated *cis*-splicing, indicating that these two processes are catalytically similar, if not identical (Hannon *et al.*, 1990b; Maroney *et al.*, 1990).

In both nematodes and trypanosomes, SL RNAs exist as RNP particles with striking similarity to the U snRNPs involved in catalysis of *cis*-splicing. This is particularly clear in nematodes where SL RNAs possess both the cap structure and Sm binding sequences characteristic of snRNAs. The Sm binding site is required since it has been demonstrated that assembly of the nematode SL RNA into an Sm snRNP is a prerequisite for its participation in *trans*-splicing *in vitro* (Maroney *et al.*, 1990). The resemblance of SL RNAs to U snRNAs has led to the proposal that SL RNAs represent a unique class of snRNP where 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 Hirsh, 1988). Unlike the snRNPs involved in *cis*-splicing, the SL RNA is consumed during the *trans*-splicing reaction.

An important unresolved question with respect to the SL snRNP is whether both domains (exon and snRNA-like sequence) have functional significance in *trans*-splicing. Despite a lack of sequence similarity, the SL RNAs of nematodes and trypanosomes are predicted to form nearly identical secondary structures consisting of three stem–loops (Bruzik *et al.*, 1988; Nilsen *et al.*, 1989). In these structures, stem–loop I is separated from the second stem by a short single stranded region. Stems II and III are separated by another single stranded region consisting of the Sm binding site and adjacent nucleotides. An interesting feature of this structure, as noted by Bruzik *et al.* (1988) is that the 5′ splice sites of the SL RNAs are involved in base pairing within stem I. By analogy to the intermolecular base pairing of U1 snRNP to *cis* 5′ splice sites, it was proposed that this intramolecular base pairing could autonomously activate the SL RNA 5′ splice site and render *trans*-splicing independent of the U1 snRNP (Bruzik *et al.*, 1988). Recently, this notion has received strong experimental support from studies of the processing of chimeric transcripts containing SL RNA splice donor sites linked *in cis* to an adenoviral splice acceptor site and 3′ exon (Bruzik and Steitz, 1990). Such transcripts, containing either trypanosomatid or nematode SL RNA sequences, were efficiently spliced in HeLa cell extracts. By several criteria this splicing did not require the 5′ end of U1 snRNA (Bruzik and Steitz, 1990). These observations suggest that the exon sequence and consequent structure may be important for appropriate utilization of SL RNAs in *trans*-splicing.

Here, we have examined the role of intramolecular base pairing of the SL sequence and its 5′ splice site in nematode *trans*-splicing in a homologous whole cell extract. Surprisingly, in this system, the predicted base pairing interactions within stem I of the nematode SL RNA are not required for SL RNA participation in *trans*-splicing. Further, the snRNA-like domain of the SL RNA can productively deliver a wide variety of heterologous exons via *trans*-splicing to an appropriate acceptor. These results indicate that the 'intron' of the SL RNA functions as a quasi-independent Sm snRNP in *trans*-splicing whose role is to deliver exons to the *trans*-spliceosome.

Results

Mutational analysis of 'stem I' of the A.lumbricoides SL RNA

We have previously shown that cell free extracts derived from synchronous embryos of the parasitic nematode *Ascaris lumbricoides* catalyze accurate *trans*-splicing between endogenous or synthetic *A.lumbricoides* SL RNA and a synthetic pre-mRNA containing a *trans*-splice acceptor site (Hannon *et al.*, 1990b; Maroney *et al.*, 1990). Here, *trans*-

© Oxford University Press

splicing of the same substrates (shown schematically in Figure 1) was assayed using labelled synthetic SL RNA.

Since synthetic SL RNA participates in *trans*-splicing, it was possible to assess the effect of sequence alterations on SL RNA function. It is important to note that several factors must be taken into account in the interpretation of SL RNA mutant phenotypes. First, in order to function in *trans*-splicing, the synthetic SL RNA must assemble into an Sm snRNP (Maroney *et al.*, 1990). In reactions described here, the levels of added SL RNAs were well below those which saturate the assembly capacity of the extract. Second, all assays with synthetic SL RNAs were performed in the presence of equal or greater amounts of endogenous SL RNP. Thus, to participate in *trans*-splicing, synthetic SL RNPs must compete with this endogenous pool of SL RNP for components of the *trans*-splicing machinery. Finally, under our assay conditions, SL RNPs are not limiting for *trans*-splicing since only a small fraction of the total SL RNA pool is consumed during a 60 min reaction (Maroney *et al.*, 1990).

