

## Identification of a New Class of Exonic Splicing Enhancers by In Vivo Selection

LYDIA R. COULTER, MARK A. LANDREE, AND THOMAS A. COOPER\*

*Department of Pathology and Program in Cell and Molecular Biology,  
Baylor College of Medicine, Houston, Texas 77030*

Received 21 November 1996/Returned for modification 30 December 1996/Accepted 23 January 1997

**In vitro selection strategies have typically been used to identify a preferred ligand, usually an RNA, for an identified protein. Ideally, one would like to know RNA consensus sequences preferred in vivo for as-yet-unidentified factors. The ability to select RNA-processing signals would be particularly beneficial in the analysis of exon enhancer sequences that function in exon recognition during pre-mRNA splicing. Exon enhancers represent a class of potentially ubiquitous RNA-processing signals whose actual prevalence is unknown. To establish an approach for in vivo selection, we developed an iterative scheme to select for exon sequences that enhance exon inclusion. This approach is modeled on the in vitro SELEX procedure and uses transient transfection in an iterative procedure to enrich RNA-processing signals in cultured vertebrate cells. Two predominant sequence motifs were enriched after three rounds of selection: a purine-rich motif that resembles previously identified splicing enhancers and a class of A/C-rich splicing enhancers (ACEs). Individual selected ACEs enhanced splicing in vivo and in vitro. ACE splicing activity was competed by RNAs containing the purine-rich splicing enhancer from cardiac troponin T exon 5. Thus, ACE activity is likely to require a subset of the SR splicing factors previously shown to mediate activity of this purine-rich enhancer. ACE motifs are found in two vertebrate exons previously demonstrated to contain splicing enhancer activity as well as in the well-characterized *Drosophila doublesex* (*dsx*) splicing enhancer. We demonstrate that one copy of the *dsx* repeat enhances splicing of a vertebrate exon in vertebrate cells and that this enhancer activity requires the ACE motif. We suggest the possibility that the *dsx* enhancer is a member of a previously unrecognized family of ACEs.**

Most metazoan pre-mRNAs contain multiple exons that are precisely joined during splicing. Remarkable progress toward understanding the biochemistry of intron removal has been made (30); however, many questions regarding the initial events of spliceosome assembly remain. Conserved sequences at the intron/exon borders (splice sites) are required for splicing, but these sequences do not contain sufficient information to distinguish bona fide splice sites from cryptic splice sites of similar sequence found in introns and exons. How the authentic splice sites are initially recognized, how cryptic splice sites are ignored, and how exons are contiguously joined without exon skipping remain essentially unknown (reviewed in reference 3).

Recent evidence suggests that exons are defined early in the splicing reaction by factors that bind to the splice sites and communicate across the exon (2). Cooperative interactions of a large number of factors that bind to multiple *cis* elements may compensate for the lack of information in individual splicing elements (33). Some exons are defined by auxiliary elements in addition to the splice sites. For example, intronic elements have been found to be associated with exons that are unusually small and/or alternatively spliced (4, 6, 9, 19, 25, 35, 37). Some of these elements are associated with cell-specific splicing, while others appear to be ubiquitously recognized. Auxiliary splicing elements have also been identified within exons (1, 5, 11, 17, 20, 22, 36, 38, 40, 46, 48, 50, 51). Most of these exonic splicing enhancers are purine rich; however, non-purine-rich enhancers have also been described (15, 39, 46, 47).

Activation of purine-rich exonic splicing enhancers requires binding of SR proteins, a conserved family of essential splicing factors (22, 32, 39, 40). Differences in substrate specificity among SR protein family members have been revealed by SELEX analysis and have been demonstrated directly by RNA binding and in vitro splicing assays (reviewed in references 16 and 29). For example, we demonstrated that a subset of SR proteins bound and activated the purine-rich splicing enhancer in cardiac troponin T (cTNT) exon 5 (32). Protein-protein interactions between the Arg-Ser domains of SR proteins and other splicing factors are likely to contribute to the cooperative assembly of early splicing complexes (2, 29, 33).

