

Enhancer-dependent interaction between 5' and 3' splice sites in trans

JAMES P. BRUZIK*[†] AND TOM MANIATIS[‡]

*Case Western Reserve University, Department of Molecular Biology and Microbiology, Center for Molecular Parasitology, 10900 Euclid Avenue, Cleveland, OH 44106; and [‡]Harvard University, Department of Molecular and Cellular Biology, 7 Divinity Avenue, Cambridge, MA 02138

Contributed by Tom Maniatis, May 3, 1995

ABSTRACT Splice-site selection and alternative splicing of nuclear pre-mRNAs can be controlled by splicing enhancers that act by promoting the activity of upstream splice sites. Here we show that RNA molecules containing a 3' splice site and enhancer sequence are efficiently spliced in trans to RNA molecules containing normally cis-spliced 5' splice sites or to normally trans-spliced spliced leader RNAs from lower eukaryotes. In addition, we show that this reaction is stimulated by (Ser + Arg)-rich splicing factors that are known to promote protein–protein interactions in the cis-splicing reaction. Thus, splicing enhancers facilitate the assembly of protein complexes on RNAs containing a 3' splice site, and this complex is sufficiently stable to functionally interact with 5' splice sites located on separate RNAs. This trans-splicing is mediated by interactions between (Ser + Arg)-rich splicing factors bound to the enhancer and general splicing factors bound to the 5' and 3' splice sites. These same interactions are likely to play a crucial role in alternative splicing and splice-site selection in cis.

Specific interactions between 5' and 3' splice sites in nuclear pre-mRNA occur during the formation of the earliest known functional pre-spliceosome complex (1–3). Both small nuclear ribonucleoprotein particles (snRNPs) (1–3) and protein factors (2–5) have been implicated in this interaction. Among the protein factors required are (Ser + Arg)-rich splicing factors (SR proteins) (3–5), a family of general splicing factors (6, 7). Recent studies have shown that the utilization of weak splice sites can be dramatically stimulated by downstream regulatory elements termed splicing enhancers (8–11) that require binding of SR proteins for activity (8, 11). In addition, SR proteins have been demonstrated to promote splice-site switching in constructs with duplicated junctions (12–14), mediate alternative splicing (8, 11), and commit pre-mRNAs to the splicing pathway (15).

Several recent studies have been designed to address the question of specific interactions between members of the SR protein family as well as between SR proteins and other known splicing factors (5, 16, 17). A model for splice-site selection was proposed (5) on the basis of these experiments, the observation that SR proteins may facilitate splice-site interactions (4), and the determination of the protein composition of early functional spliceosome complexes (2, 3). In this model U1 snRNP interacts with the 5' splice site via base pairing with the U1 snRNA, and the heterodimeric splicing factor U2AF binds specifically to the pyrimidine tract at the 3' splice site. SR proteins then mediate the joining of the two complexes by simultaneously interacting with the integral U1 snRNP 70-kDa protein at the 5' splice site and the 35-kDa subunit of U2AF at the 3' splice site. Specific protein–protein interactions were also implicated in the functional interaction between splicing enhancers and weak 3' splice sites (5, 16). In this case SR

proteins and the splicing regulators Transformer (Tra) and Transformer 2 (Tra2), which bind specifically to splicing enhancers, were proposed to facilitate binding of U2AF to the upstream 3' splice site through interactions with the 35-kDa subunit of U2AF.

On the basis of these observations we carried out experiments to determine whether interactions between proteins bound to splicing enhancers and 3' splice sites lead to the formation of complexes that are sufficiently stable to functionally interact with 5' splice sites located on separate RNA molecules. We find that *in vitro* trans splicing with both weak and strong 3' splice sites is dramatically stimulated when the 3' exon contains a splicing enhancer.

