RNA (2001), *7*:793–805. Cambridge University Press. Printed in the USA.
Copyright © 2001 RNA Society.

Functional selection of splicing enhancers that stimulate *trans*-splicing in vitro

LISA A. BOUKIS and JAMES P. BRUZIK

Center for RNA Molecular Biology, Department of Molecular Biology and Microbiology,
Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA

ABSTRACT

The role of exonic sequences in naturally occurring *trans*-splicing has not been explored in detail. Here, we have identified *trans*-splicing enhancers through the use of an iterative selection scheme. Several classes of enhancer sequences were identified that led to dramatic increases in *trans*-splicing efficiency. Two sequence families were investigated in detail. These include motifs containing the element $^{G}/_{C}GAC^{G}/_{C}$ and also 5′ splice site-like sequences. Distinct elements were tested for their ability to function as splicing enhancers and in competition experiments. In addition, discrete *trans*-acting factors were identified. This work demonstrates that splicing enhancers are able to effect a large increase in *trans*-splicing efficiency and that the process of exon definition is able to positively enhance *trans*-splicing even though the reaction itself is independent of the need for the 5′ end of U1 snRNA. Due to the presence of internal introns in messages that are *trans*-spliced, the natural arrangement of 5′ splice sites downstream of *trans*-splicing acceptors may lead to a general promotion of this unusual reaction.

Keywords: exon definition; nematode; splicing enhancer; *trans*-splicing

INTRODUCTION

Pre-mRNA splicing in nematodes includes both conventional *cis*-splicing as well as *trans*-splicing (for reviews, see Blumenthal & Steward, 1997; Nilsen, 1997). In the *trans*-splicing reaction, the 5′ exon is donated from a separate transcript, the spliced leader RNA (SL RNA), which contains hallmarks of the small nuclear RNAs (snRNAs) involved in standard splicing reactions (Bruzik et al., 1988; Thomas et al., 1988; Nilsen et al., 1989). Required cofactors for *trans*-splicing include the U2, U5, and U4/U6 small nuclear ribonucleoprotein particles (snRNPs; Hannon et al., 1991; Maroney et al., 1996) as well as serine/arginine-rich splicing factors (SR proteins; Sanford & Bruzik, 1999b). Although extensive similarities between the required cofactors for *cis*- and *trans*-splicing have been documented, there is a clear difference in the necessity for U1 snRNP at the 5′ splice site (Hannon et al., 1991). Thus, the *trans*-splicing reaction naturally occurring in nematodes has been deemed independent of the need for the 5′ end of U1 that normally base pairs with the 5′ splice site.

Reprint requests to: James P. Bruzik, Center for RNA Molecular Biology, Department of Molecular Biology and Microbiology, Case Western Reserve University School of Medicine, 10900 Euclid Avenue, Cleveland, Ohio 44106, USA; e-mail: jxb83@po.cwru.edu.

A large body of work has elucidated many of the *cis*-acting elements involved in pre-mRNA splicing in higher eukaryotes (for review, see Reed & Palandjian, 1997). These include the 3′ splice site AG dinucleotide whose necessity is coupled to the length of the preceding polypyrimidine tract (Reed, 1989). Because *trans*-splicing substrates lack the potential for spliceosomal interactions across an intron, they have been suggested to be exquisitely AG-dependent at the 3′ splice site (Wu et al., 1999). Additionally, the effects of 3′ exonic sequences on splicing of an upstream intron were documented over a decade ago (Somasekhar & Mertz, 1985; Reed & Maniatis, 1986). More recently, discrete elements termed exonic splicing enhancers (ESEs) have been identified that promote upstream splicing in both alternatively spliced (for review, see Wang & Manley, 1997) as well as constitutive pre-mRNAs (Schaal & Maniatis, 1999a). Many of these ESEs have been shown to interact with SR proteins. The two major domains of SR proteins, an N-terminal RNA recognition motif(s) (RRM) and a C-terminal region rich in RS dipeptides, allow for RNA binding and protein–protein interactions, respectively (for review, see Graveley, 2000). In addition to the predominantly purine-rich ESEs, several examples of other general classes of splicing enhancers have been identified. These include the A/C-rich splicing enhancers (Coulter et al.,

1997), intronic splicing enhancers (for examples, see Wang et al., 1997), and splicing enhancers that modulate 5′ splice site usage (Humphrey et al., 1995; Bourgeois et al., 1999; Côté et al., 1999; Selvakumar & Helfman, 1999). The concept of elements in an exon being able to promote complex formation on and splicing of a nearby intron can also be extended to the positive influence seen when the next 5′ splice site downstream is included in the substrate (for review, see Berget, 1995). This exon definition effect (Robberson et al., 1990) can also be observed in pre-mRNAs where a 5′ splice site-like sequence is present at a location other than a bona-fide splice junction (Lou et al., 1995). Interestingly, the ability of U1 snRNP to promote splicing of an upstream intron has also been observed when the upstream intron does not require U1 (Wu & Krainer, 1996). In this case, the splicing of a minor class (AT-AC) intron was stimulated via the presence of a classical 5′ splice site downstream. Therefore, a splicing reaction that did not require U1 snRNP for recognition of the 5′ splice site was actually dependent upon U1 action from a downstream position for efficient splicing.

The role(s) of exonic sequences in authentic *trans*-splicing has not been explored at length (for review, see Bruzik, 1996). A recent report from work performed in a trypanosomatid system has demonstrated that exonic sequences can modulate usage of the upstream *trans* acceptor. In this example, both changes in or substitution of the exon of the α-tubulin gene in *Trypanosoma brucei* led to differences in the utilization of its 3′ splice site in a splice site competition assay (López-Estraño et al., 1998). A second set of experiments documented the presence of *cis*-spliced introns in the poly-A polymerase gene of two different trypanosomatids. Intriguingly, the same 3′ splice site acted as an acceptor for both *cis*- and *trans*-splicing (Mair et al., 2000).

In vitro selection has proven to be a powerful tool in the examination of *cis*-acting elements that are important for pre-mRNA splicing. Several studies have determined the preferred binding sites for splicing factors that interact near the 3′ splice site, including $U2AF^{65}$, Sex-lethal, and polypyrimidine tract binding protein (Singh et al., 1995) as well as the U2AF heterodimer (Wu et al., 1999). Recent work has also identified selected sequences at the branchpoint in mammals (Lund et al., 2000). In addition, consensus sequences have been derived for the binding of SR proteins, with these elements able to function as splicing enhancers when introduced into appropriate transcripts (Heinrichs & Baker, 1995; Tacke & Manley, 1995; Tacke et al., 1997; Cavaloc et al., 1999). Selection has also been applied to the role(s) of SR proteins in splicing function, as opposed to RNA binding, with splicing of substrates containing a randomized region driven by the presence of an exogenously added, specific, SR protein (Liu et al., 1998, 2000). Open ended (non-SR protein driven) screens have also been performed on pre-mRNA splicing both in vitro (Tian & Kole, 1995; Schaal & Maniatis, 1999b) and in vivo (Coulter et al., 1997) with progressive increases in splicing efficiency being observed. It is notable that although a 5′ splice site downstream has been shown to promote splicing of the upstream intron (see above), 5′ splice site-like sequences were not reported in these splicing activity-based assays. Recently, 5′ splice site-like sequences were found to promote splicing in the yeast *Saccharomyces cerevisiae* (Libri et al., 2000). In this system, a region within the intron that normally interacts via base pairing to promote splicing through secondary structure was randomized. The resulting selected RNAs contained many sequences that resembled 5′ splice sites.

Here, we have employed a selection procedure to identify exonic sequences capable of acting as *trans*-splicing enhancers in vitro. This work was performed in the homologous *trans*-splicing system derived from the nematode *Ascaris lumbricoides* (Hannon et al., 1990). A short region in the 3′ exon of a poor *trans*-splicing substrate was randomized. This initial, heterogeneous pool was incubated under splicing conditions, the *trans*-spliced product was isolated, and the exon was reintroduced into the parental construct. Over the course of seven rounds of selection, the splicing efficiency improved dramatically. Individual clones were isolated and were found to comprise several distinct sequence families. These include both protein-binding motifs as well as 5′ splice site-like elements. Competition experiments have revealed potential differences in the ability of these selected sequences to interfere with *trans*- and *cis*-splicing. One of the identified sequence elements is seen in up to three copies within the short (18 nt) randomized region and crosslinks to discrete proteins in nematode whole-cell extract as well as to purified *A. lumbricoides* SR proteins. A second set of closely related sequences include members that are good matches to 5′ splice site consensus elements. In this case, *trans*-splicing in vitro, which is independent of the need for the 5′ end of U1 (Hannon et al., 1991), can be made to rely on the interaction of U1 with exon sequences downstream of the 3′ acceptor. This may be an endogenous mechanism of *trans*-splicing enhancement utilized in vivo.

RESULTS

Selection of exonic elements capable of enhancing *trans*-splicing in vitro

To monitor the selection of positive acting sequence elements during the course of successive rounds of *trans*-splicing, we needed to begin with an inefficient pre-mRNA substrate. We took advantage of the fact that a well characterized, efficient nematode *cis*-splicing substrate became a poor *trans*-splicing substrate when

its 5′ exon and half of the intron were removed. Within the 3′ exon of this half construct, we engineered a stretch of 18 random nucleotides (18*n*), extending from position +103 to +121 relative to the 3′ splice site (Fig. 1). The position of the randomized region was based on the distance-dependent effects of splicing enhancers on their ability to promote utilization of a 3′ splice site (for example, see Tian & Maniatis, 1994; Graveley et al., 1998). The initial pool, round 0 (R0), of randomized substrates was examined to determine its composition. Twenty-four individual clones were sequenced. The purine and pyrimidine contents were 48% and 52%, respectively. Also, the individual base contents were as follows: A = 13%, C = 21%, G = 35%, and T = 31%. In general, RNA was transcribed from a mixed pool, *trans*-spliced in vitro, reverse transcribed, and PCR amplified using a downstream primer and the 22-nt spliced leader exon as a tag and reintroduced into the parental construct (Fig. 1). Although the introduction of a subcloning step limits the complexity of the subsequent rounds, our aim was to determine whether sequence elements could be selected that would promote *trans*-splicing and not to survey all of the possible variants. As shown in Figure 2, the efficiency of the *trans*-splicing reaction increased dramatically through the use of this technique.