For initial mutagenic analysis of the SL RNA, we focused on stem–loop I since, as indicated above, it has been suggested that intramolecular base pairing within this stem could contribute to 5′ splice site activation in *trans*-splicing. Computer predictions of secondary structure (Zuker and Stiegler, 1981, as modified by Turner *et al.*, 1987) suggest that stem I of the wild type *A.lumbricoides* SL RNA could contain nine base pairs spanning the splice donor site. In this potential structure, there is one bulged adenosine residue (A_{21}) immediately preceding G_{22}, the last base of the SL sequence. Directed mutagenesis, designed to destabilize this potential structure, was used to change bases in the 5′ side of the stem (Figure 2A, G_8G_9 and $G_8G_9A_{11}G_{12}$). In these altered SL RNAs, the splice donor site and immediately surrounding bases were left intact. As shown in Figure 2B, SL RNAs lacking either two or four of the predicted base pairs in stem I (including those base pairs spanning the splice site) functioned in *trans*-splicing. Computer analysis predicted that the quadruple mutation would completely destabilize base pairing within stem I. In a parallel set of experiments, bases in the 3′ side of stem I lying immediately downstream of the conserved GU of the splice donor site were altered (Figure 2A, $C_{25}A_{28}$ and $C_{25}C_{26}A_{27}A_{28}$). These changes were designed both to destabilize stem I and to minimize potential base pairing with the 5′ end of U1 snRNA. As with the mutant SL RNAs altered in the 5′ side of stem I, SL RNAs containing either two or four altered bases in the 3′ side of stem I were used as *trans*-splice donors although these RNAs were used less efficiently than the wild type SL RNA. Densitometry indicated that the $C_{25}A_{28}$ and $C_{25}C_{26}A_{27}A_{28}$ mutant RNAs were used with ~90% and ~30% of the efficiency of wild type. Since assembly of the SL RNA into an Sm snRNP is a prerequisite for use in *trans*-splicing (see above), inefficient use of the $C_{25}C_{26}A_{27}A_{28}$ RNA could reflect either reduced assembly or a defect directly related to splicing. To distinguish between these possibilities, we measured assembly both by assaying cap trimethylation of mutant SL RNAs (Maroney *et al.*, 1990) and by mobility shift on native gels. These analyses revealed that the $C_{25}C_{26}A_{27}A_{28}$ mutant assembled with ~30% of the efficiency of the normal SL RNA, and the $C_{25}A_{28}$ mutant assembled nearly as well as the wild type SL RNA (Figure 2C and data not shown). Both of these values are in good agreement with the level to which these mutants participate in *trans*-splicing. We do not know whether the reduction in assembly of the $C_{25}C_{26}A_{27}A_{28}$ mutant results from a direct effect of the altered bases or from inappropriate folding of the mutant RNA.

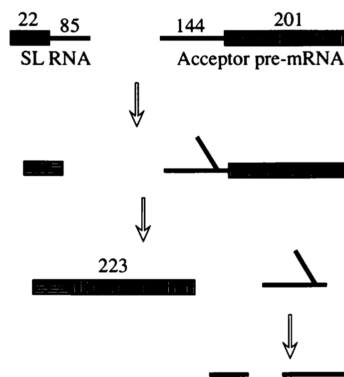

Fig. 1. Schematic representation of *in vitro trans*-splicing. The reaction pathway of the *in vitro trans*-splicing reaction is shown schematically. The numbers indicate sizes in nucleotides.

In addition to base substitutions, we constructed SL RNAs in which blocks of the SL sequence were deleted (Figure 2A, Δ3–12 and Δ11–21). Removal of bases 3–12 or 11–21 of the SL sequence did not reduce the efficiency of *trans*-splicing (Figure 2B, lanes Δ3–12 and Δ11–21). In these reactions, truncated exons of 12 and 11 nucleotides respectively were transferred to the acceptor molecule. This result further strengthened the notion that secondary structure involving the 5′ splice site was not relevant to *trans*-splicing, since neither deletion construct was predicted to form any base pairing interactions at or around the 5′ splice site. Use of SL RNAs with deleted exons also demonstrated that an intact 22 nt SL sequence was not a prerequisite for *trans*-splicing.

Mutational analysis of SL RNA sequence between the splice site and stem II

The foregoing experiments suggested that intramolecular base pairing within stem I was not an important determinant for use of the SL RNA in *trans*-splicing even under competition conditions with endogenous SL RNA. It thus became of interest to determine which sequence features of the SL RNA (in addition to the Sm binding site) were required for *trans*-splicing. In the following experiments we focused on the 11 base region separating stem II and the 5′ splice site. This region was chosen for analysis for three reasons. First, alteration of this region of a trypanosomatid SL RNA sequence restored U1 dependence to certain *cis*-constructs in the HeLa cell-free system (Bruzik and Steitz, 1990). Second, appropriate functioning of U1 snRNA has been shown to depend upon precise spacing between its conserved 5′ end and stem–loop I (You and Weiner, 1989). Thus, by analogy, it seemed possible that correct positioning of the SL RNA 5′ splice site with respect to the rest of the molecule could be critical. Third, we noted a phylogenetically conserved potential base pairing interaction between this region of the *Ascaris* and *Caenorhabditis elegans* SL RNAs and their respective U2 snRNAs. Directed mutagenesis was used to introduce two different five base substitutions just 5′ to stem II (Figure 3A, $G_{29}A_{30}T_{31}A_{32}C_{33}$