MATERIALS AND METHODS

Plasmid Construction. Appropriate fragments of the *Drosophila melanogaster* *dsx* and *ftz* genes were cloned into vector pSP73 (Promega) for *in vitro* transcription. The *dsx* 5' splice-site-containing fragment (*dsx* 5' SS) was subcloned from the *dsx* construct D1 (18) by removing an *Xho* I–*Bst*XI fragment containing 48 bp of intron sequence and 96 bp of exon 3 (including polylinker sequence at the 5' end) and inserting this into *Xho* I–*Pvu* II-digested pSP73. The *ftz* 5' splice-site-containing fragment (*ftz* 5' SS) was subcloned from the construct T7-*ftz* (19) by removing a *Pvu* II–*Acc* I fragment containing 69 bp of intron sequence and 93 bp of exon 1 and inserting this into *Pvu* II-digested pSP73. The spliced leader (SL) RNA clones were as described (20). The *dsx* 3' SS/300RE clone (where RE represents repeat element) was isolated as a *Bst*XI–*Dra* I fragment from clone D1 (18) containing 62 bp of intron sequence and 541 bp of exon 4 and inserted into *Pvu* II-digested pSP73. The *dsx* 3' SS/100RE was cloned as a *Bst*XI–*Bam*HI fragment from clone D4 (18) containing 62 bp of intron sequence and 338 bp of exon 4 (including polylinker sequence at the 3' end) and was cloned into *Pvu* II–*Bam*HI-digested pSP73. Both the *dsx* 3' SS up/ASLV and *dsx* 3' SS up constructs contain an improved polypyrimidine tract [as in clone D17 (18)], either with the avian sarcoma-leukosis virus (ASLV) enhancer or no enhancer downstream.

In Vitro Splicing Reactions. Individual RNAs were transcribed *in vitro* and gel-purified before the splicing reaction. Stocks of unlabeled 3' SS substrates were also reexamined on polyacrylamide gels after experiments to ensure that they were not degraded (data not shown). Splicing reactions contained 400,000 cpm (15 ng) of uniformly labeled 5' SS RNAs or SL RNAs and 100 ng of unlabeled 3' splice-site RNA, 60% nuclear extract, 0.45 μ g of poly(A), 7.2 mM MgCl₂, 20 mM creatine phosphate, 1 mM ATP, 3% PEG 8000, 1 μ l (40 units) of RNasin (Promega), and 4 mM UTP in a 25 μ l reaction [poly(A) and equal concentrations of UTP and MgCl₂ were

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

Abbreviations: snRNPs, small nuclear ribonucleoprotein particles; SS, splice-site-containing fragment; *dsx*RE, *dsx* repeat element; ASLV, avian sarcoma-leukosis virus; SL, spliced leader; Tra, Transformer; SR proteins, (Ser + Arg)-rich splicing factors.

[†]To whom reprint requests should be addressed.

added as carrier to reduce nonspecific background on the gel]. Reactions were incubated at 30°C for 2 hr followed by treatment with proteinase K and phenol/chloroform/isoamylalcohol, 50:49:1, extraction. RNA was run on 5% polyacrylamide/7 M urea/0.5× TBE (90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 8.3) gels and visualized by autoradiography. Trans-spliced product resulting from the coincubation of labeled *dsx* 5' SS and unlabeled *dsx* 3' SS/100RE was sequenced and found to be accurate (data not shown). Sizes of the products with labeled *dsx* 5' SS are 402 nt *dsx* 3' SS/100RE, 260 nt *dsx* 3' SS up/ASLV, and 369 nt for splicing to the cryptic 3' splice site in *dsx* 3' SS/300RE. Sizes of the products with the *ftz* 5' SS are all 9 nt longer than with the *dsx* 5' SS. Sizes of the products with the *Leptomonas collosoma* (trypanosomatid) SL RNA are 339 nt *dsx* 3' SS/100RE, 197 nt *dsx* 3' SS up/ASLV, and 306 nt for splicing to the cryptic 3' splice site in *dsx* 3' SS/300RE. Products with the *Caenorhabditis elegans* (nematode) SL RNA are all 9 nt shorter than with the *L. collosoma* SL RNA. Reactions in S100 were similar, containing 400,000 cpm (15 ng) of uniformly labeled *dsx* 5' SS transcript and 100 ng of unlabeled 3' SS transcript, 40% S100 extract, 0.45 μg of poly(A), 10 mM MgCl₂, 5 mM creatine phosphate, 5 mM ATP, 3% PEG 8000, 1 μl (40 units) of RNasin (Promega), and 4 mM UTP in a 25-μl reaction (carrier as above) with or without 0.5 μg of purified SR proteins (6). Reactions were incubated for 2 hr at 30°C and were terminated and analyzed as described.