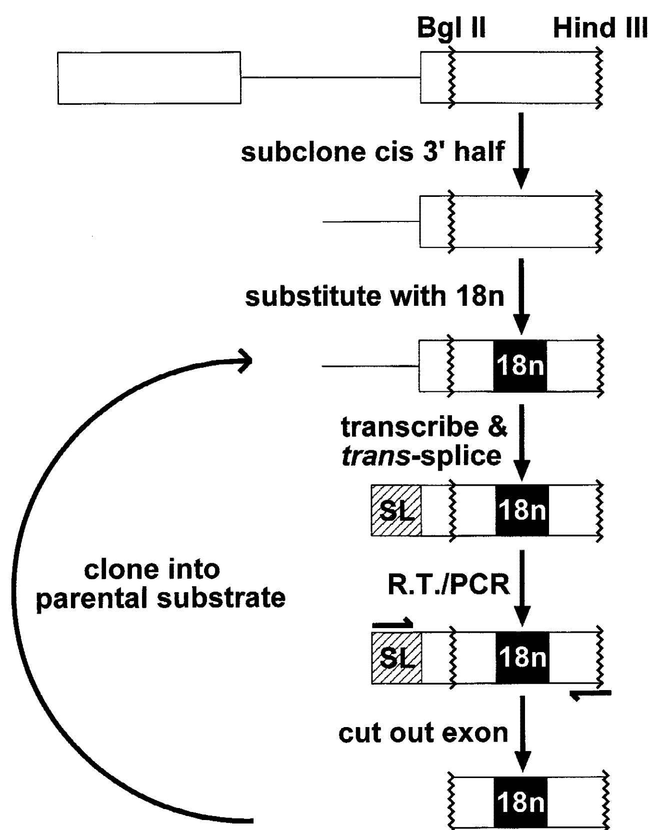


FIGURE 1. Schematic representation of the selection procedure. The parental *cis*-spliced nematode transcript was modified to contain a unique *Bgl*II restriction site. The 3′ half of this substrate was subcloned and served as the starting material for the introduction of the sequence 18*n*. Pools of mixed RNAs were *trans*-spliced, gel purified, and RT-PCR amplified with the exon sequences represented in the *trans*-spliced product reintroduced into the starting construct. Seven rounds of selection were performed.

The parental, nonrandomized, wild-type clone that encompasses the 3′ half of a normally *cis*-spliced substrate (see above) splices with an efficiency of 5% after 90 min (Fig. 2A, lanes 2 and 3). Starting with R0, the turnover immediately began to increase, with the R0 mixed pool of substrates exhibiting 10% splicing efficiency (Fig. 2A, lanes 4 and 5). This effect of increasing splicing efficiency solely through the introduction of a randomized region within a substrate, prior to any selection procedure, has been observed before (Liu et al., 1998, 2000). As the rounds of selection and amplification progressed, there was a steady increase in the fraction of the transcript that underwent *trans*-splicing (see R1 through R7, Fig. 2A, lanes 6–19). The efficiency of an authentic *trans*-splicing substrate is seen in Figure 2A, lanes 20 and 21, for comparison. This set of splicing reactions was performed three times, with the average splicing efficiency shown graphically in Figure 2B. As can be seen from the bar graph, the percent turnover plateaued after R4, within error. To more thoroughly examine the splicing of R0 versus R7, a splicing time course was performed. In this experiment, it is clear that the splicing reaction of the R7 pool exhibits enhancement over the R0 pool at all times examined (Fig. 2C, compare lanes 2–5 with lanes 6–9). In addition, R7 was more efficient at the early, 30-min time point than the authentic *trans*-splicing substrate (Fig. 2C, compare lane 7 with lane 11) as well as at 90 min (Fig. 2C, compare lane 9 with lane 13).

Individual members of the mixed populations of R2, R4, and R7 were isolated and sequenced. Upon examination of these clones, several classes of sequence families were identified. These include four major groupings (Fig. 3A). The first of these is composed of sequences containing the motif SGACS (where S = $^{G}/_{C}$). Interestingly, variants were obtained that include one (clones 2.2, 2.10, and 2.11), two (clones 2.7 and 2.9), and three copies (clones 4.2 and 7.1). Individual SGACS repeats are shown at the bottom of the SGACS panel in Figure 3A. Although the GAC core is invariant, the terminal G and C bases are fairly evenly divided. The next group of sequences include extended stretches that contain good homology to 5′ splice sites. The 5′ end of *A. lumbricoides* U1 snRNA is shown in a proposed base-paired interaction. These stretches of complementarity range from 9/9 matches (clones 4.1 and 7.2) to six contiguous potential base pairs (clone 2.4). Clone 2.5 appeared to represent an interesting problem, with 7 bp to U1 interrupted by a single base bulge. It has previously been reported that *A. lumbricoides* actually contains approximately 10 U1 genes, with at least 2 being expressed (Shambaugh et al., 1994). One of the expressed U1 snRNAs contains a base change

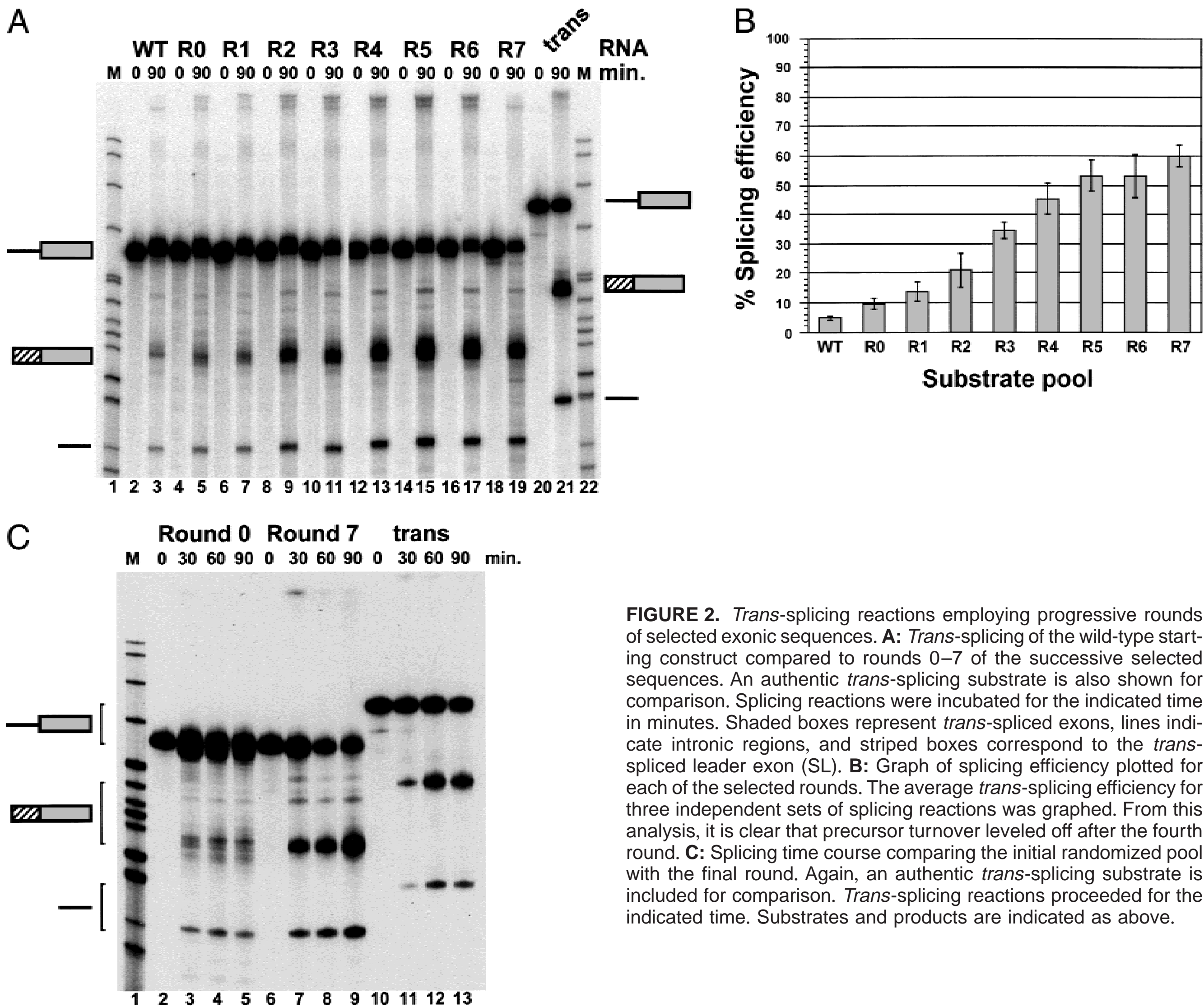


FIGURE 2. *Trans*-splicing reactions employing progressive rounds of selected exonic sequences. **A:** *Trans*-splicing of the wild-type starting construct compared to rounds 0–7 of the successive selected sequences. An authentic *trans*-splicing substrate is also shown for comparison. Splicing reactions were incubated for the indicated time in minutes. Shaded boxes represent *trans*-spliced exons, lines indicate intronic regions, and striped boxes correspond to the *trans*-spliced leader exon (SL). **B:** Graph of splicing efficiency plotted for each of the selected rounds. The average *trans*-splicing efficiency for three independent sets of splicing reactions was graphed. From this analysis, it is clear that precursor turnover leveled off after the fourth round. **C:** Splicing time course comparing the initial randomized pool with the final round. Again, an authentic *trans*-splicing substrate is included for comparison. *Trans*-splicing reactions proceeded for the indicated time. Substrates and products are indicated as above.

at position six that allows this 7-bp stretch containing a bulge to interact perfectly through a run of 8 bp. Other repeated sequence motifs observed include A/C-rich stretches and repeats of the purine-rich element RRGAGS. These last two families appear to correspond to the A/C-rich (Coulter et al., 1997) and purine-rich (for review, see Wang & Manley, 1997) enhancer sequences reported in numerous *cis*-splicing substrates.