One of the best-characterized exonic splicing enhancers is in the fourth exon of the *Drosophila doublesex* (*dsx*) gene (reviewed in reference 34). This enhancer contains six copies of a 13-bp sequence (the *dsx* repeat) and a purine-rich element located between the fifth and sixth repeats. In contrast to the splicing enhancers identified in vertebrate exons which appear to be ubiquitously recognized, the *dsx* enhancer promotes female-specific splicing of exon 4 (34). RNA-protein and protein-protein interactions between the *dsx* repeats and two *Drosophila* SR-like proteins, Tra2 and Tra, recruit essential SR proteins into a complex that promotes *dsx* enhancer activity in HeLa cell nuclear extracts (18, 26, 27, 44, 45).

The prevalence of exonic splicing enhancers is unknown. Most enhancers have been found in exons that are associated with an alternative splicing event, within either an alternative exon or an adjacent constitutive exon. The association of enhancers with alternative splicing may reflect a mechanistic link: alternative splicing generally requires weak splice sites to allow regulated usage, and enhancers would ensure that these signals are not completely ignored. Alternatively, this association may reflect the bias towards investigations of alternatively spliced

\* Corresponding author. Phone: (713) 798-3141. Fax: (713) 798-5838. E-mail: tcooper@bcm.tmc.edu.

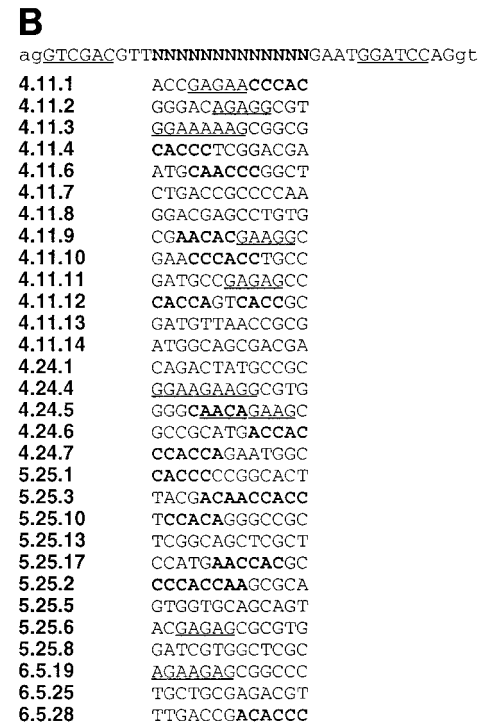
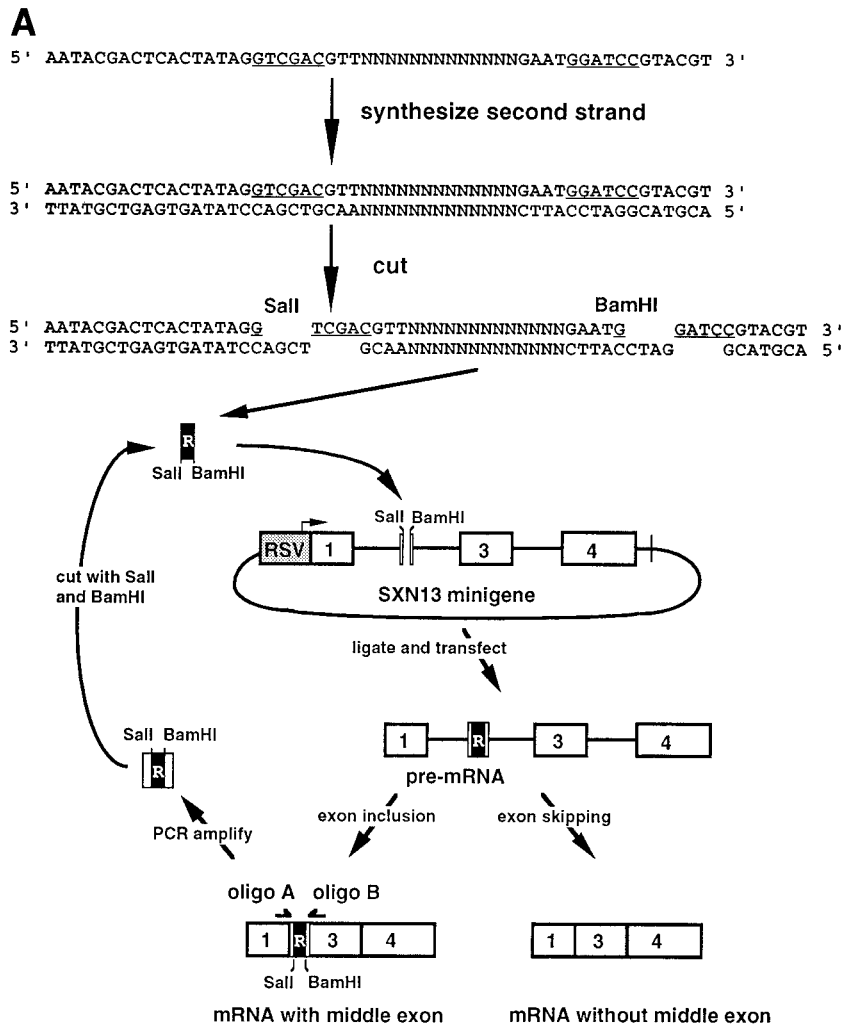