RESULTS

Enhancer-Dependent Trans Splicing. To determine whether splicing enhancers can promote trans splicing *in vitro*, we tested RNA substrates containing the *D. melanogaster* doublesex (*dsx*) repeat element (*dsxRE*), a splicing enhancer located downstream from a weak 3' splice site that does not function in cis splicing in the absence of an enhancer (18, 21, 22). Previous studies have shown that the position of the *dsxRE* relative to this 3' splice site plays a critical role in its ability to function as a splicing enhancer (18). When the *dsxRE* is located at its normal position 300 nt downstream from the 3' splice site, it does not function as a splicing enhancer *in vitro* in the absence of the splicing regulators Tra and Tra2 (18). However, when the *dsxRE* is located within 100 nt of the 3' splice site, it is active as a splicing enhancer in the absence of Tra and Tra2 (18).

Fig. 1 shows that the 5' splice site-containing RNA substrates tested include *dsx* 5' SS and *ftz* 5' SS, fragments of normally cis-spliced pre-mRNAs containing a 5' splice site, but lacking a 3' splice site, as well as the *C. elegans* (nematode) and *L. collosoma* (trypanosomatid) SL RNAs. In contrast, the 3' splice-site-containing fragments lack a 5' splice site and include *dsx* 3' SS/300RE and *dsx* 3' SS/100RE (containing the *dsx* 3' splice site and exon 4 with the *dsxRE* located 300 and 100 nt, respectively, downstream from the 3' splice site) as well as up mutants of the *dsx* 3' splice site containing either the ASLV enhancer (*dsx* 3' SS up/ASLV) or no enhancer (*dsx* 3' SS up) (Fig. 1). RNA substrates containing the 3' splice-site up-mutation function in cis splicing *in vitro* in the absence of a splicing enhancer (18).

Radiolabeled 5' splice-site-containing RNAs were separately incubated with each of the unlabeled 3' splice-site-containing RNA substrates in HeLa cell nuclear extracts, and the products were examined by electrophoresis on denaturing polyacrylamide gels (Fig. 2). Spliced products were observed when unlabeled *dsx* 3' SS/100RE RNA, which contains a constitutively active splicing enhancer, was incubated with either of the pre-mRNA fragments containing a 5' splice site (Fig. 2A, lanes 4 and 10). The sizes of these products are exactly those expected for bimolecular splicing reactions between the 5' and 3' SS RNA substrates. By contrast, no spliced products were observed when either of these labeled 5' splice-site-containing RNAs was incubated alone (Fig. 2A, lanes 2 and 8)