We decided to further investigate individual SGACS- and 5′ splice site-like elements. Splicing of three discrete clones, the wild-type starting construct, clone 7.1 (3XSGACS) and 7.2 (5′ ss) was examined over a time course (Fig. 3B). As previously seen at 90 min (Fig. 2A, lane 3), the wild-type construct splices very poorly at all times examined (Fig. 3B, lanes 2–5). By comparison, clones 7.1 and 7.2 both splice extremely well over this span (clone 7.1: Fig. 3B, lanes 6–9; clone 7.2: Fig. 3B, lanes 10–13). These levels of splicing efficiency represent >12-fold splicing enhancement.

The ability of individual exonic sequences to act as competitors in the splicing reaction was examined. The competitors employed included identical length fragments of the 3′ exon, beginning at position +82 relative to the 3′ splice site. The only differences are within the 18 nt region that was the subject of the selection experiments, with the competitors containing the wild-type, 7.1 (3XSGACS), and 7.2 (5′ss) sequence elements. Initially, the effects of these competitors were examined on the splicing of the 7.1 and 7.2 substrates. Both of these *trans*-splicing substrates demonstrate only a modest (1.5 times) reduction in splicing efficiency in the presence of the highest concentration of the wild-type exon competitor used (Fig. 4A (7.1), lanes 4–7; Fig. 4B (7.2), lanes 4–7). Thus, this fragment of the parental substrate does not exhibit a high level of competition for *trans*-splicing on these two RNAs when preincubated with the splicing reaction, prior to the addition of the substrate. The exonic fragment containing clone

A

SGACS Sequences

W.T.: acuggUGUGAAGAUUCGUGAUCUaguua

7.1 : acuggAGGACCGACGCCGGACGAaguua

4.2 : acuggACGAAUCGACCACGACCGaguua

2.2 : acuggAUACCGGWGCAAUGGACGaguua

2.7 : acuggAGCACCGACGAUGCGACGaguua

2.9 : acuggNAUNAAGGACGACGACCGaguua

2.10: acuggACCCCAAAAGUAACCGAaguua

2.11: acuggACCGAGAACGGUACCCCAaguua

7.1:	GGACC	2.2:	GGACG
	CGACG	2.7:	CGACG
	GGACG		CGACG
4.2:	CGACC	2.9:	GGACG
	CGACC		CGACG
	GGACG	2.10:	ggACC
		2.11:	ggACC

cons: SGACS

ACE-like Sequences

2.2 : acuggAUACCGGWGCAAUGGACGaguua

2.4 : acuggUGAGGCGCACAUACACGCaguua

2.6 : acuggGACAAAGGCGAGAUAACGaguua

2.10: acuggACCCCAAAAGUAACCGAaguua

2.11: acuggACCGAGACCGGUACCCCAaguua

2.12: acuggAGUCGUGCGAAGACCAUAaguua

Ψ-5' Splice Site

W.T.: acuggUGUGAAGAUUCGUGAUCUaguua

UCCAUUCAA ApppGm3
7.2 : acuggAACGUACAAGAAAAGGUAaguua

UCCAUUCAA ApppGm3
4.1 : acuggACUGAGGUAGGUUGUAGAaguua

UCCAUUCAA ApppGm3
2.3 : acuggGAGGAACGCGCCGUCGAGaggua agu

UCCAUUCAA ApppGm3
2.5 : acuggUCCCUACGCAGGUUGGUAaguua

UCCAUUCAA ApppGm3
2.4 : acuggUGAGGCGCACAUACACGCaguua

RRGAGS Sequences

2.1 : acuggGAAGAGGGGCCAGCGUUGaguua

2.3 : acuggGAGGAACGGGCCGUCGAGaggua

2.4 : acuggUGAGGCGCACAUACACGCaguua

2.8 : acuggGAGCCAUGAGNAUGCGCAaguua

2.1 :	AAGAGG
2.3 :	ggGAGG
	GAGagg
2.4 :	gUGAGG
2.8 :	ggGAGC

cons: RRGAGS

B

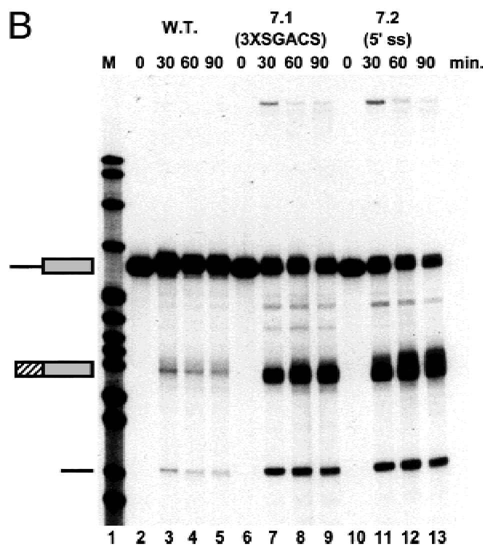


FIGURE 3. Isolation of individual sequences from the mixed pools. **A:** Discrete sequences are grouped according to the elements that they contain. For the SGACS, ACE-like, and RRGAGS elements, individual motifs are underlined. S = G or C; N = A, C, G or U; W = A or U. Upper case letters indicate bases within the 18*n* randomized region, lower case letters indicate surrounding sequences. Outlined letters (clone 2.3) indicate a change in the sequence of the surrounding sequence. Sequences that fall into the 5′ splice site-like category are drawn in a proposed base paired interaction with the 5′ end of *A. lumbricoides* U1 snRNA (Shambaugh et al., 1994) **B:** *Trans*-splicing of individual selected sequences versus the wild-type starting construct. Clone 7.1 contains three copies of the SGACS motif (3XSGACS) and clone 7.2 contains a 9 out of 9 complementarity to the 5′ end of U1 (5′ ss). Splicing reactions were incubated for the indicated times. Precursors and products are indicated as in the legend to Figure 2.

7.1 (3XSGACS) sequences acts as an efficient competitor for both self-competition (Fig. 4A, lanes 8–11) as well as on the splicing of clone 7.2 (5′ss; Fig. 4B, lanes 8–11). Thus, because the SGACS-containing competitor acts as an effective inhibitor of the 5′ splice-containing substrate (7.2), it presumably titrates out a factor(s) required for general *trans*-splicing and not just a factor(s) necessary for the specific *trans*-splicing of substrates containing the SGACS motif. A similar profile was observed when the exon from clone 7.2 (5′ss)

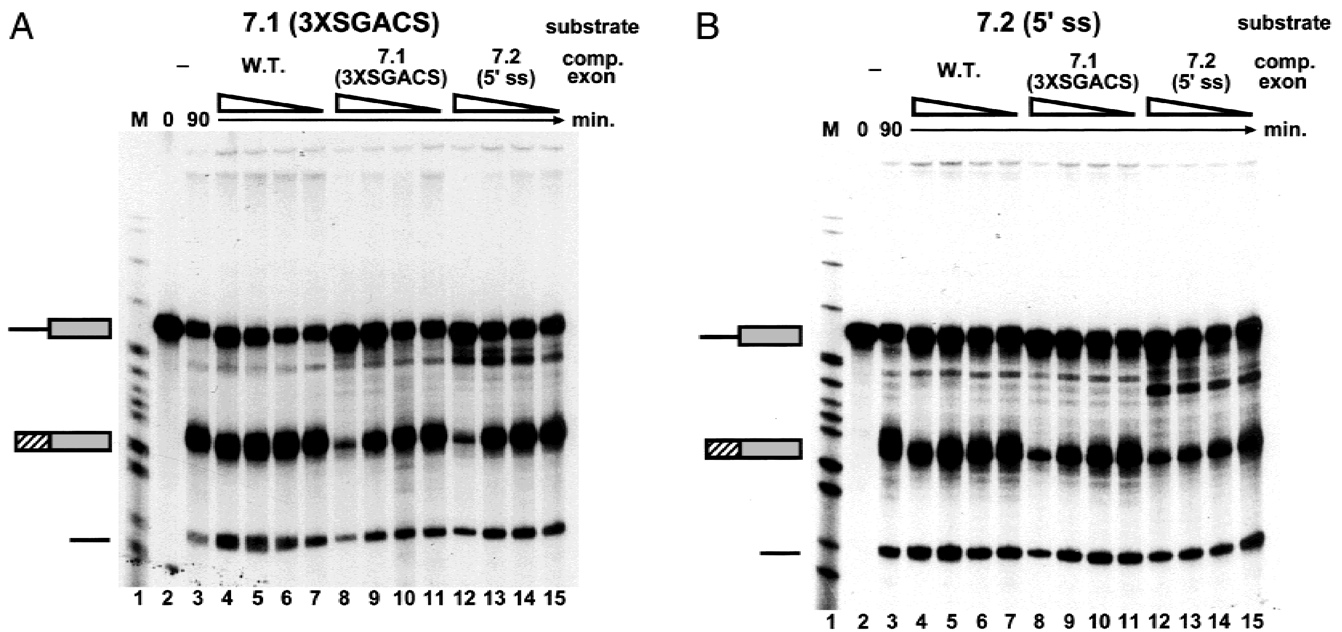
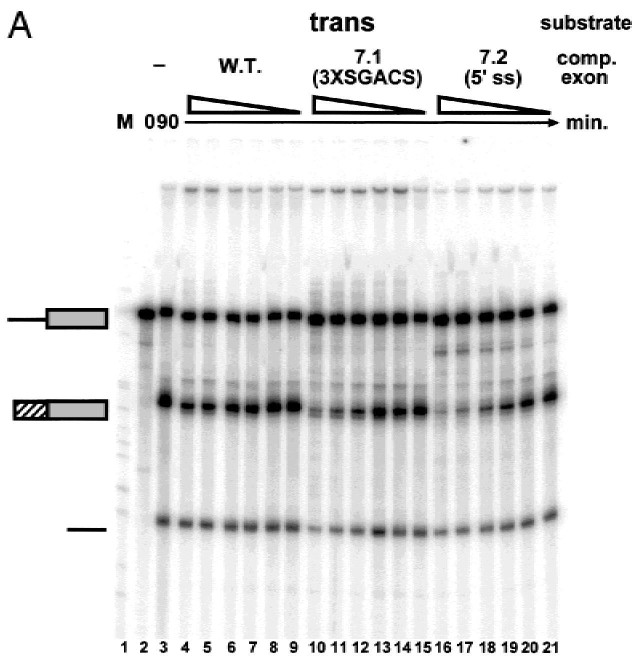