FIG. 1. In vivo selection for splicing enhancers. (A) See Materials and Methods for details of the selection procedure. (B) The sequence of the SXN13 middle exon is shown at the top of the column. The intron sequence is in lowercase, and the exon sequence is in uppercase. All sequenced third-round exons are shown at the bottom. In the selected sequences, purine-rich motifs are underlined and ACEs are in boldface (these motifs are defined in Results). To avoid selecting against premature stop codons which can result in exon skipping (10), the natural globin translation start codon was removed from the SXN13 minigene. Selected sequences from all three rounds contain equivalent numbers of stop codons in all three reading frames, indicating that there was no selection against stop codons. In both panels, randomized positions are indicated by the letter N and the *Sall* and *Bam*HI restriction sites are underlined.

exons. It is possible that splicing enhancers are common among both constitutive and alternative exons.

Tian and Kole used an in vitro selection scheme to enrich for sequences that enhance inclusion of the middle exon of a two-intron, three-exon transcript in HeLa cell nuclear extracts (42). Most of the enriched pre-mRNAs contained purine-rich tracts. However, exons that lacked purine-rich tracts were also enriched, supporting the contention that many different sequence motifs may serve as splicing enhancers.

Here we describe an iterative procedure to select for exon sequences that enhance inclusion of an alternative exon during splicing in vivo. We used this approach to identify a novel class of A/C-rich exonic splicing enhancers (ACEs). Splicing enhancers were selected from a 13-nucleotide randomized cassette inserted into an alternative exon that is predominantly skipped in the absence of a splicing enhancer. Exons that were spliced into the mRNA were selectively amplified by reverse transcription (RT)-PCR; repeated cycles of ligation, transfection, and amplification were used to enrich for sequences that enhanced exon inclusion. After three rounds of selection, two sequence motifs predominated: purine-rich motifs resembling the previously defined enhancers and an A/C-rich motif. Exons that were neither purine rich nor A/C rich were also isolated,

indicating that a variety of sequences can enhance splicing in vivo.

## MATERIALS AND METHODS

**Selection procedure.** One hundred nanograms of oligonucleotide containing 13 randomized positions (Fig. 1A, top) was made double stranded and amplified by PCR with Vent DNA polymerase and oligonucleotides that flanked the randomized region (GACGTACGGATCCATTC and TAATACGACTCACTA TA). The PCR product was phenol-chloroform extracted and digested with *Sall* and *Bam*HI. The 24-nucleotide *Sall*/*Bam*HI fragment was isolated from a 6% nondenaturing polyacrylamide gel and quantitated. Ligation reaction mixtures (100  $\mu$ l each) contained 1  $\mu$ g of gel-isolated minigene with an equal molar amount of insert. Ligation reaction mixtures were transfected directly into QT35 quail fibroblast cultures (two 60-mm-diameter plates containing  $10^6$  cells) by a calcium phosphate protocol as described previously (50). QT35 cells were chosen for their high transfection efficiency. Total RNA was harvested 40 to 48 h following transfection with guanidine thiocyanate (49) and treated with DNase. RNAs containing the alternative exon were selectively amplified by RT-PCR with oligonucleotides A (GTGGTGAGGCCCTGGGCAGGTC) and B (AAGGGTAGACCACCAGCAGCCTGGA) (Fig. 1A), which prime within the alternative exon (nucleotides complementary to the alternative exon are underlined). The PCR product was phenol-chloroform extracted and processed through two more rounds. Conditions for ligation and preparation of vector and insert were optimized by using a bacterial transformation assay.