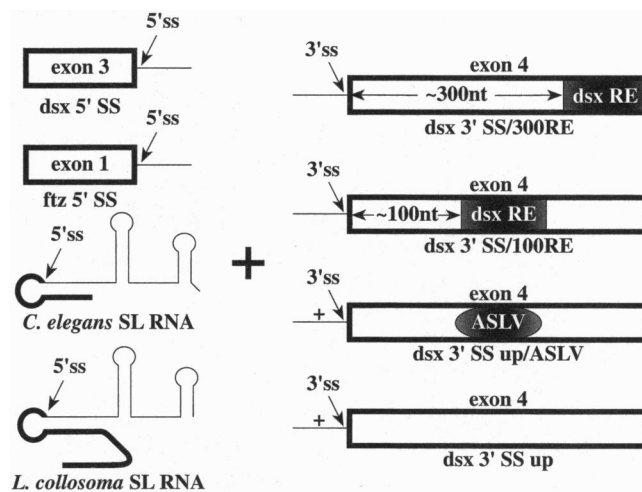


FIG. 1. RNA substrates used to assay enhancer-dependent trans splicing. The 5' and 3' half molecules were generated from *dsx* constructs D1 and D4 (18) and from *ftz* clone T7-*ftz* (19). In cis, *in vitro* splicing of the D1 transcript requires Tra and Tra2 proteins, whereas the D4 construct is spliced in the absence of Tra and Tra2. The *ftz* pre-mRNA is efficiently spliced in HeLa cell extracts (19). The 5' splice-site-encompassing fragments contain exon 3 from the *dsx* gene (the last common exon) or exon 1 of the *ftz* gene, and the trypanosomatid (*L. collosoma*) or nematode (*C. elegans*) SL RNAs. The 3' splice-site fragments contain exon 4 (the female-specific exon) of the *dsx* gene with the *dsxRE* located either 300 nt downstream of the 3' splice site (*dsx* 3' SS/300RE) or 100 nt downstream (*dsx* 3' SS/100RE). The *dsxRE* (shaded box) is a well characterized splicing enhancer (8, 18, 22). In addition, two other 3' splice-site-containing fragments of *dsx* were utilized as controls. These are the *dsx* 3' SS up/ASLV and *dsx* 3' SS up RNAs containing an improved 3' splice site (up) [as in clone D17 (18)] and either the ASLV enhancer (3) (shaded oval) or no enhancer, respectively. Exon sequences are denoted by thick lines or boxes.

or with the *dsx* 3' SS/300RE RNA (Fig. 2A, lanes 3 and 9). The *dsx* 3' SS/300RE RNA should not function in trans splicing because it does not function in cis in the absence of Tra and Tra2 (18) (the faint bands observed in the *dsx* 3' SS/300RE lanes that migrate below the *dsx* 3' SS/100RE products correspond in size to weak utilization of a cryptic 3' splice site that is much closer to the *dsxRE*). Interestingly, addition of Tra and Tra2 does not rescue trans splicing with these substrates, although control of cis splicing in the same extract is greatly stimulated (data not shown). Thus, the Tra- and Tra2-dependent splicing enhancer complex formed on the 3' splice site is not sufficient for trans splicing *in vitro*. An up mutant in the pyrimidine tract of the *dsx* 3' splice site (*dsx* 3' SS up) which is efficiently cis spliced in the absence of Tra and Tra2 (18) does not trans splice in the absence of a splicing enhancer (Fig. 2A, lanes 6 and 12). However, this 3' splice site is trans spliced in the presence of a nearby downstream splicing enhancer (*dsx* 3' SS up/ASLV, Fig. 2A, lanes 5 and 11).

Previously, we demonstrated that trans splicing could occur at low levels with SL RNAs in HeLa nuclear extract (20). Therefore we asked whether a splicing enhancer would improve the efficiency of this reaction. As seen in Fig. 2B, both the *C. elegans* (lanes 7–11) and *L. collosoma* (lanes 1–5) SL RNAs are efficiently trans spliced (~15 fold better than pre-mRNA 5' splice-site-containing fragments) in HeLa nuclear extract only when an active splicing enhancer is present in the 3' exon. We conclude that trans splicing is promoted by a constitutively active splicing enhancer located downstream of a 3' splice site.