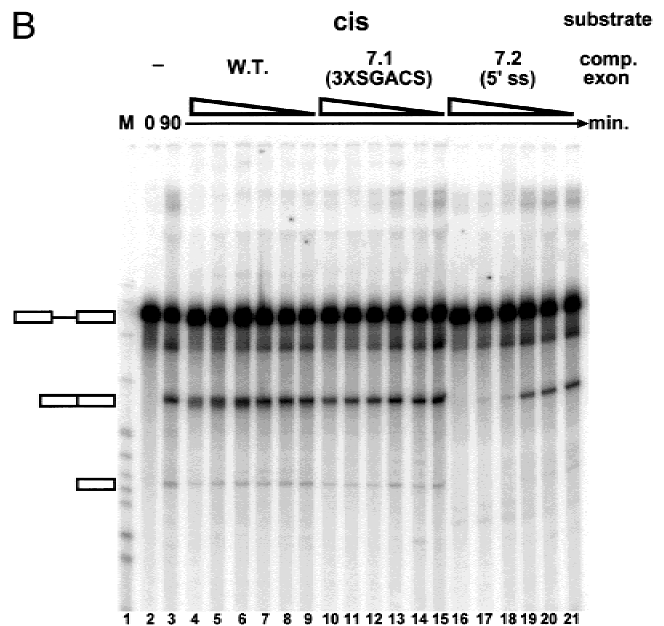
FIGURE 4. Competition experiments performed using selected sequence-containing substrates and competitors. **A:** Splicing of the selected clone 7.1 (3XSGACS) in the presence of either the wild-type, 7.1(3XSGACS) or 7.2(5′ ss) sequence-containing competitor exon fragment. Splicing reactions were preincubated with increasing amounts (0.125 ng, 0.25 ng, 0.5 ng, and 1 μg) of the indicated competitor-containing fragments. Splicing of the 7.1(3XSGACS) substrate in the absence of a competitor is shown for comparison. *Trans*-splicing in the presence of the indicated competitor RNAs proceeded for 90 min. Precursors and products are indicated as in the legend to Figure 2. **B:** Splicing of clone 7.2 (5′ ss) with competition experiments performed as in **A**.

was examined. This competitor led to a substantial reduction in the *trans*-splicing efficiency of both the 7.1 (3XSGACS) (Fig. 4A, lanes 12–15) and self 7.2 (5′ss) substrates (Fig. 4B, lanes 12–15). This result is intriguing because it demonstrates that a 5′ splice site-containing RNA is able to compete for *trans*-splicing, even though the reaction does not require the 5′ end of U1 (Hannon et al., 1991). Therefore, other factors required for *trans*-splicing are presumably titrated away by the presence of excess 5′ splice sites. Possible candidates include p220/Prp8 (Wyatt et al., 1992; Reyes et al., 1996) and U6 snRNP (Sawa & Abelson, 1992; Wassarman & Steitz, 1992). Because the only difference between the three competitor fragments lies within the 18-nt region that was the focus of the selection procedure, the effects on *trans*-splicing efficiency can be directly attributed to this region.

To extend the characterization of the competitor RNAs, we investigated their effects on the splicing of authentic nematode *trans*- and *cis*-splicing substrates. Their effects on a *trans*-splicing substrate, unrelated in sequence to the *cis*-half molecules examined above, were comparable. Both the 7.1 (3XSGACS)- and 7.2 (5′ss)-containing competitors led to similar decreases in *trans*-splicing efficiency (Fig. 5A, lanes 10–15 and lanes 16–21, respectively). Again, the wild-type *cis* exon led to a nominal ~1.5 times reduction in turnover (Fig. 5A, lanes 4–9). Therefore, both SGACS and 5′ splice site sequence elements can serve as general inhibitors of *trans*-splicing. This result demonstrates that factors that interact with the SGACS repeats and with 5′ splice sites are necessary for authentic *trans*-splicing even though the 5′ end of U1 snRNA is not required.

Lastly, we examined the effects of both competitor RNAs on the splicing of the full-length nematode *cis*-splicing construct. Here, an interesting difference with the other substrates was observed. As expected, the 7.2 (5′ss) competitor efficiently competed for *cis*-splicing (Fig. 5B, lanes 16–21) as compared to the wild-type exon (Fig. 5B, lanes 4–9). In addition to other potential factors described above, this competitor would also be expected to titrate the U1 snRNP present in the extract, therefore effecting a large decrease in *cis*-splicing. Surprisingly, when the 7.1 (3XSGACS) competitor was tested, the effect on *cis*-splicing was indistinguishable from the wild-type control (Fig. 5B, lanes 10–15). Each of these RNAs led to a nominal reduction (~1.5 times) in *cis*-splicing efficiency. We have also examined the ability of both the 7.1 (3XSGACS) and 7.2 (5′ss) sequences to act as splicing enhancers in the context of the *cis*-spliced construct incorporating a weakened pyrimidine tract (UUUUUU → UUAAUU). Both of these sequence elements served as strong splicing enhancers in this substrate (L. Boukis & J.P. Bruzik, unpubl. data). Thus, although the SGACS-containing selected sequence was able to promote *cis*-splicing of a weakened substrate, this element, added as a competitor, was not able to compete for *cis*-splicing. This result





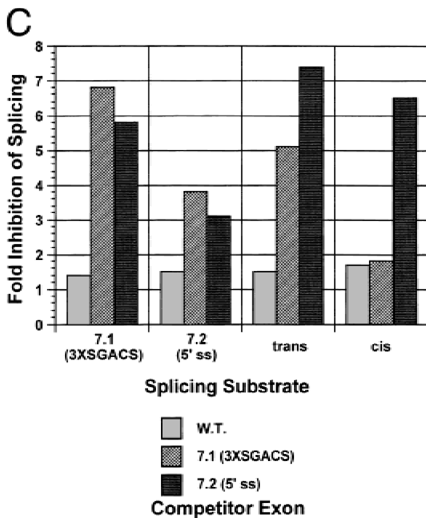


FIGURE 5. Competition experiments performed on authentic *trans*- and *cis*-splicing substrates. **A:** Effects of competitor-containing fragments on an authentic *trans*-splicing substrate. Competitor-containing reactions were performed as in Figure 4 with the following exception. The range of competitor fragment preincubated with the splicing reaction included 0.125 ng, 0.25 ng, 0.5 ng, 1 μg, 1.5 μg, and 2 μg. Substrates and products are as indicated in the legend to Figure 2. **B:** Effects of competitors on a nematode *cis*-splicing substrate. Reactions are as described in **A**. Open boxes indicate *cis*-spliced exons. **C:** Graph of splicing inhibition effected by preincubation of the extract with the indicated competitors. Gray bars represent the splicing inhibition observed at the highest level of wild-type exon competitor, striped boxes the 7.1 (3XSGACS)-containing competitor, and stippled boxes the 7.1 (5′ ss)-containing competitor as assayed on the indicated substrates.

suggests that the factor(s) recognizing the SGACS repeats is essential for the *trans*-splicing reaction, but not for *cis*-splicing of this substrate. The effects of each of the competitor RNAs tested on the four different substrates is depicted graphically in Figure 5C. When presented in this manner, it is clear that each of the selected sequences employed as a competitor exhibit a dramatic decrease in splicing efficiency, except for the SGACS-containing RNA when tested on the *cis*-spliced substrate.