**Minigene constructs, transient transfection, and RNA analysis.** The SXN13 selection minigene contains a 34-nucleotide alternative exon flanked by duplicated intron 1 from human  $\beta$ -globin such that the first and third exons of the

minigene are globin exons 1 and 2. The 5' splice site of the middle exon is the noncanonical natural sequence of globin intron 1 (CAG/GTTGGT). The fragment containing exons 1 to 3 of this minigene was constructed by using a PCR approach to a design described by Dominski and Kole (13). This construct was previously used to demonstrate splicing enhancer activity of cTNT exon 5 (50). The level of alternative exon inclusion can be modulated by changing exon size (13, 50). The 34-nucleotide exon shows a low but detectable level of enhancer-independent inclusion in QT35 cells. We chose to use this exon by reasoning that more sequences would have the capacity to boost inclusion of a weakly recognized exon than to activate inclusion of an exon that is completely skipped. Minigene intron 3 and exon 4 are derived from cTNT intron 6 and exon 18 (details are available upon request). The cTNT minigene (see Fig. 4) was derived from  $\Delta$ PB (50) by replacing natural exon 4 sequence (to remove a previously introduced *Bam*HI site). Exon 5 was modified as shown in Fig. 4 to accommodate the *Sal*I/*Bam*HI cassette containing the randomized region. Transient transfections of cloned minigenes were done with 5  $\mu$ g of plasmid DNA per 60-mm plate of  $10^6$  cells. Total cellular RNA from half of a plate of cells was assayed by primer extension with 2 ng of  $^{32}$ P-, end-labeled oligonucleotides (50) (AGAACCTCTGGGTCCAAGGGTAG is complementary to globin exon 2 and GTCTCCTCTTCCTCCTCGTCTACCTGATCC is complementary to cTNT exon 6). The products of primer extension were quantitated directly from the gel with a Betagen Betascope 603 analyzer. The percent exon inclusion was calculated as follows: [cpm exon inclusion/(cpm exon inclusion + cpm exon skipping)]  $\times$  100. Except for some of the nonselected clones in Fig. 2, results for all minigenes were confirmed by at least two transfections.

**In vitro RNA assays.** The in vitro splicing substrates in Fig. 5 were constructed by an *Eco*RI/*Nde*I collapse of the transfection plasmids in Fig. 4 that brought exon 4 just downstream of the T3 polymerase promoter in the pBluescript plasmid. *Eco*RI is located 25 nucleotides downstream of the T3 promoter, and *Nde*I is located 8 nucleotides upstream from exon 4. These RNAs contain (5' to 3') 28 nucleotides of polylinker, 8 nucleotides of cTNT intron 3, the 18-nucleotide cTNT exon 4, cTNT intron 4, and the selection exon shown in Fig. 4. Uniformly labeled splicing substrates were synthesized in 20- $\mu$ l reaction mixtures containing 500 ng of DNA (linearized with *Bam*HI), 40 mM Tris (pH 7.5), 6 mM MgCl<sub>2</sub>, 2 mM spermidine, 50 mM NaCl, 10 mM dithiothreitol, 0.5 mM each GTP, ATP, CTP, and UTP, 1.25 mM GpppG, 20 U of RNasin, 25  $\mu$ Ci of [ $\alpha$ - $^{32}$ P]GTP, and 20 U of T3 RNA polymerase. All RNAs were gel isolated. Splicing reaction mixtures (10  $\mu$ l) contained 40,000 cpm of substrate (approximately 0.01 pmol), 0.8% polyethylene glycol, 1 mM MgCl<sub>2</sub>, 0.625 mM ATP, 25 mM creatine phosphate, and 40% HeLa cell nuclear extract (10 to 12  $\mu$ g/ $\mu$ l of protein). Reactions were stopped by addition of 90  $\mu$ l of urea solution (7 M urea, 0.35 M NaCl, 0.01 M Tris [pH 7.4], 0.01 M EDTA, 0.05 g of sodium dodecyl sulfate per ml) and 90  $\mu$ l of water, and the products were phenol-chloroform extracted and then run on a 4% denaturing polyacrylamide gel.