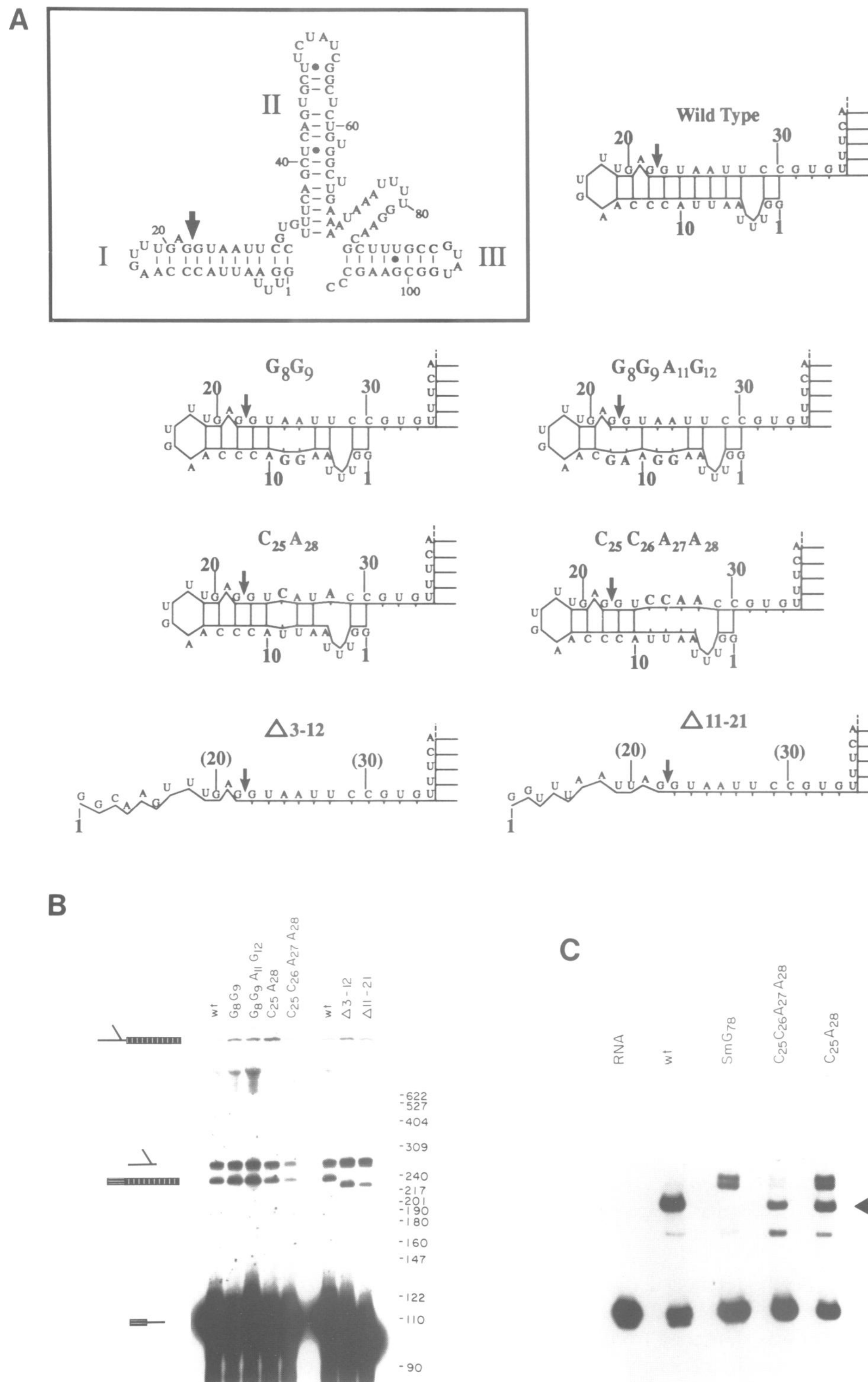

Fig. 2. *Trans*-splicing of SL RNAs altered in stem I. (**A**) A computer generated secondary structure of the *Ascaris* SL RNA is shown schematically as is the sequence and potential structure of the unaltered stem I (wild type). Mutations were introduced into SL RNA stem I by site directed mutagenesis as described in Materials and methods. In each mutant construct, altered bases are shown in bold type and the splice site is indicated by an arrow. The structures shown for each of the stem I mutant SL RNAs are intended to illustrate the mutations and do not necessarily represent predicted structures. Stem I sequences in constructs $G_8G_9A_{11}G_{12}$, $C_{25}C_{26}A_{27}A_{28}$, $\Delta 3-12$ and $\Delta 11-21$ are predicted to be completely unstructured. (**B**) Unaltered (wt) or the mutant SL RNAs (shown schematically in panel A) were assayed for the ability to serve as *trans*-splice donors. Intermediates and products of the reaction are depicted as in Figure 1. Indicated sizes correspond to the positions of labelled restriction fragments electrophoresed in parallel lanes. (**C**) The assembly of wild type or mutant SL RNAs (as indicated) was assayed by native gel electrophoresis. The $C_{25}C_{26}A_{27}A_{28}$ and $C_{25}A_{28}$ mutants are depicted in panel A. In the SmG_{78} mutant, the second uridine of the SL RNA Sm binding site has been changed to guanosine. This mutation has previously been shown to abolish the assembly of this mutant into an Sm snRNP (Maroney *et al.*, 1990). For this analysis, *trans*-splicing reactions were incubated for only 20 min and then treated with high salt (see Materials and methods) before native gel analysis. Under these conditions, the wild type SL RNA does not completely shift into the specific complex (indicated by the arrow). Quantitatively similar results were obtained under conditions where the wild type SL RNA was completely assembled (data not shown).