Characterization of selected enhancers containing the SGACS motif

We further investigated the properties of SGACS-containing sequences by comparing the activity of individual clones containing one (Fig. 3A, 2.10), two (2.7), and three (7.1) copies. As seen before, the wild-type sequence *trans*-splices very poorly over a short time course (Fig. 6, lanes 2–5). As the number of SGACS repeats increases from one (Fig. 6, lanes 6–9) to two (Fig. 6, lanes 10–13) and finally to three (Fig. 6, lanes 14–17), the splicing efficiency also increases. At the 90-min time point, the splicing efficiencies are 27.4%, 35.2%, and 45.5% for the substrates containing one, two, and three copies, respectively. The response to increasing the number of SGACS elements is linear at every interval (30, 60, and 90 min) examined. This is in agreement with published work that demonstrated an additive rather than synergistic effect when the number of individual enhancer elements was increased (Hertel & Maniatis, 1998). Therefore, the SGACS-containing sequences increase *trans*-splicing efficiency in a manner that corresponds to the effect of characterized en-

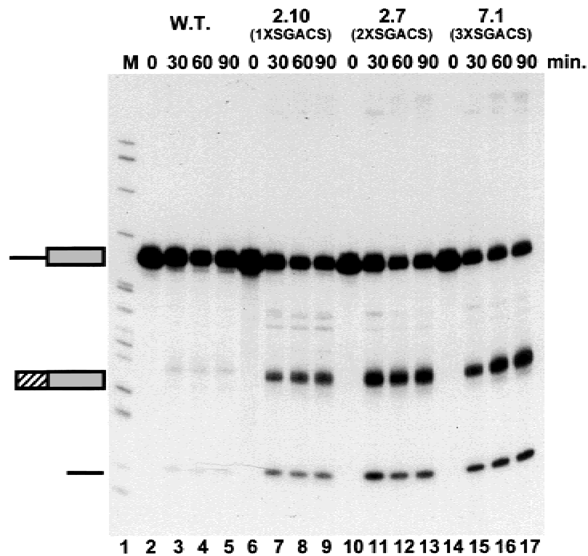


FIGURE 6. Enhancement of *trans*-splicing as a result of increasing the number of SGACS motifs in the substrate. Splicing of the indicated clones was assayed over a time course. Clones 2.10, 2.7, and 7.1 contain one, two, and three copies of SGACS, respectively. Precursors and products are denoted as in Figure 2.

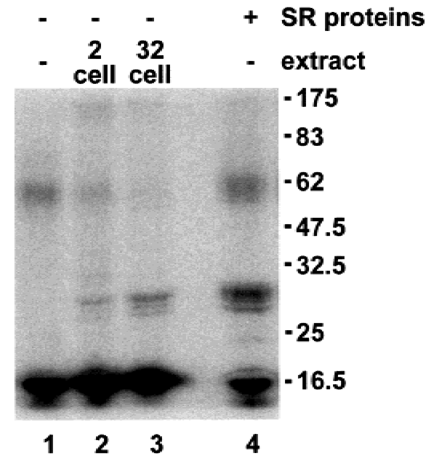


FIGURE 7. Crosslinking of factors to site-specifically labeled SGACS-containing substrate. Crosslinking reactions were performed with identical RNAs in the absence of extract (−), early (presplicing activity) extract (2-cell), active splicing extract prepared from a later developmental stage (32-cell) and in the absence of extract but with purified SR proteins.

hancer elements with respect to the inclusion of multiple copies.

Due to the effectiveness of the SGACS sequence in promoting *trans*-splicing as well as the differences between the ability of the SGACS-containing competitor RNA to inhibit *cis*- and *trans*-splicing of the substrates examined, we wanted to determine what *trans*-acting factor(s) interacted with this element. To look at the SGACS element directly, we employed site-specifically labeled transcripts in a UV crosslinking/RNase protection assay (Wu & Green, 1997). In the absence of the whole-cell extract used in *trans*-splicing assays, only a small amount of undigested or partially digested RNA remains (Fig. 7, lane 1). Upon addition of active, ~32 cell-stage whole-cell extract, a triplet appears at ~30 kDa (Fig. 7, lane 3). This result is specific for the SGACS-containing element, as the reactions were performed with a site-specifically labeled RNA and in addition, only cold RNAs containing the SGACS element act as efficient competitors for the generation of the crosslinked species (data not shown). Since total *A. lumbricoides* SR proteins contain species that migrate at ~30 kDa and are required for in vitro *trans*-splicing (Sanford & Bruzik, 1999b), we also performed crosslinking reactions in the presence of homologous, purified SR proteins. The SR proteins alone (no extract) also generated a profile containing three prominent bands at ~30 kDa (Fig. 7, lane 4). We have previously demonstrated that pre-mRNA splicing activity is developmentally regulated in *A. lumbricoides*, and that this activity correlates with the state of phosphorylation of the embryonic SR proteins (Sanford & Bruzik, 1999a). Early (two-cell stage) extracts contain SR proteins that are inactive, hyperphosphorylated, and predominantly cytoplasmic (J.R. Sanford & J.P. Bruzik, submitted) relative to SR proteins isolated from ~32-cell-stage embryos. When we performed the crosslinking assay with two-cell-stage whole-cell extract, the slowest migrating band of the set of three is most prominent (Fig. 7, lane 2). This is in agreement with the altered mobility of the ~30-kDa SR protein species at this early developmental point as a result of high levels of phosphorylation (Sanford & Bruzik, 1999a). Thus, it appears that the SGACS-containing enhancer elements interact with either a single or multiple SR protein(s) in nematode extracts.

Characterization of selected enhancers containing 5′ splice site-like sequences

The second family of selected sequences that function as effective enhancers in *trans*-splicing are 5′ splice site-like elements (Fig. 3A). To assess the potential role of U1 snRNP in the activation of *trans*-splicing via interaction with a downstream 5′ splice site, we blocked the 5′ end of U1 with a 2′-*O*-methyl oligonucleotide complementary to nucleotides 1–14. Because *trans*-splicing is independent of the need for the 5′ end of U1, this allowed us the opportunity to assess the role of U1 in splicing enhancement, separate from its role in the splicing reaction. Preincubation of the splicing extract with this oligo led to a marked reduction in *cis*-splicing activity (Fig. 8, lanes 14–16) and had no effect on *trans*-splicing (lanes 11–13). *Trans*-splicing of both the weak wild-type substrate (Fig. 8, lanes 2–4) and the strong 7.1 (3XSGACS) clone (lanes 5–7) were unaffected by

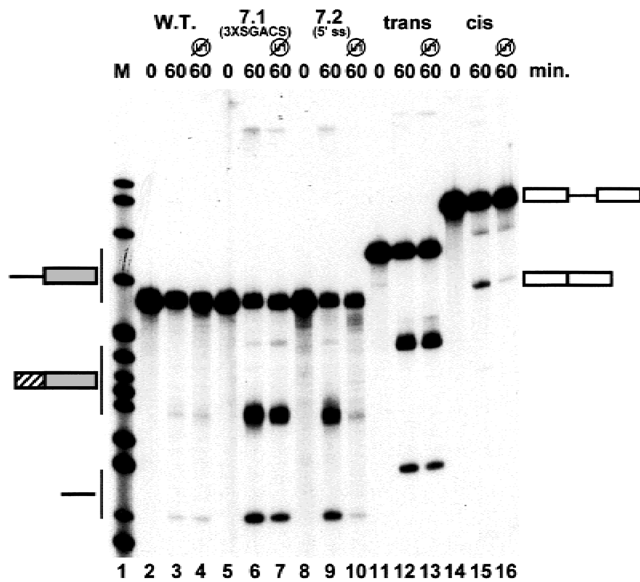


FIGURE 8. Sensitivity of *trans*-splicing to blockage of the 5′ end of U1 snRNA. A 2′-*O*-methyl oligonucleotide targeting the 5′ end of U1 was preincubated in splicing reactions performed on the indicated substrates. Lanes marked by ∅ overwriting U1 indicate the U1 blocked reactions. Splicing reactions were incubated for the indicated times. Precursors and products are denoted as in Figures 2 and 5.

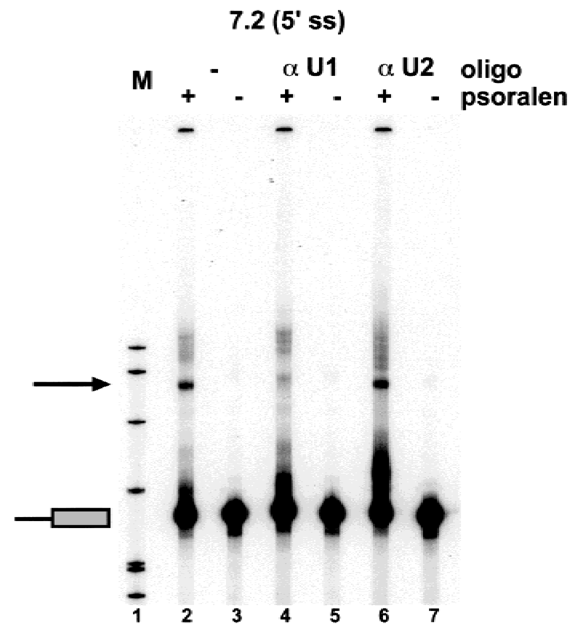


FIGURE 9. Psoralen crosslinking to the 7.2 (5′ss) substrate. Crosslinking reactions were performed in the presence or absence of AMT-psoralen (+ or −). Reactions were preincubated with no oligonucleotide (−), blockage of the 5′ end of U1 (α U1) or blockage of the branchpoint interaction sequence in U2 (α U2). The *trans*-splicing substrate and the crosslinked species (arrow) are indicated.

blocking U1. In contrast, the 7.2 (5′ss) construct was inhibited by this preincubation (Fig. 8, lanes 8–10). This result demonstrates that although binding of U1 is not required for authentic *trans*-splicing, this reaction can be made to rely on the interaction of U1 with the downstream exon.