SR Proteins Promote the Functional Interaction of 5' and 3' Splice Sites in Trans. SR proteins are required for cis splicing (7) and have been shown to stimulate enhancer-

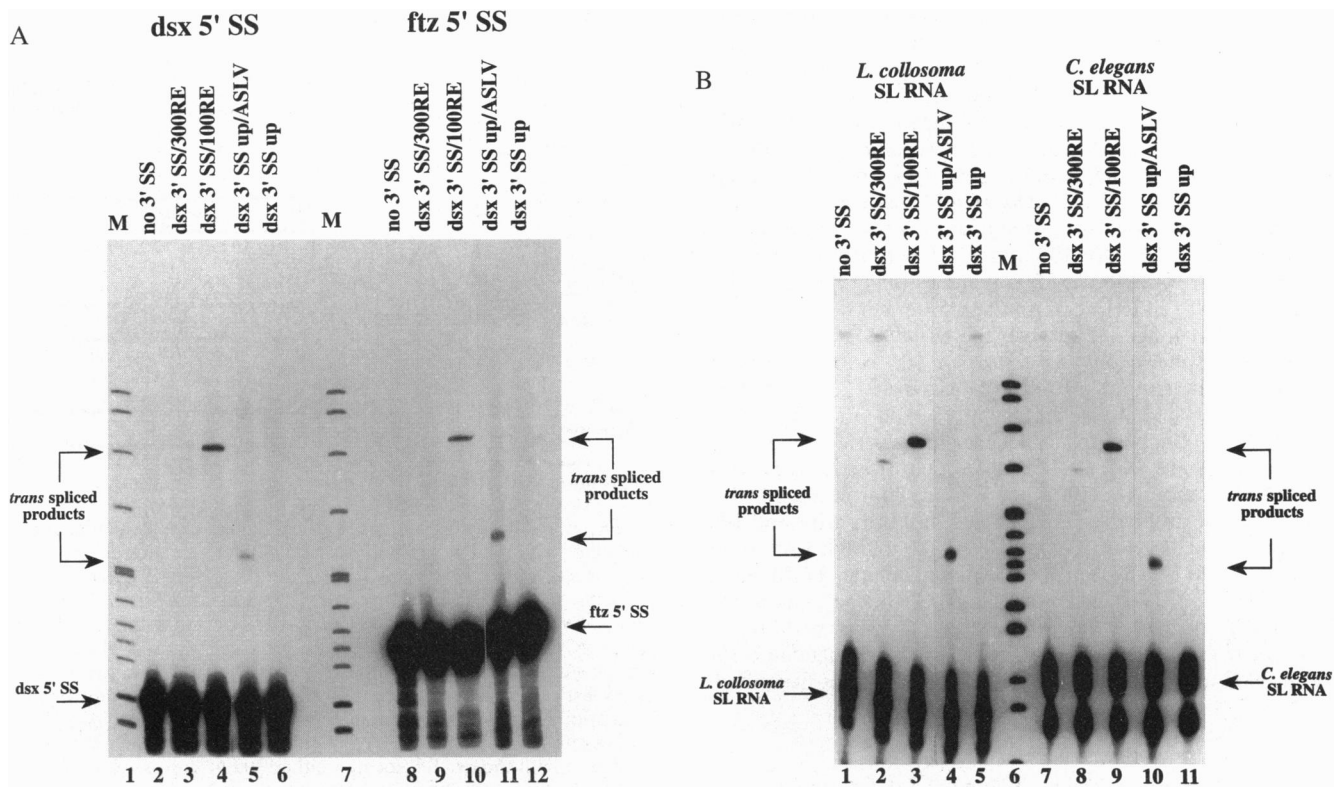


FIG. 2. Trans splicing of 5' and 3' splice-site-containing molecules *in vitro*. (A) Radiolabeled 5' splice-site-containing fragments (*dsx* 5' SS or *ftz* 5' SS) were synthesized *in vitro* and incubated in HeLa cell nuclear extracts alone or with the 3' SS half molecules illustrated in Fig. 1. Lanes: 2, labeled *dsx* 5' SS RNA was incubated in nuclear extract alone; 3, unlabeled *dsx* 3' SS/300RE was incubated with labeled *dsx* 5' SS RNA; 4, labeled *dsx* 5' SS RNA was incubated with *dsx* 3' SS/100RE RNA; 5, labeled *dsx* 5' SS RNA was incubated with *dsx* 3' SS up/ASLV RNA; 6, unlabeled *dsx* 3' SS up RNA was incubated with labeled *dsx* 5' SS RNA; 8–12, the same order of unlabeled 3' SS fragments as lanes 2–6 except that they are incubated with labeled *ftz* 5' SS; 1 and 7, pBR322 *Msp* I-digested labeled DNA size markers. (B) Radiolabeled SL RNAs (*C. elegans* or *L. collosoma*) were synthesized *in vitro* and incubated in HeLa cell nuclear extracts alone or with the 3' SS half molecules described above. The order of the 3' SS acceptors is the same as in A, with the *L. collosoma* reactions in lanes 1–5 and the *C. elegans* reactions in lanes 7–11. pBR322 *Msp* I-digested labeled DNA size markers are in lane 6. Positions of the trans-spliced products and the labeled 5' SS RNAs and SL RNAs are indicated.

dependent cis splicing (8, 10, 11). To determine whether SR proteins can also stimulate enhancer-dependent trans splicing we carried out the splicing reaction in S100 extracts that lack SR proteins but contain all other components necessary for splicing (23). The addition of SR proteins to the S100 extracts did not result in trans splicing between the *dsx* 5' SS and *dsx* 3' SS/300RE RNA (Fig. 3, lanes 4 and 5). Thus, SR proteins are not sufficient to