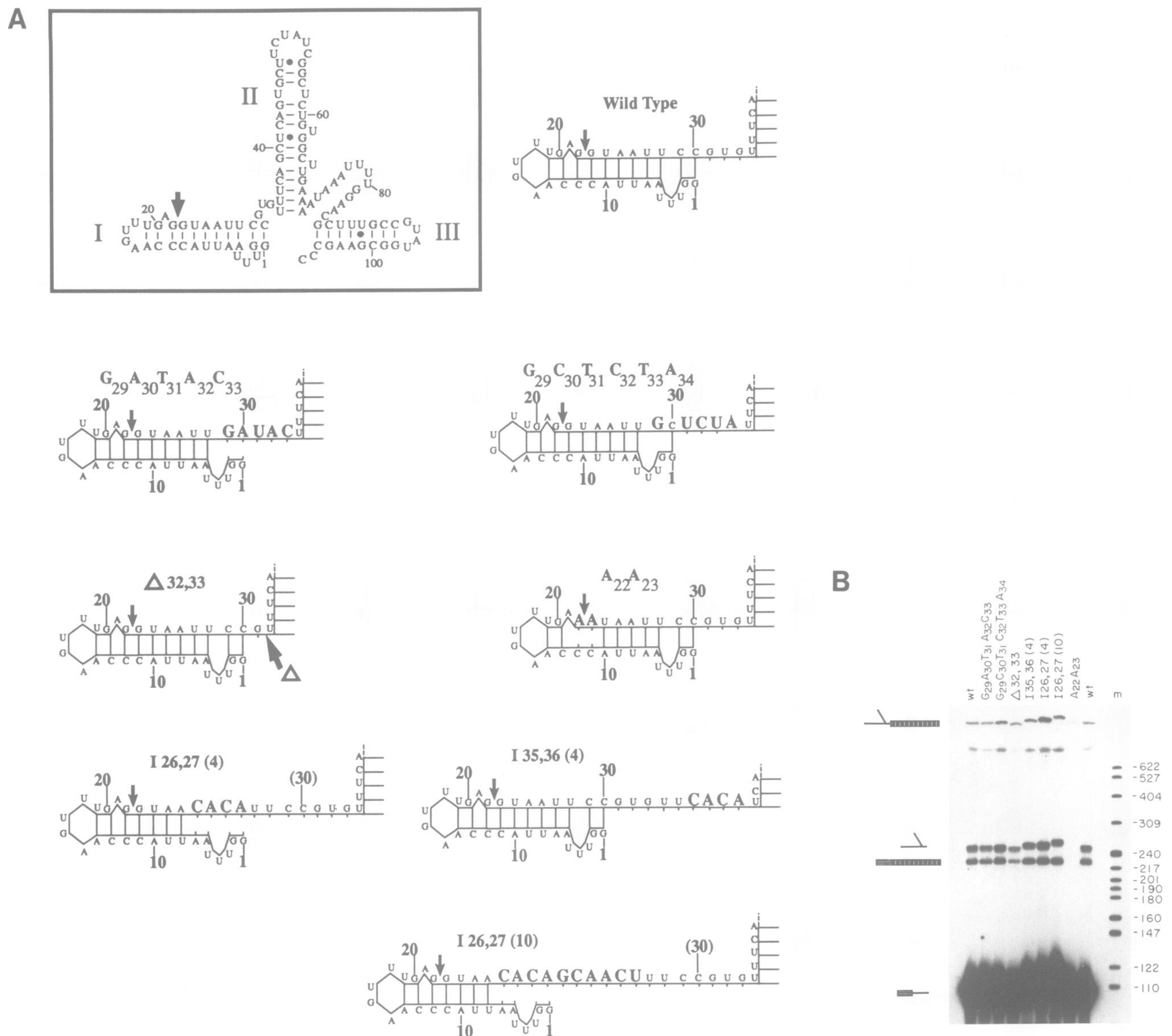

Fig. 3. Mutational analysis of the *Ascaris* SL RNA in the region between the splice donor site and stem II. (**A**) Schematic representation of alterations of the *Ascaris* SL RNA. Alterations introduced by site directed mutagenesis are indicated in bold type and expressed as DNA. As in Figure 2, indicated structures are for illustration purposes only. In all cases, numbering is with respect to the unaltered SL RNA. (**B**) Each altered SL RNA (as indicated) was assayed for the ability to participate in *trans*-splicing as described in Materials and methods. Schematic designations of splicing intermediates and products are shown as in Figure 1. The positions of labelled restriction fragments electrophoresed in parallel are shown (lane m).

and $G_{29}C_{30}T_{31}C_{32}T_{33}A_{34}$). Neither sequence substitution caused a significant change in *trans*-splicing efficiency (Figure 3B). Next, the region was shortened from 11 bases to 9 bases by removing two bases immediately 5′ to stem II (Figure 3A, Δ32,33). This mutation caused a modest reduction (30%) in utilization of the SL RNA, but as with the $C_{25}C_{26}A_{27}A_{28}$ mutation in stem I, the reduction in utilization was directly correlated with inefficient assembly of the shortened SL RNA (data not shown). We attempted to decrease further the distance between the 5′ splice site and stem II by deleting four bases (bases 29–32). However, this RNA failed to assemble into an Sm snRNP and consequently was not used in *trans*-splicing (data not shown). The 'hinge' region was considerably more tolerant to lengthening. The region was lengthened from 11 to 15 nucleotides by insertion of four bases at two separate sites downstream of the 5′ splice site [Figure 3A, I26,27(4) and I35,36(4)] or lengthened from 11 to 21 nucleotides by insertion of ten bases [I26,27(10)]. Each of these RNAs participated productively in *trans*-splicing (Figure 3B). Finally, the splice site guanosines were changed to adenosines (Figure 3A, $A_{22}A_{23}$). This SL RNA was not used as a *trans*-splice donor (Figure 3B) even though it was efficiently assembled into an Sm snRNP (data not shown). We concluded from these experiments that although an intact splice site was required, neither the primary sequence of the 'hinge' region nor the absolute position of the splice site relative to the rest of the SL RNA molecule affected splice site selection or utilization.