To directly demonstrate that U1 interacts with the 5′ splice site present in the exon, we performed crosslinking reactions in the presence of AMT psoralen. The 7.2 (5′ss) substrate was also site-specifically labeled with the single ^{32}P present in the middle of the potential 9/9 complementarity. We detected a strong crosslinked species migrating above the precursor that was absolutely dependent upon the addition of AMT psoralen (Fig. 9, lane 2 (+ psoralen) and lane 3 (− psoralen)). Preincubation of the crosslinking reaction with the 2′-*O*-methyl oligonucleotide complementary to the first 14 bases of U1 abolished this band (Fig. 9, lanes 4 and 5, + and − psoralen, respectively). Addition of a control 2′-*O*-methyl oligonucleotide complementary to the branchpoint interaction region of U2 had no effect on this species (Fig. 9, lanes 6 and 7, + and − psoralen, respectively). Control crosslinking reactions employing the 7.1 (3XSGACS) substrate did not generate any U1-dependent species (data not shown). Thus, the 5′ end of U1 snRNP is required for both activation of *trans*-splicing and the appearance of the distinct crosslinked species. Furthermore, we have gel purified the crosslink and targeted it for degradation with RNase H. Oligonucleotides complementary to U1 lead to cleavage of this band, whereas an oligonucleotide complementary to U2 has no effect (data not shown). This set of crosslinking experiments demonstrates that the enhancement of *trans*-splicing that results from the incorporation of a 5′ splice site-like sequence in the downstream exon is correlated with the binding of U1 to that sequence. The other potential factors that interact with 5′ splice sites that allow them to act as effective competitors of *trans*-splicing are required for the reaction. Therefore, U1 snRNP might be a useful factor for regulating *trans*-splicing based on the inclusion of a 5′ splice site downstream of the acceptor. As *trans*-splicing occurs at the 5′ end of the transcript, nematodes usually include internal introns, and thus 5′ splice sites, downstream of the 3′ *trans* acceptor.

DISCUSSION

We have investigated the role of exonic sequences in naturally occurring *trans*-splicing. Through the use of a selection scheme, sequence families were identified that greatly increased *trans*-splicing efficiency over the parental construct. One of these elements (SGACS) can act as an effective competitor of *trans*- but not *cis*-splicing for the substrates examined. In addition, protein(s) in the range of ~30 kDa directly interact with this sequence in homologous splicing-competent extracts as well as with purified SR proteins. Intriguingly, 5′ splice site-like sequences were also isolated multiple times. The activity of a 5′ splice site in enhance-

ment of *trans*-splicing is dependent upon the 5′ end of U1 snRNP. Therefore, although other factors that interact with 5′ splice sites are necessary for *trans*-splicing, the reaction can be made dependent on interaction with U1 snRNP from a position within the downstream exon. This is the first demonstrated role for U1 in *trans*-splicing and highlights the potential for positive interactions between the *cis*- and *trans*-splicing machinery.

Comparison of selected enhancers that promote *trans*-splicing with characterized elements

Many different sequence elements have been described over the past several years that can serve as splicing enhancers. A majority of these act through the interaction of SR proteins with the pre-mRNA. Whether they were identified through SR protein binding assays or via functional selection of splicing activity, it is clear that the initial generalization of splicing enhancers as purine-rich sequences did not reflect the true complexity of active enhancer elements (for a recent compilation of selected sequences, see Tacke & Manley, 1999). The sequence SGACS that we have identified as an active *trans*-splicing enhancer in homologous *A. lumbricoides* extracts has also been identified previously (Tacke & Manley, 1995; Shi et al., 1997; Liu et al., 1998; Bourgeois et al., 1999; Schaal & Maniatis, 1999b). It has been isolated in single clones deduced through binding assays with SF2/ASF, splicing assays in S100 extracts supplemented with SF2/ASF, and also in open-ended screens for active splicing enhancers. In only one of these reports did the sequence emerge enough times to be included in the consensus (Schaal & Maniatis, 1999b). The first of these sets of experiments examined sequences that could interact with histidine-tagged SF2/ASF ΔRS through retention on Ni^{2+} agarose (Tacke & Manley, 1995). Here the consensus sequence (AGGACAGAGC) is close to SGACS and in fact 11 out of 36 sequences in this category contain perfect SGACS motifs. Functional selection, driven by the presence of exogenous SF2/ASF has also yielded a similar consensus (SRSASGA) and as in the previous example, a subset (5 out of 28) of the members of this class contain SGACS elements (Liu et al., 1998). These two reports both implicate SF2/ASF as a candidate for interaction with the sequence SGACS. Two other sets of experiments that noted SGACS elements reported that they responded to SC35 (Schaal & Maniatis, 1999b) or served as a component of an element that interacted strongly with *Drosophila melanogaster* B52 (SRp55; Shi et al., 1997). Finally, a natural example of an element containing an SGACS motif has been implicated in bidirectional activation of both an upstream 3′ splice site and a downstream 5′ splice site in adenovirus E1A pre-mRNA (Bourgeois et al., 1999). Mediation of downstream 5′ splice site activation was shown to correlate with binding of the SR protein 9G8. In our site-specific crosslinking experiments, we observe crosslinks to an SR protein(s) that migrates at ~30 kDa, the same size as SF2/ASF, SC35, and 9G8.

Recent work in our lab has involved an examination of the regulation of splicing of the *fert-1* gene in the nematode *A. lumbricoides*. In this pre-mRNA, a single 3′ splice site can participate in both *trans*- and *cis*-splicing reactions, with apparent developmental regulation (Spicher et al., 1994). This transcript contains three copies of the SGACS element within the first ~100 nt downstream of the common 3′ splice site. Mutation of the core GAC of these elements leads to the inactivation of the 3′ splice site in vitro (J.R. Sanford & J.P. Bruzik, unpubl. results). Thus, the SGACS sequence uncovered via in vitro selection does appear to function as an enhancer element in endogenous, homologous pre-mRNAs in nematodes.

Mechanistically, the basis for the enhancing effect observed with both SGACS-containing and 5′ splice site-like elements has been determined (Romfo et al., this issue). Both types of *trans*-splicing enhancers function to dramatically increase the levels of U2AF and U2 snRNP recruited to the 3′ acceptor complex. Upon blockage of the branchpoint interaction region of U2, a build-up of U2AF is observed. Perhaps most intriguing is the fact that, upon blockage of the 5′ end of U1, the 5′ splice site-like enhancer does not promote U2 or U2AF addition, demonstrating that interaction of U1 with the exon is required for U2AF binding upstream.

Interplay between U1 snRNP and the *trans*-splicing machinery

In addition to the classic role of U1 snRNP in base pairing to 5′ splice sites as an initial event in the *cis*-splicing pathway, other functions have also been ascribed. These include the 5′ splice site-independent interaction between U1 and U2 (Zillmann et al., 1987; Barabino et al., 1990; Fu & Maniatis, 1992). This interplay was further determined to be refractory to either RNase H-targeted removal of the 5′ end of U1 (Fu & Maniatis, 1992) or blockage of this same region with a 2′-*O*-methyl oligonucleotide (Barabino et al., 1990). In fact, participation of U1 snRNP in some aspect of pre-mRNA splicing, other than base pairing to the 5′ splice site, was demonstrated through the use of nonstandard substrates (Seiwert & Steitz, 1993).

Perhaps the most intriguing of the proposed roles of U1 in pre-mRNA splicing relates to the ability of U1 snRNP to promote splicing of upstream introns through exon definition (Robberson et al., 1990). Direct interaction of the 5′ end of U1 snRNA was shown to mediate this effect (Talerico & Berget, 1990; Kreivi et al., 1991; Kuo et al., 1991). In this case, U1 is able to positively effect the splicing of an upstream intron through promotion of complex formation on the 3′ splice

site of the preceding intron (Hoffman & Grabowski, 1992; Chiara & Reed, 1995). It has been suggested that this role for U1 might actually indicate its major function (Cohen et al., 1994; Hwang & Cohen, 1996). In *trans*-splicing, where there is no required base pairing interaction between the 5′ splice site and U1 (Hannon et al., 1991), we have demonstrated that U1 interaction with a sequence downstream of the 3′ *trans* acceptor can positively affect the reaction. Due to the fact that the 5′ splice site in the *trans*-splicing reaction is contained on a single SL RNP in *A. lumbricoides*, there is no possibility of differences in the sequence of the 5′ splice site that might allow for regulation of *trans*-splicing. The inclusion of discrete splicing enhancers or 5′ splice sites downstream of the *trans*-spliced 3′ splice site might allow for some level of *trans*-splicing regulation. This would be especially important in cases where there is a decision between *cis*- and *trans*-splicing as in the case of the *fert-1* gene (Spicher et al., 1994). In addition, the fact that 5′ splice sites normally follow 3′ *trans* acceptors in nematodes may allow for exon definition to actively promote constitutive *trans*-splicing, within distance constraints, in a large number of transcripts.

Complexes formed on 3′ splice sites

In the past several years, the complexity of proteins that assemble on 3′ splice sites has grown with the identification of new protein factors as well as their interactions with both *cis*-acting elements and each other. The region of the pre-mRNA from the branch point to the 3′ splice site has been shown to be important for the interaction of SF1/BBP (Berglund et al., 1997), U2AF^{65} (Ruskin et al., 1988), U2AF^{35} (Merendino et al., 1999; Wu et al., 1999; Zorio & Blumenthal, 1999), PUF60 (Query et al., 1997; Page-McCaw et al., 1999), SRp54 (Query et al., 1997; Page-McCaw et al., 1999; or with py-rich intronic elements upstream of the branch point; Kennedy et al., 1998), and several other proteins (for review, see Reed & Palandjian, 1997). Two pairs of these factors act together in the form of the U2AF heterodimer and a complex of SRp54 and PUF60 (Zamore & Green, 1989; Page-McCaw et al., 1999). Interestingly, poly U-depleted splicing extracts cannot be complemented with U2AF^{65} alone, but in fact require U2AF^{65} as well as PUF60 and SRp54 (Page-McCaw et al., 1999). It has also been suggested that SRp54 might play a large role in intron bridging in organisms that contain short introns (Kennedy et al., 1998). In part, this model incorporates SRp54's unique protein–protein interaction profile, particularly, its interaction with U2AF^{65} and its lack of interactions with U1 70kDa and U2AF^{35} (Zhang & Wu, 1996). Very recently, Blumenthal and colleagues have determined that both complexes, U2AF heterodimer as well as PUF60/SRp54, are necessary components of a 3′ splice site recognition complex through the use of an assay based on nuclear retention of transcripts containing multiple copies of a nematode 3′ splice site (MacMorris et al., 1999, T. Blumenthal, pers. comm.). All of this complexity related to factors at the 3′ splice site has raised the possibility of transcript-specific assembly of early spliceosomal constituents (Kennedy et al., 1998). In fact, if *trans*-splicing is an extreme example of intron definition (Bruzik, 1996), then it is possible that nematodes employ both exon definition, as seen here with the selection of 5′ splice site-like elements, as well as intron definition, the *trans*-splicing reaction itself. In addition, it is also possible that the nature of the enhancing element and its *trans*-acting factors might allow for different types of 3′ splice site complexes to form, thus allowing for selectivity in the types of splicing reactions that ensue as well as their regulation.