stimulate trans splicing in the absence of an active splicing enhancer. Low levels of trans splicing were observed with the *dsx* 3' SS/100RE RNA in S100 extracts lacking additional SR proteins (Fig. 3, lane 6). By contrast, addition of purified HeLa SR proteins significantly increased the level of trans splicing with this 3' SS substrate (lane 7). We conclude that SR proteins promote enhancer-dependent trans splicing.

DISCUSSION

Low levels of trans splicing of normally cis-spliced RNAs were previously reported (24, 25), and in at least one of these cases base-pair interactions between the two substrates were thought to be required for trans splicing (26, 27). In addition, trans splicing between natural substrates from lower eukaryotes was demonstrated in mammalian cells *in vivo* and *in vitro* (20). Here we show that fragments of normally cis-spliced pre-mRNA substrates, as well as SL RNAs, are trans spliced in mammalian nuclear extracts and that this reaction is promoted by splicing enhancers and SR proteins.

Previous studies have shown that SR proteins associate with complexes that form on both the 5' and 3' splice sites (3, 4, 17,

28). Recently, it was proposed that SR proteins are able to substitute for the functions of the U1 snRNP when added in excess to either U1 snRNP-depleted or -inactivated extracts (29, 30). In addition, SR proteins have been shown to influence the utilization of different 5' splice sites by regulating the interaction of U1 snRNP with the pre-mRNA (31, 32). Moreover, the presence of either a downstream 5' splice site or splicing enhancer has been shown to promote the formation of complexes on the 3' splice site (3, 33, 34). In addition, the formation of these complexes is promoted by SR proteins (3). Taken together, these results indicate that SR proteins play a central role in spliceosome assembly and splice-site interactions. SR proteins also bind to splicing enhancers (3, 8, 10, 11), and they are required for enhancer activity (8, 11).