Competitor RNAs in Fig. 6 contain (5' to 3') 20 nucleotides of vector, the last 50 nucleotides of cTNT intron 4, and the selected exon to the *Bam*HI site (Fig. 4). UP, WT, and DOWN competitor RNAs containing cTNT exon 5 enhancer mutants were described previously (32). Competitor RNAs were synthesized in 100- $\mu$ l reaction mixtures containing 6  $\mu$ g of linearized DNA, 40 mM Tris (pH 7.5), 20 mM MgCl<sub>2</sub>, 2 mM spermidine, 50 mM NaCl, 10 mM dithiothreitol, 4 mM each GTP, ATP, CTP, and UTP, 40 U of RNasin, 2  $\mu$ Ci of [ $\alpha$ - $^{32}$ P]GTP (for quantitation), and 60 U of T7 RNA polymerase. Competitor RNAs were gel isolated and quantitated by either scintillation counting or UV absorbance determination.

## RESULTS

The selection scheme is diagrammed in Fig. 1A. A synthetic DNA cassette containing 13 contiguous randomized positions (approximately  $6.7 \times 10^7$  sequences) was directionally ligated into the middle exon of a three-exon minigene derived from human  $\beta$ -globin. Ligation reaction mixtures were transfected directly into QT35 quail fibroblast cultures. In the absence of a splicing enhancer, the middle exon is predominantly skipped (50) due to its small size (34 nucleotides) and a noncanonical 5' splice site (14). We have shown previously that inclusion of this exon is induced by a purine-rich splicing enhancer (50). Following transfection, RNAs that contained the middle exon were selectively amplified by RT-PCR with oligonucleotides complementary to the junctions of exons 1 and 2 (oligonucleotide A [Fig. 1A]) and exons 2 and 3 (oligonucleotide B). To enrich for the strongest splicing enhancers, the selected pool of randomized cassettes was excised from the RT-PCR product and cycled through two additional rounds of ligation, transfection, and amplification.

Two predominant sequence motifs were enriched by the

TABLE 1. Enrichment for ACEs and/or purine-rich sequences during selection for exonic enhancers

Round of selection	% of sequenced clones containing:		n
	ACE	ACE and/or PRE <sup>a</sup>	
None (nonselected) <sup>b</sup>	16	23	31
First	25	39	28
Second	34	47	32
Third	53	73	30

<sup>a</sup> PRE, purine-rich element.

<sup>b</sup> To obtain nonselected exons, ligation mixtures containing the *Sal*I/*Bam*HI randomized cassette and SXN13 minigene vector were transformed into bacteria and colonies were picked at random for sequencing.

selection (Fig. 1B and Table 1). One motif resembles the previously characterized purine-rich splicing enhancers, thereby validating the in vivo selection approach. The second motif is an ACE. Sequences rich in A/C and G/A were clearly enriched; however, no consensus for either motif was discernible. Therefore, the ACE and purine-rich sequence are defined as at least five consecutive A+C or G+A motifs, respectively. To exclude homopolymers, which are unlikely to have enhancer activity (41), the motif must contain both nucleotides at a ratio of  $\geq 1:5$ . This criterion was based on the length and sequence requirements for activity of natural and synthetic purine-rich motifs (12, 32, 41). It should be noted that a 5-nucleotide purine-rich sequence contributes significantly to the enhancer in bovine growth hormone exon 5 (12).

After three rounds of selection, 16 of 30 (53%) of selected exons contained an ACE (Fig. 1B), compared to 5 of 31 (16%) of randomly picked nonselected exons (data not shown). Strikingly, 22 of the 30 (73%) third-round exons that were sequenced contained either an ACE, a purine-rich motif, or both motifs.