MATERIALS AND METHODS

Selection of *trans*-splicing enhancers

The starting *cis*-spliced construct used in the selection procedure (Fig. 1) has been previously described (Hannon et al., 1991). A unique *Bgl*II restriction site was introduced 82 nt downstream of the 3′ splice site through overlap PCR (Ho et al., 1989). A fragment of this *cis*-splicing construct was subcloned into pSP73 (Promega) whose *Bgl*II site was eliminated. The *cis* 3′ half clone (wild-type starting clone) contained 121 nt of the *cis* intron followed by 151 nt of exon sequence. A stretch of 18 random nucleotides was subsequently introduced between position +103 and +121 of the 3′ exon via overlap PCR (Ho et al., 1989). After selection for precursors that allowed for efficient *trans*-splicing, the product of the reaction was gel purified and served as the material for RT/PCR using an antisense 3′ primer including the *Hin*dIII site and a sense primer corresponding to bases 1–20 of the *A. lumbricoides* SL exon (Nilsen et al., 1989). The DNA fragment corresponding to the *trans*-spliced product was then cut with *Bgl*II and *Hin*dIII to generate a 69-bp fragment that was reinserted into the starting wild-type *trans*-splicing substrate. This procedure (*trans*-splicing, gel purification of the product, RT/PCR, subcloning back into the starting material) was performed for a total of seven rounds.

A. lumbricoides whole-cell extracts and body labeled transcripts were prepared as previously described (Hannon et al., 1990). Splicing reactions were performed as described (Hannon et al., 1990), except that the reactions containing mixed populations were scaled up to 25 μL and were supplemented with 50 ng of unlabeled transcript. In the initial splicing reaction, this allowed for approximately five times coverage of the 18*n* possible sequences. The randomized starting construct pool was found to contain 48% purine and 52% pyrimidine (13% A, 21% C, 35% G, 31% T). Competitor RNA exon fragments were 121 nt in length and included sequences from the *Bgl*II site (+82) to the *Hin*dIII site (+151) as well as polylinker sequences. Each of the competitor RNAs was identical except for the region within the 18 nt that were the subject of the selection process. Splicing reactions containing indicated amounts of competitor RNAs were incubated on ice for 30 min followed by the addition of the labeled substrate to

initiate the splicing reaction. The 5′ end of U1 snRNA was blocked in splicing reactions by preincubating a 2′-*O*-methyl oligonucleotide (200 ng) complementary to bases 1–14 with the splicing reaction for 5 min at room temperature prior to the addition of the labeled substrate. All splicing reactions were separated on 5% denaturing PAGE and visualized by autoradiography.

Crosslinking to site-specifically labeled substrates

Site-specifically labeled RNAs clone 7.1 (3XSGACS) or 7.2 (5′ ss) were prepared essentially as described (Maroney et al., 2000). The single ^{32}P label was positioned at the C in the central SGAC*S element (7.1) and in the middle of the 5′ splice site in clone 7.2 (AGGU*AAGUU). RNase H (USB/Pharmacia) cleavage of 40 pmol of transcript was performed with 60 pmol of chimeric 2′-*O*-methyl/deoxy oligonucleotide (Lapham & Crothers, 1996; Lapham et al., 1997). Hybridization of the chimeric oligonucleotide to the target RNA was found to improve cleavage efficiency. Oligo and RNA were incubated at 90 °C for 30 s and slow cooled from 90 °C to 37 °C over 30 min prior to the addition of RNase H. Protein–RNA crosslinking reactions were as described (Wu & Green, 1997). Samples were resolved on 11% SDS-PAGE and visualized by autoradiography. RNA–RNA crosslinking reactions were performed as follows. Splicing reactions (10 μL) were incubated with 30,000 cpm of site-specifically labeled RNA for 5 min on ice and then exposed to 365 nm light for 20 min in the presence or absence of 40 μg/mL AMT psoralen. Reactions were proteinase K treated and phenol/chloroform/isoamyl alcohol (50/49/1) extracted. Samples were separated on a 5% denaturing PAGE and visualized by autoradiography. Blockage with 2′-*O*-methyl oligonucleotides complementary to either the 5′ end of U1 snRNA (nt 1–14) or the branchpoint interaction region of U2 (nt 29–45) was accomplished by preincubating the reaction in the presence of the oligonucleotide for 10 min at room temperature prior to the addition of the labeled substrate.

ACKNOWLEDGMENTS

We thank Tom Blumenthal and Tim Nilsen for communication of results prior to publication, the Nilsen lab for helpful discussions and reagents, and Tom Blumenthal, Tom Maniatis, Tim Nilsen, and members of the Bruzik lab for helpful comments on the manuscript. J.P. Bruzik is supported by National Institutes of Health Grant R01 54204.

Received February 22, 2001; returned for revision March 12, 2001; revised manuscript received March 20, 2001

REFERENCES

Barabino SML, Blencowe BJ, Ryder U, Sproat BS, Lamond A. 1990. Targeted snRNP depletion reveals an additional role for mammalian U1 snRNP in spliceosome assembly. *Cell 63*:293–302.

Berget SM. 1995. Exon recognition in vertebrate splicing. *J Biol Chem 270*:2411–2414.

Berglund JA, Chua K, Abovich N, Reed R, Rosbash M. 1997. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell 89*:781–787.

Blumenthal T, Steward K. 1997. RNA processing and gene structure. In: Riddle DL, Blumenthal T, Meyer BJ, Priess JR, eds. *C. Elegans II*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 117–146.

Bourgeois CF, Popielarz M, Hildwein G, Stévenin J. 1999. Identification of a bidirectional splicing enhancer: Differential involvement of SR proteins in 5′ or 3′ splice site activation. *Mol Cell Biol 19*:7347–7356.

Bruzik JP. 1996. Splicing glue: A role for SR proteins in *trans*-splicing? *Microbial Path 21*:149–155.

Bruzik JP, Van Doren K, Hirsh D, Steitz JA. 1988. *Trans*-splicing involves a novel form of small nuclear ribonucleoprotein particles. *Nature 335*:559–562.

Cavaloc Y, Bourgeois CF, Kister L, Stévenin J. 1999. The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. *RNA 5*:468–483.

Chiara MD, Reed R. 1995. A two-step mechanism for 5′ and 3′ splice-site pairing. *Nature 375*:510–513.

Cohen JB, Snow JE, Spencer SD, Levinson AD. 1994. Suppression of mammalian 5′ splice-site defects by U1 small nuclear RNAs from a distance. *Proc Natl Acad Sci USA 91*:10470–10474.

Côté J, Simard MJ, Chabot B. 1999. An element in the 5′ common exon of the NCAM alternative splicing unit interacts with SR proteins and modulates 5′ splice site selection. *Nucleic Acids Res 27*:2529–2537.

Coulter LR, Landree MA, Cooper TA. 1997. Identification of a new class of exonic splicing enhancers by in vivo selection. *Mol Cell Biol 17*:2143–2150.

Fu X-D, Maniatis T. 1992. The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3′ splice site. *Proc Natl Acad Sci USA 89*:1725–1729.

Graveley BR. 2000. Sorting out the complexity of SR protein functions. *RNA 6*:1197–1211.

Graveley BR, Hertel KJ, Maniatis T. 1998. A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. *EMBO J 17*:6747–6756.

Hannon GJ, Maroney PA, Denker JA, Nilsen TW. 1990. *Trans*-splicing of nematode pre-messenger RNA in vitro. *Cell 61*:1247–1255.

Hannon GJ, Maroney PA, Nilsen TW. 1991. U small nuclear ribonucleoprotein requirements for nematode *cis*- and *trans*-splicing in vitro. *J Biol Chem 266*:22792–22795.

Heinrichs V, Baker BS. 1995. The *Drosophila* SR protein RBP1 contributes to the regulation of *doublesex* alternative splicing by recognizing RBP1 RNA target sequences. *EMBO J 14*:3987–4000.

Hertel KJ, Maniatis T. 1998. The function of multisite splicing enhancers. *Mol Cell 1*:449–455.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene 77*:51–59.

Hoffman BE, Grabowski PJ. 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3′ splice site by a network of interactions spanning the exon. *Genes & Dev 6*:2554–2568.

Humphrey MB, Bryan J, Cooper TA, Berget SM. 1995. A 32-nucleotide exon-splicing enhancer regulates usage of competing 5′ splice sites in a differential internal exon. *Mol Cell Biol 15*:3979–3988.

Hwang DY, Cohen JB. 1996. Base pairing at the 5′ splice site with U1 small nuclear RNA promotes splicing of the upstream intron but may be dispensable for splicing of the downstream intron. *Mol Cell Biol 16*:3012–3022.