SL RNAs containing artificial exons of 2–246 nucleotides participate in trans-splicing

The experiments presented in Figure 2 indicated that an intact 22 nt SL sequence was not a prerequisite for *trans*-splicing. Since numerous studies have suggested that both exon

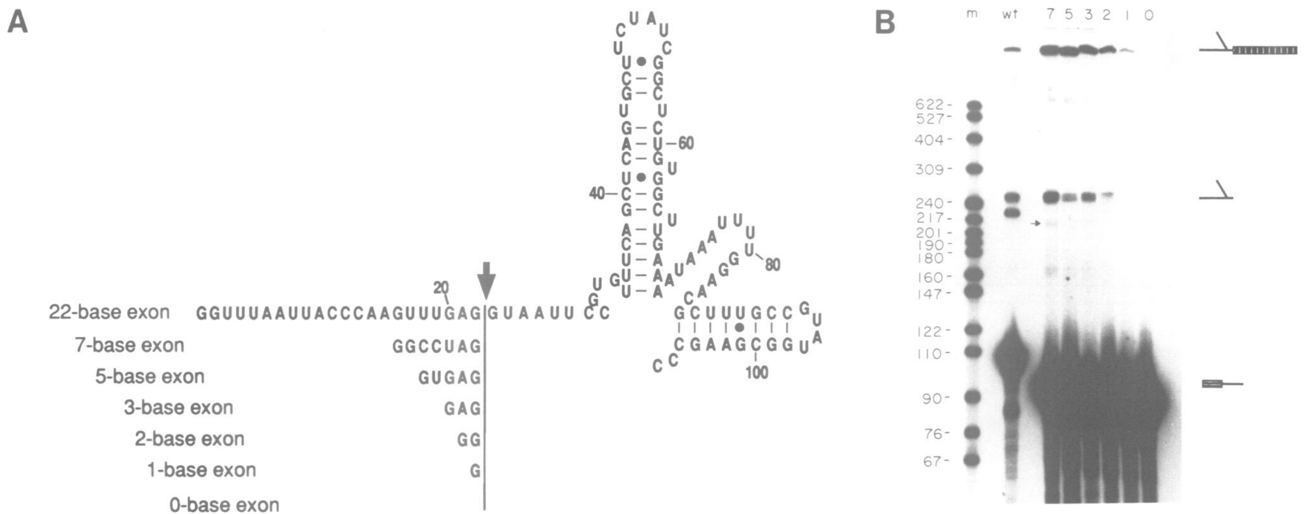

Fig. 4. The 22 nucleotide SL sequence is dispensable for *trans*-splicing. Transcription templates encoding SL RNAs containing 0, 1, 2, 3, 5 or 7 base exons were constructed by site directed mutagenesis. The sequence of the truncated constructs is shown in (**A**). Each of these RNAs as indicated was assayed for participation in *trans*-splicing (**B**). An unaltered SL RNA (wt) was analyzed in parallel. 1 μl of the wt reaction was electrophoresed whereas 5 μl of each of the other reactions were analyzed. Schematic designations of splicing intermediates are shown. The *trans*-spliced product produced upon incubation of the SL RNA containing a seven base exon is denoted by an arrow. Indicated sizes (M) show the positions of labelled restriction fragments.

sequence and exon length can be important in *cis*-splicing, it was of interest to ascertain the degree of exon flexibility that could be accommodated in *trans*-splicing. We first determined the minimal exon size required for *trans*-splicing by synthesizing SL RNAs containing artificial exons of 0, 1, 2, 3, 5 and 7 bases (Figure 4A). As with experiments shown above, each RNA was tested for its ability to participate in *trans*-splicing in the presence of equal or greater amounts of endogenous SL RNA. While the SL RNA lacking an exon (0, Figure 4B) did not participate in *trans*-splicing, all of the other SL RNAs with short artificial exons were used as *trans*-splice donors (Figure 4B). Densitometric scanning of Y intron RNAs generated in these reactions showed that SL RNAs containing 7, 5, 3, 2 and 1 base exons were used with 46, 10, 16, 6 and <1% the efficiency of the wild type SL RNA. The reduction in efficiency probably reflects effects of exon size rather than sequence content since SL RNAs with longer exons bearing no resemblance to the 22 nt SL sequence were used with efficiencies comparable to that of wild type SL RNA (see below). Somewhat surprisingly, with the exception of the SL RNA containing a seven base exon, we did not observe the expected *trans*-spliced products in reactions containing SL RNAs with truncated exons. Since both steps of splicing (*cis* and *trans*) are thought to occur by transesterification reactions, we considered it highly unlikely that release of the Y intron intermediate could be uncoupled from product formation. It seemed more likely that the failure to observe product was due to underlabelling in the 5′ end of the synthetic RNAs, in part because the artificial exons contain fewer labelled guanosines and additionally, although we do not understand the effect, by fingerprint analysis we routinely observe underlabelling in the 5′ terminus of RNAs synthesized by T3 RNA polymerase. As a more sensitive assay for product formation, 5′ end-labelled SL RNAs (see Materials and methods) containing truncated exons were used in *trans*-splicing reactions. Products of the expected size were observed for those RNAs containing exons of 2, 3, 5 and 7 nucleotides (data now shown). We have not observed a product with the RNA containing a one base exon, presumably because this RNA is used extremely inefficiently for *trans*-splicing (see Figure 4). The ability of SL RNAs bearing short exons to function in both steps of *trans*-splicing contrasts with previous determinations of minimal exon size for *cis*-splicing. Domdey and coworkers (Duchenne *et al.*, 1988) showed in a yeast system that constructs containing first exons of 1, 2, 3 or 5 nucleotides participated only in the first step of splicing and did not participate in the second step.