These observations in conjunction with the demonstration here that splicing enhancers and SR proteins promote trans splicing *in vitro* show that the stable complexes assembled on 5' and 3' splice sites located on separate RNA molecules can functionally interact and that SR proteins play a critical role in this interaction. This reaction dramatically illustrates the ability of splicing enhancers to direct the splicing machinery to specific 3' splice sites in alternatively spliced pre-mRNAs and suggests that if trans splicing is required for the production of certain vertebrate mRNAs (35–39), then splicing enhancers and SR proteins are likely to be involved. Significantly, SR proteins isolated from naturally trans-splicing organisms promote authentic, SL RNA-dependent trans splicing *in vitro* (J.P.B., P. Maroney, and T. Nilsen, unpublished data). Thus, the role of SR proteins in mediating splice-site interactions in splicing complexes is highly conserved in evolution.

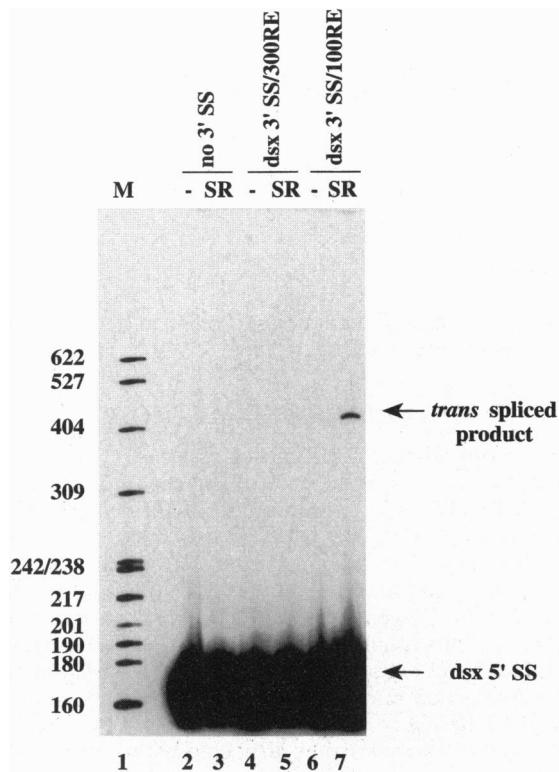


FIG. 3. Stimulation of trans splicing by purified SR proteins. Radiolabeled *dsx* 5' SS RNA was incubated either alone or together with unlabeled 3' splice-site-containing transcripts in HeLa cytoplasmic extract S100. Lanes: 2, incubation of labeled *dsx* 5' SS alone; 4 and 6, incubation of *dsx* 5' SS RNA with each of the indicated 3' SS-containing RNAs. Each of these combinations was also incubated in the S100 extract with purified HeLa SR proteins (6). These reactions are seen in lanes 3, 5, and 7, with the same order of 3' SS fragments as the nonsupplemented reactions. Lane 1 contains pBR322 *Msp* I-digested labeled DNA size markers. Positions of the trans-spliced product and labeled *dsx* 5' SS RNA are indicated.

We thank Ming Tian for providing the *dsx* clones used to construct the 5' and 3' half molecules and Ping Zuo for control clones. We also thank Joan Steitz, Pat Maroney, Tim Nilsen, Jo Ann Wise, and members of the Maniatis laboratory for helpful discussions and comments on the manuscript. This research was supported in part by Postdoctoral Fellowship 61-862 from the Jane Coffin Childs Memorial Fund for Medical Research (J.P.B.), start-up funding from Case Western Reserve University School of Medicine (J.P.B.), and Grant GM42231 from the National Institutes of Health (T.M.).

1. Wise, J. A. (1993) *Science* **262**, 1978–1979.
2. Michaud, S. & Reed, R. (1993) *Genes Dev.* **5**, 2534–2546.
3. Staknis, D. & Reed, R. (1994) *Mol. Cell. Biol.* **14**, 7670–7682.