Kennedy CF, Krämer A, Berget SM. 1998. A role for SRp54 during intron bridging of small introns with pyrimidine tracts upstream of the branch point. *Mol Cell Biol 18*:5425–5434.

Kreivi JP, Zefrivitz K, Akusjärvi G. 1991. A U1 snRNA binding site improves the efficiency of in vitro pre-mRNA splicing. *Nucleic Acids Res 19*:6956.

Kuo HC, Nasim FH, Grabowski PJ. 1991. Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. *Science 251*:1045–1050.

Lapham J, Crothers DM. 1996. RNase H cleavage for processing of in vitro transcribed RNA for NMR studies and RNA ligation. *RNA 2*:289–296.

Lapham J, Yu Y-T, Shu MD, Steitz JA, Crothers DM. 1997. The position of site-directed cleavage of RNA using RNase H and 2′-*O*-methyl oligonucleotides is dependent on the enzyme source. *RNA 3*:950–951.

Libri D, Lescure A, Rosbash M. 2000. Splicing enhancement in the yeast rp51b intron. *RNA 6*:352–368.

Liu HX, Chew SL, Cartegni L, Zhang MQ, Krainer AR. 2000. Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. *Mol Cell Biol 20*:1063–1071.

Liu HX, Zhang M, Krainer AR. 1998. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes & Dev 12*:1998–2012.

López-Estraño C, Tschudi C, Ullu E. 1998. Exonic sequences in the 5′ untranslated region of alpha-tubulin mRNA modulate *trans*-splicing in *Trypanosoma brucei*. *Mol Cell Biol 18*:4620–4628.

Lou H, Yang Y, Cote GJ, Berget SM, Gagel RF. 1995. An intron enhancer containing a 5′ splice site sequence in the human calcitonin/calcitonin gene-related peptide gene. *Mol Cell Biol 15*:7135–7142.

Lund M, Tange TO, Dyhr-Mikkelsen H, Hansen J, Kjems J. 2000. Characterization of human RNA splice signals by iterative functional selection of splice sites. *RNA 6*:528–544.

MacMorris MA, Zorio DA, Blumenthal T. 1999. An exon that prevents transport of a mature mRNA. *Proc Natl Acad Sci USA 96*:3813–3818.

Mair G, Shi H, Li H, Djikeng A, Aviles HO, Bishop JR, Falcone FH, Gavrilescu C, Montgomery JL, Santori MI, Stern LS, Wang Z, Ullu E, Tschudi C. 2000. A new twist in trypanosome RNA metabolism: *Cis*-splicing of pre-mRNA. *RNA 6*:163–169.

Maroney PA, Romfo CM, Nilsen TW. 2000. Functional recognition of 5′ splice site by U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly. *Mol Cell 6*:317–328.

Maroney PA, Yu Y-T, Jankowska M, Nilsen TW. 1996. Direct analysis of nematode *cis*- and *trans*-spliceosomes: A functional role for U5 snRNA in spliced leader addition *trans*-splicing and the identification of novel Sm snRNPs. *RNA 2*:735–745.

Merendino L, Guth S, Bilbao D, Martinez C, Valcárcel J. 1999. Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3′ splice site AG. *Nature 402*:838–841.

Nilsen TW. 1997. *Trans*-splicing. In: Krainer AR, ed. *Eukaryotic mRNA processing*. Oxford: IRL Press. pp 310–334.

Nilsen TW, Shambaugh J, Denker J, Chubb G, Faser C, Putnam L, Bennett K. 1989. Characterization and expression of a spliced leader RNA in the parasitic nematode *Ascaris lumbricoides* var. *suum*. *Mol Cell Biol 9*:3543–3547.

Page-McCaw PS, Amonlirdviman K, Sharp PA. 1999. PUF60: A novel U2AF65-related splicing activity. *RNA 5*:1548–1560.

Query CC, McCaw PS, Sharp PA. 1997. A minimal spliceosomal complex A recognizes the branch site and polypyrimidine tract. *Mol Cell Biol 17*:2944–2953.

Reed R. 1989. The organization of 3′ splice-site sequences in mammalian introns. *Genes & Dev 3*:2113–2123.

Reed R, Maniatis T. 1986. A role for exon sequences and splice-site proximity in splice-site selection. *Cell 46*:681–690.

Reed R, Palandjian L. 1997. Spliceosome assembly. In: Krainer AR, ed. *Eukaryotic mRNA processing*. Oxford: IRL Press. pp 103–129.

Reyes JL, Kois P, Konforti BB, Konarska MM. 1996. The canonical GU dinucleotide at the 5′ splice site is recognized by p220 of the U5 snRNP within the spliceosome. *RNA 2*:213–225.

Robberson BL, Cote GJ, Berget SM. 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol Cell Biol 10*:84–94.

Ruskin B, Zamore PD, Green MR. 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell 52*:207–219.

Sanford JR, Bruzik JP. 1999a. Developmental regulation of SR protein phosphorylation and activity. *Genes & Dev 13*:1513–1518.

Sanford JR, Bruzik JP. 1999b. SR proteins are required for nematode *trans*-splicing in vitro. *RNA 5*:918–928.

Sawa H, Abelson J. 1992. Evidence for a base-pairing interaction between U6 small nuclear RNA and 5′ splice site during the splicing reaction in yeast. *Proc Natl Acad Sci USA 89*:11269–11273.

Schaal TD, Maniatis T. 1999a. Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA. *Mol Cell Biol 19*:261–273.

Schaal TD, Maniatis T. 1999b. Selection and characterization of pre-mRNA splicing enhancers: Identification of novel SR protein-specific enhancer sequences. *Mol Cell Biol 19*:1705–1719.

Seiwert SD, Steitz JA. 1993. Uncoupling two functions of the U1 small nuclear ribonucleoprotein particle during in vitro splicing. *Mol Cell Biol 13*:3135–3145.

Selvakumar M, Helfman DM. 1999. Exonic splicing enhancers contribute to the use of both 3′ and 5′ splice site usage of rat beta-tropomyosin pre-mRNA. *RNA 5*:378–394.

Shambaugh JD, Hannon GE, Nilsen TW. 1994. The spliceosomal U small nuclear RNAs of *Ascaris lumbricoides*. *Mol Biochem Parasit 64*:349–352.

Shi H, Hoffman BE, Lis JT. 1997. A specific RNA hairpin loop structure binds the RNA recognition motifs of the *Drosophila* SR protein B52. *Mol Cell Biol 17*:2649–2657.

Singh R, Valcárel J, Green MR. 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science 268*:1173–1176.

Somasekhar MB, Mertz JE. 1985. Exon mutations that affect the choice of splice sites used in processing the SV40 late transcripts. *Nucleic Acids Res 13*:5591–5609.

Spicher A, Etter A, Bernard V, Tobler H, Müller F. 1994. Extremely stable transcripts may compensate for the elimination of the gene *fert-1* from all *Ascaris lumbricoides* somatic cells. *Dev Biol 164*:72–86.

Tacke R, Chen Y, Manley JL. 1997. Sequence-specific RNA binding by an SR protein requires RS domain phosphorylation: Creation of an SRp40-specific splicing enhancer. *Proc Natl Acad Sci USA 94*:1148–1153.

Tacke R, Manley JL. 1995. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *EMBO J 14*:3540–3551.

Tacke R, Manley JL. 1999. Determinants of SR protein specificity. *Curr Opin Cell Biol 11*:358–362.

Talerico M, Berget SM. 1990. Effect of 5′ splice site mutations on splicing of the preceding intron. *Mol Cell Biol 10*:6299–6305.

Thomas JD, Conrad RC, Blumenthal T. 1988. The *C. elegans trans*-spliced leader RNA is bound to Sm and has a trimethylguanosine cap. *Cell 54*:533–539.

Tian H, Kole R. 1995. Selection of novel exon recognition elements from a pool of random sequences. *Mol Cell Biol 15*:6291–6298.

Tian M, Maniatis T. 1994. A splicing enhancer exhibits both constitutive and regulated activities. *Genes & Dev 8*:1703–1712.

Wang J, Manley JL. 1997. Regulation of pre-mRNA splicing in metazoa. *Curr Opin Genet Dev 7*:205–211.

Wang Y-C, Selvakumar M, Helfman DM. 1997. Alternative pre-mRNA splicing. In: Krainer AR, ed. *Eukaryotic mRNA processing*. New York: Oxford University Press. pp 242–279.

Wassarman DA, Steitz JA. 1992. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science 257*:1918–1925.

Wu Q, Krainer AR. 1996. U1-mediated exon definition interactions between AT-AC and GT-AG introns. *Science 274*:1005–1008.

Wu S, Green MR. 1997. Identification of a human protein that recognizes the 3′ splice site during the second step of pre-mRNA splicing. *EMBO J 16*:4421–4432.

Wu S, Romfo CM, Nilsen TW, Green MR. 1999. Functional recognition of the 3′ splice site AG by the splicing factor $U2AF^{35}$. *Nature 402*:832–835.

Wyatt JR, Sontheimer EJ, Steitz JA. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5′ splice site before the first step of pre-mRNA splicing. *Genes & Dev 6*:2542–2553.

Zamore PD, Green MR. 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc Natl Acad Sci USA 86*:9243–9247.

Zhang W-J, Wu JY. 1996. Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. *Mol Cell Biol 16*:5400–5408.

Zillmann M, Rose SD, Berget SM. 1987. U1 small nuclear ribonucleoproteins are required early during spliceosome assembly. *Mol Cell Biol 7*:2877–2883.

Zorio DA, Blumenthal T. 1999. Both subunits of U2AF recognize the 3′ splice site in *Caenorhabditis elegans*. *Nature 402*:835–838.