To test directly whether the selected sequences functioned as splicing enhancers, individual clones containing selected or nonselected exons were picked at random and transiently transfected into QT35 cultures and the level of exon inclusion was assayed by primer extension. The selected exons showed a significantly higher level of exon inclusion than nonselected exons (representative results are in Fig. 2). The median level of exon inclusion for selected exons was 39% ( $n = 25$ ) compared to 9% ( $n = 20$ ) for exons containing nonselected sequences. We conclude that the selection procedure enriched for sequences with splicing enhancer activity.

The enhancer activities of three individual selected sequences are shown in Fig. 3. These clones, which contain an ACE (4.11.12 [lane 3]), a purine-rich motif (4.24.4 [lane 5]), or both motifs (4.11.1 [lane 1]) show higher levels of exon inclusion than individual nonselected exons (Fig. 2). In addition, exon inclusion is enhanced in these clones above a baseline level of 22%, determined by transfecting a plasmid pool from >400 bacterial colonies containing nonselected exons (Fig. 3, lane 7). To directly determine whether the ACE was required for enhanced exon inclusion, we introduced point mutations into two selected clones. Mutations which reduced the A/C content of 4.11.1 (lane 2) and 4.11.12 (lane 4) each decreased the level of exon inclusion compared to that of the selected sequence. A synthetic ACE sequence also enhanced exon inclusion, further supporting a role for this motif as a splicing enhancer (lane 6).

To test the enhancer activity in a different minigene, selected purine-rich and ACE motifs were inserted in place of the purine-rich enhancer in exon 5 of the cTNT gene. Splicing of

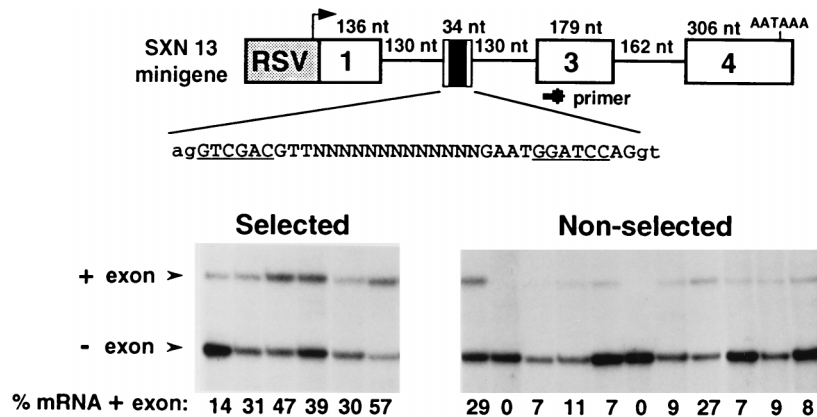


FIG. 2. Exons selected in the SXN13 minigene enhanced exon inclusion compared to nonselected exons. Individual clones were transfected into QT35 cultures and assayed by primer extension. The selected clones are, from left to right, 5.25.10, 5.25.13, 5.25.17, 6.5.19, 5.25.5, and 5.25.8. All are from the third round, and their sequences are shown in Fig. 1B. Nonselected clones were obtained as described in Results. The sequences of the nonselected exons, from left to right, are CATCCGAGCACCT, AGCAAAGGGGAAG, ACCCAGTAAGGA, ACCGGTAGCTGT, GCGTAATGAGACT, AAAACAGGTCGAG, TTCGATGTCTGG, TAAGCCAGCAGAG, AAGGCCAAAAGAT, GTCATCGATTAGA, and CGTGAATGAGACT. The primer extension products are 175 and 209 nucleotides for exon skipping and inclusion, respectively. The *Sall* and *Bam*HI restriction sites are underlined. RSV, Rous sarcoma virus; nt nucleotides.

cTNT exon 5 requires an exonic enhancer, making this a good context in which to test selected sequences (7, 50). All three of the selected exons shown in Fig. 3 as well as the synthetic ACE enhanced splicing of the cTNT exon compared to pooled nonselected exons (Fig. 4). Three independent mutations in the ACE decreased splicing of 4.11.12, demonstrating a sequence-specific requirement of the ACE for enhancer activity. We conclude that splicing enhancer activity of the selected motifs is independent of the minigene used for selection.