The use of SL RNAs with extremely short exons in *trans*-splicing demonstrated definitively that the 22 nt SL sequence was not required for this process and suggested the possibility that the 'intron' portion of the SL RNA might be capable of productively delivering heterologous exons to the *trans*-spliceosome. To test this possibility, a restriction endonuclease cleavage site was created just upstream of the SL RNA splice donor site by site directed mutagenesis and fragments of either pBR322 or pBS M13+ were cloned into this site to generate 'exons' ranging in size from 29 to 246 nucleotides (Figure 5B). The resultant constructs were transcribed *in vitro* and remarkably, all of these artificial exons were *trans*-spliced to the acceptor molecule (Figure 5B, lanes +). As expected for authentic *trans*-splicing, the Y intron + exon and Y intron intermediates observed in these reactions were identical in mobility regardless of exon size (Figure 5B). Also, function of these *trans*-splicing substrates was dependent upon assembly into an Sm snRNP since molecules with inactivated Sm binding sites failed to participate in *trans*-splicing (Figure 5B, lanes −).

Discussion

5′ splice site identification in nematode trans-splicing

Use of altered SL RNAs containing base substitutions and truncated exons demonstrates that the predicted base pairing interactions between the SL sequence and the splice donor site within stem I of the *A.lumbricoides* SL RNA are not important for *trans*-splicing *in vitro*. While these results would appear to indicate that intramolecular base pairing

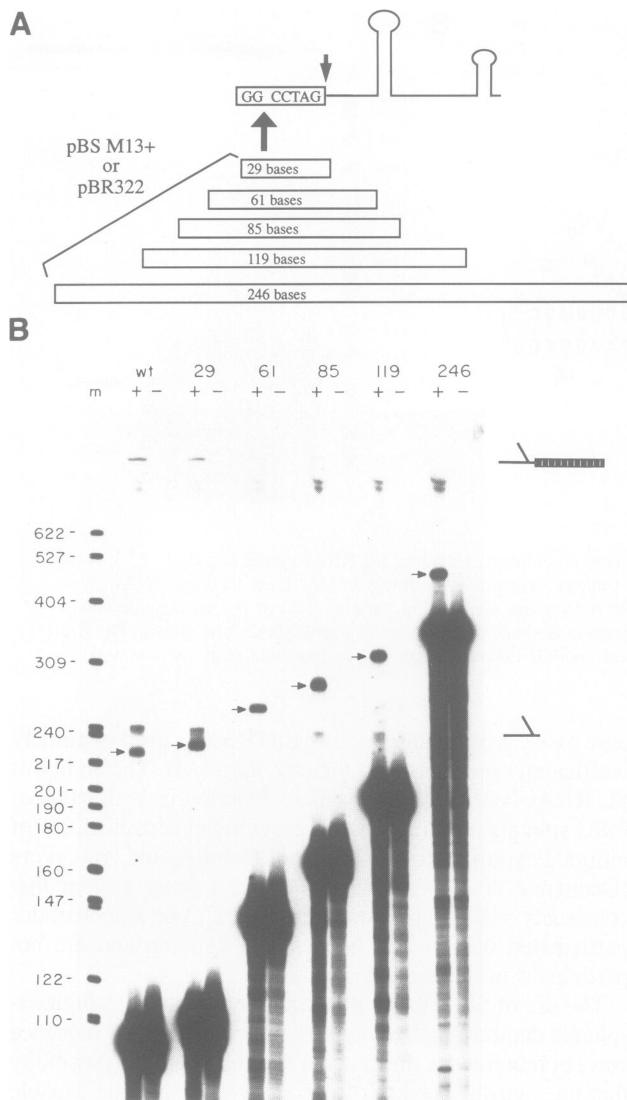

Fig. 5. The 'intron' of the SL RNA can productively deliver heterologous artificial exons via *trans*-splicing. The 22 nt SL sequence of the *Ascaris* SL RNA was replaced with a *Stu*I recognition sequence as described in Materials and methods. Random fragments of pBR322 or PBS M13+ were sub-cloned into this site to yield transcription templates with artificial exons of the indicated sizes (**A**). Equal c.p.m. and masses of unaltered SL RNA (wt) or SL RNAs with artificial exons (as indicated) were assayed for *trans*-splicing as described in Materials and methods (**B**). (Since equal c.p.m. were used, longer SL RNAs were present in proportionally lower molar concentrations.) Lanes designated + correspond to RNAs with intact Sm binding sites. In lanes designated − the second U of each Sm binding site sequence was changed to G to prevent assembly of the RNAs into Sm snRNPs (analogous to G_{78}, Maroney *et al.*, 1990). Schematic designations of splicing intermediates are shown. The different splicing products are indicated by arrows. The position of labelled restriction fragments electrophoresed in parallel are shown.

does not contribute to 5′ splice site activation in nematode *trans*-splicing, other interpretations are formally possible. It is conceivable that the computer-generated secondary structure of the SL RNA could be in error and that the splice donor site might actually pair with a different region of the SL RNA molecule. It is also possible that disruption or deletion of base pairs within stem I could promote an alternative secondary structure in which the 5′ splice site is again paired with a different region of the SL RNA. We consider the first possibility unlikely since any alternative intramolecular base pairing would have to accommodate the base changes present in the mutant SL RNAs with alterations at positions +3 to +6 with respect to the splice donor site. The second possibility also seems unlikely both for this reason and because any induced alteration of secondary structure would have to allow for efficient assembly of the SL RNA into an Sm snRNP (which is an absolute prerequisite for SL RNA function) (Maroney *et al.*, 1990 and see Figure 5).