4. Fu, X.-D. & Maniatis, T. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1725–1729.
5. Wu, J. Y. & Maniatis, T. (1993) *Cell* **75**, 1061–1070.
6. Zahler, A. M., Lane, W. S., Stolk, J. A. & Roth, M. B. (1992) *Genes Dev.* **6**, 837–847.
7. Horowitz, D. S. & Krainer, R. (1993) *Trends Genet.* **10**, 100–106.
8. Tian, M. & Maniatis, T. (1993) *Cell* **74**, 105–114.
9. Watakabe, A., Tanaka, K. & Shimura, Y. (1993) *Genes Dev.* **7**, 407–418.
10. Lavigne, A., La Branche, H., Kornblihtt, A. R. & Chabot, B. (1993) *Genes Dev.* **7**, 2405–2417.
11. Sun, Q., Mayeda, A., Hampson, R. K., Krainer, A. R. & Rottman, F. M. (1993) *Genes Dev.* **7**, 2598–2608.
12. Ge, H. & Manley, J. L. (1990) *Cell* **62**, 25–34.
13. Krainer, A. R., Conway, G. C. & Kozak, D. (1990) *Cell* **62**, 35–42.
14. Fu, X.-D., Mayeda, A., Maniatis, T. & Krainer, A. R. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 11224–11228.
15. Fu, X.-D. (1993) *Nature (London)* **365**, 82–85.
16. Amrein, H., Hedley, M. L. & Maniatis, T. (1994) *Cell* **76**, 735–746.
17. Kohtz, J. D., Jamison, S. F., Will, C. L., Zuo, P., Lührmann, R., Garcia-Blanco, M. A. & Manley, J. L. (1994) *Nature (London)* **368**, 119–124.
18. Tian, M. & Maniatis, T. (1994) *Genes Dev.* **8**, 1703–1712.
19. Reed, R. & Maniatis, T. (1985) *Cell* **41**, 95–105.
20. Bruzik, J. P. & Maniatis, T. (1992) *Nature (London)* **360**, 692–695.
21. Hedley, M. L. & Maniatis, T. (1991) *Cell* **65**, 579–586.
22. Tian, M. & Maniatis, T. (1992) *Science* **256**, 237–240.
23. Krainer, A. R., Conway, G. C. & Kozak, D. (1990) *Genes Dev.* **4**, 1158–1171.
24. Solnick, D. (1985) *Cell* **42**, 157–164.
25. Konarska, M. M., Padgett, R. A. & Sharp, P. A. (1985) *Cell* **42**, 165–171.
26. Solnick, D. (1986) *Cell* **44**, 211.
27. Sharp, P. A. & Konarska, M. M. (1986) *Cell* **44**, 211.
28. Eperon, I. C., Ireland, D. C., Smith, R. A., Mayeda, A. & Krainer, A. R. (1993) *EMBO J.* **12**, 3607–3617.
29. Crispino, J. D., Blencowe, B. J. & Sharp, P. A. (1994) *Science* **265**, 1866–1869.
30. Tarn, W.-Y. & Steitz, J. A. (1994) *Genes Dev.* **8**, 2704–2717.
31. Tarn, W.-Y. & Steitz, J. A. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 2504–2508.
32. Zahler, A. M. & Roth, M. B. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 2642–2646.
33. Hoffman, B. E. & Grabowski, P. J. (1992) *Genes Dev.* **6**, 2554–2568.
34. Wang, Z., Hoffman, H. M. & Grabowski, P. J. (1995) *RNA* **1**, 21–35.
35. Perbal, B. & Vellard, M. (1990) *C. R. Acad. Sci. Paris* **311**, 467–472.
36. Sureau, A. & Perbal, B. (1991) *C. R. Acad. Sci. Paris* **312**, 323–328.
37. Shimizu, A., Nussenzweig, M. C., Mizuta, T.-R., Leder, P. & Honjo, T. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8020–8023.
38. Shimizu, A., Nussenzweig, M. C., Han, H., Sanchez, M. & Honjo, T. (1991) *J. Exp. Med.* **173**, 1385–1393.
39. Shimizu, A. & Honjo, T. (1993) *FASEB J.* **7**, 149–154.