As discussed by Bruzik *et al.* (1988), in almost all types of splicing, including the self-catalyzed splicing of group I and group II introns and snRNP mediated pre-mRNA splicing, the 5′ splice site is identified by either intra- or intermolecular base pairing interactions. Since our results indicate that *trans*-splicing does not require intramolecular base pairing interactions within the SL RNA, is any base pairing necessary to identify the 5′ splice site region? An obvious candidate for such a role in 5′ splice site identification is the nematode U1 snRNP. We have attempted to address the role of U1 snRNP in *trans*-splicing by using oligodeoxynucleotide mediated RNase H digestion. When ~90% of U1 snRNA was degraded and *cis*-splicing was inhibited by ~90%, *trans*-splicing using either wild type SL RNA or any of the mutant derivatives described here was unaffected (G.J.Hannon, P.A.Maroney and T.W. Nilsen, unpublished observation). While these experiments suggest that *trans*-splicing might not require U1 snRNA, they cannot be considered definitive. In addition to assessing the role, if any, of U1 snRNP in *trans*-splicing, it will be equally important to determine whether U5 snRNA is required in *trans*-splicing. This question is especially relevant in light of the recent demonstration by Newman and Norman (1991) that U5 plays a role in 5′ splice site identification in yeast. The most straightforward way to assess the possible role of U1 or U5 in *trans*-splicing would be to analyze *trans*-spliceosomal constituents. However, to date, we have been unable to establish conditions which make this (still hypothetical) complex stable to either gradient centrifugation or native gel electrophoresis.

While it remains possible that U1 and U5 snRNPs are involved in nematode *trans*-splicing, it seems equally likely that these U snRNPs, which are apparently absent from trypanosomes (see Agabian, 1990), are not required for *trans*-splicing. In this view, interactions between the SL snRNP and other spliceosomal factors may be sufficient to place the 5′ splice site in a position where it can be acted upon by the catalytic machinery of the spliceosome. Regardless of the mechanism of splice site identification, this process is remarkably flexible in *trans*-splicing. The region which separates the splice site from the rest of the SL RNA abides substantial variation in length (9−21 nucleotides) and sequence content (Figure 3) and the sequence immediately surrounding the splice site itself is highly tolerant to mutation (Figure 2, and unpublished data).

Exon flexibility in trans-splicing

We have shown that the 22 nt SL sequence is dispensable for *trans*-splicing. This was surprising in light of the fact that the SL sequence has been stringently conserved in widely diverged nematodes (Bektesh *et al.*, 1988; Zeng *et al.*, 1990; Nilsen, 1989). However, we have previously shown that the SL sequence itself is an essential promoter element for SL RNA transcription by RNA polymerase II. Thus, SL

sequence conservation can, at least in part, be explained by constraints imposed by the binding specificity of a transcription factor (Hannon *et al.* 1990a). It also remains possible that the SL sequence imparts specific properties to the mRNAs to which it becomes affixed.

The SL RNA 'intron' appears to be capable of productively delivering any exon to the *trans*-spliceosome (Figures 4 and 5 and unpublished data). The function of SL RNAs bearing heterologous exons provides strong support for the idea, originally discussed by Sharp (1987), that SL RNAs can be considered chimeric molecules in which an exon has been fused to an snRNA sequence. It is now necessary to determine the features of the snRNA domain which confer its specificity and ability to function in *trans*-splicing. To date, we recognize only two absolute requirements for SL RNA utilization. The first is a splice donor site, although the sequences which constitute a functional *trans* 5′ splice site remain to be defined. Given the apparent lack of sequence constraints (see above) it seems possible that a GU dinucleotide some distance (still undefined) from the body of the SL RNA could be sufficient to constitute a 5′ splice site. The second requirement is for assembly into an Sm snRNP. With regard to this point, we have not yet been able to uncouple SL RNA assembly from splicing. We have tested the ability of a large number of mutant SL RNAs to function in *trans*-splicing. For each mutant that contained a functional splice site, the efficiency of *trans*-splicing was directly correlated with the extent to which the altered SL RNAs assembled into Sm snRNPs. If a functional splice site and the ability to assemble into an Sm snRNP are the only relevant characteristics of an SL RNA, it may be possible to design 'artificial' SL RNAs which function in *trans*-splicing. Experiments are in progress to test this possibility.

Materials and methods

In vitro trans-splicing and preparation of splicing substrates

RNAs used in splicing reactions were transcribed by T3 polymerase as previously described (Maroney *et al.*, 1990). Normal and mutant SL RNAs were internally labelled with [α-^{32}P]GTP to a specific activity of ~30 000 c.p.m./ng. Some SL RNAs were 5′ end labelled using vaccinia guanylyl transferase to cap uncapped, unlabelled SL RNA with [α-^{32}P]GTP (Maroney *et al.*, 1990). Alternatively, 5′ hydroxylated SL RNAs were prepared by priming T3 RNA polymerase transcription with the dinucleotide primer OH-GpG (GpG:GTP = 10:1). Transcripts were subsequently labelled with [γ-^{32}P]ATP using polynucleotide kinase. All transcripts were purified on 5% acrylamide−8 M urea gels (Hannon *et al.*, 1989).

In vitro splicing reactions in 12.5 µl contained 7.5 µl of *A.lumbricoides* extract (Hannon *et al.*, 1990b), 60 mM KCl, 4.2 mM $MgCl_2$, 2 mM ATP, 20 mM creatine phosphate, 2 mM DTT, 12 mM Tris, pH 7.6, 12% (v/v) glycerol (1× splicing buffer) and 3% PEG 8000. *Trans*-splicing reactions contained in addition 50 ng of unlabelled acceptor pre-mRNA (Hannon *et al.*, 1990b) and 200 000 c.p.m. (~6.6 ng) of either wild type or mutant SL RNA. Reactions were incubated for 60 min at 30°C then diluted to 250 µl with 250 mM NaOAc, 25 mM Tris pH 7.5, 1 mM EDTA and 0.2% SDS and digested for 15 min at 42°C with 100 µg/ml proteinase K. RNAs were recovered by phenol−chloroform extraction and ethanol precipitation, electrophoresed on 5% acrylamide−8 M urea gel and visualized by autoradiography.

Site directed mutagenesis

For site directed mutagenesis, the SL RNA T3 transcription template in pBS M13+ was transformed into *dut⁻ ung⁻ Escherichia coli* (CJ236) (Kunkel *et al.*, 1987). The plus strand of the plasmid was rescued (Vieira and Messing, 1987) and mutated as previously described (Hannon *et al.* 1990a). In each case, oligonucleotides used for mutagenesis had ~10 bases of perfect homology flanking the positions which were altered.

Native gel electrophoresis

A 12.5 µl *trans*-splicing reaction containing 200 000 c.p.m. (~6 ng) of wild type or mutant SL RNA and lacking pre-mRNA acceptor was incubated for 20 min at 30°C. Following incubation, the reaction was diluted to 120 µl with 0.5 M NaCl, 10 mM Tris pH 7.6, 5% glycerol and 200 µg/ml heparin. The diluted reaction was incubated at room temperature for 10 min before 3 µl were removed and added to 30 µl containing 60 mM KCl, 10 mM Tris pH 7.6, 5% glycerol, 200 µg/ml heparin, 0.01% xylene cyanol and 0.01% bromophenol blue. This sample was loaded onto a 1.5 mm thick, 20 cm long, 6% acrylamide, 0.5× TBE gel (acrylamide:bisacrylamide = 80:1). The gel was electrophoresed for 3 h at 25 mA. The gel was dried and complexes were visualized by autoradiography.

Construction of SL RNAs with artificial exons

Bases 3−20 of the SL RNA sequence were replaced with CCT by site directed mutagenesis to create a unique *Stu*I restriction site. This *Stu*I site included the last base of the T3 promoter (A) and the first two bases of the SL RNA transcription template (GG). Digestion with *Stu*I allowed the insertion of random fragments derived from either pBR322 or pBS m13+ upstream of the SL RNA splice donor site. In each case two guanosine residues were retained downstream of the T3 promoter to facilitate efficient transcription and the sequence, CCTAG, remained upstream of the splice donor site. In addition to these sequences the construct containing the 29 base artificial exon contained bases 968−989 of pBS m13+. The construct with the 61 base artificial exon contained bases 3195−44 of pBS m13+; the 85 base exon contained bases 26−103 of pBR322; and the 246 base exon contained bases 1022−1260 of pBR322. Point mutations were introduced into the Sm binding sites of each of these constructs, changing the second uridine residue of the Sm binding site to guanosine, as described previously (Maroney *et al.*, 1990). Each of these SL RNAs was synthesized with T3 RNA polymerase as described above.

Acknowledgements

We thank G.E.Hannon and J.A.Denker for sequence analysis of mutant constructs and Dr J.A.Steitz for critical reading of an earlier draft of the manuscript. Supported by PHS GM31428; the John D. and Catherine T.MacArthur Foundation and the Burroughs Wellcome Fund.

References

Agabian,N. (1990) *Cell*, **61**, 1157−1160.
Bektesh,S.L., van Doren,K.V. and Hirsh,D. (1988) *Genes Dev.*, **2**, 1277−1283.
Blumenthal,T. and Thomas,J. (1988) *Trends Genet.*, **4**, 305−308.
Borst,P. (1986) *Annu. Rev. Biochem.*, **55**, 701−732.
Bruzik,J.P. and Steitz,J.A. (1990) *Cell*, **62**, 889−899.
Bruzik,J.P., van Doren,K., Hirsh,D. and Steitz,J.A. (1988) *Nature*, **335**, 559−562.
Duchenne,M., Low,A., Schweizer,A. and Domdey,H. (1988) *Nucelic Acids Res.*, **16**, 7233−7239.
Hannon,G.J., Maroney,P.A., Branch,A., Benenfeld,B.J., Robertson,H.D. and Nilsen,T.W. (1989) *Mol. Cell. Biol.*, **9**, 4422−4431.
Hannon,G.J., Maroney,P.A., Ayers,D.G., Shambaugh,J.D. and Nilsen,T.W. (1990a) *EMBO J.*, **9**, 1915−1921.
Hannon,G.J., Maroney,P.A., Denker,J.A. and Nilsen,T.W. (1990b) *Cell*, **61**, 1247−1255.
Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) *Methods Enzymol.*, **154**, 367−382.
Maroney,P.A., Hannon,G.J., Denker,J.A. and Nilsen,T.W. (1990) *EMBO J.*, **9**, 3667−3673.
Newman,A. and Norman,C. (1991) *Cell*, **65**, 115−123.
Nilsen,T.W. (1989) *Exp. Parasitol.*, **4**, 305−308.
Nilsen,T.W., Shambaugh,J., Denker,J.A., Chubb,G., Faser,C., Putnam,L. and Bennett,K. (1989) *Mol. Cell. Biol.*, **9**, 3543−3547.
Sharp,P.A. (1987) *Cell*, **50**, 147−148.
Thomas,J.D., Conrad,R.C. and Blumenthal,T. (1988) *Cell*, **54**, 533−539.
Turner,D.H., Sugimoto,N., Jaeger,J.A., Longfellow,C.E., Freier,S.M. and Kierzek,R. (1987) *Cold Spring Harbor Symp. Quant. Biol.*, **52**, 123−133.
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.
Zeng,W., Alarcon,C.M. and Donelson,J.E. (1990) *Mol. Cell. Biol.*, **10**, 2765−2773.
Zuker,M. and Stiegler,P. (1981) *Nucleic Acids Res.*, **9**, 133−148.

Received on August 9, 1991