We next tested whether enhancer activities of selected exons were reproduced in vitro. A single intron splicing substrate was derived from the cTNT minigene (shown in Fig. 4) such that the downstream exon contained the selection cassette (Fig. 5). The results presented in Fig. 5 demonstrate that the selected motifs enhanced removal of the upstream intron. Mutations that disrupted ACE activity in vivo decreased the level of splicing to that observed for the nonselected sequence. As described previously for the cTNT purine-rich enhancer (50), mutations in the ACE did not affect splicing of the downstream intron (data not shown). Our data suggest that the ACE promotes exon inclusion by enhanced splicing of the upstream intron and not the downstream intron. Although a role for differences in stability in determining mRNA ratios in vivo cannot be ruled out, these in vitro results directly demonstrate that the selected enhancers affect splicing.

To determine whether the ACE enhanced splicing via direct interactions with titratable *trans*-acting factors, in vitro splicing of the 4.11.12 ACE substrate (shown in Fig. 5) was challenged with cold competitor RNAs containing the 4.11.12 downstream exon linked to the branch site and 3' splice site of intron 4. RNAs containing either the 4.11.12 or the synthetic ACE exon efficiently competed splicing (Fig. 6 and data not shown). Competitors containing 4.11.12 mutations (12 mu1 and 12 mu2) that inactivated splicing in vivo did not compete splicing, demonstrating that competition is ACE dependent. The 4.11.12 competitor RNA did not compete in vitro splicing of an enhancer-independent splicing substrate (cTNT exons 8 and 9) and therefore is not a general inhibitor of splicing (data not shown).

To determine whether the ACE and purine-rich enhancer require the same or different enhancer-binding factors for their activities, splicing of 4.11.12 was challenged with RNAs con-

taining the cTNT exon 5 enhancer and two mutants that increase or decrease exon 5 enhancer activity in vivo. These RNAs were used previously to demonstrate that the strength of the purine-rich enhancer directly correlates with binding affinity for four members of the SR protein family of essential splicing factors, SRp30a, SRp40, SRp55, and SRp75 (32). The

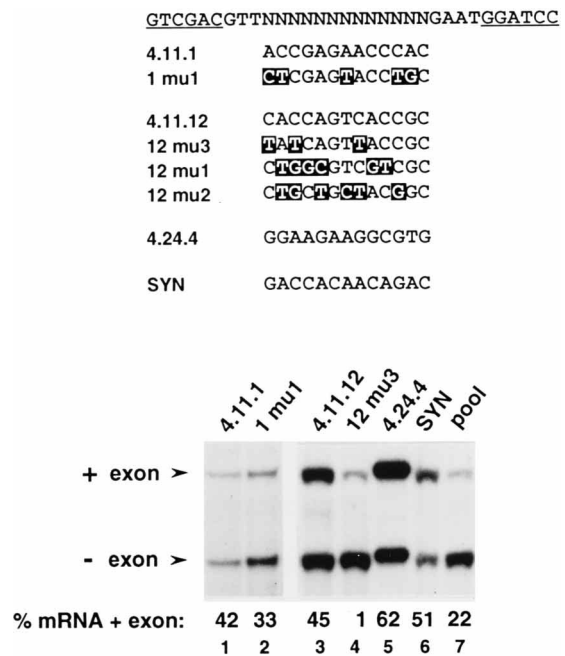


FIG. 3. In vivo analysis of individual selected exons. (Top) Sequences of selected exons and mutants thereof used for Fig. 3 to 6 are shown below the *Sall/Bam*HI selection cassette. Nucleotide substitutions are boxed in black. SYN, synthetic ACE. To test pooled clones in Fig. 3 and 4, a ligation reaction mixture containing the SXN13 or cTNT minigene and the *Sall/Bam*HI randomized cassette was transformed into bacteria. More than 400 colonies were pooled, inoculated into 100 ml of broth, and grown overnight for preparation of plasmid DNA for transfection. Lane-to-lane differences in the level of expression are primarily due to differences in transfection efficiency. We determined that the level of exon inclusion is not affected by the total level of minigene RNA (data not shown).

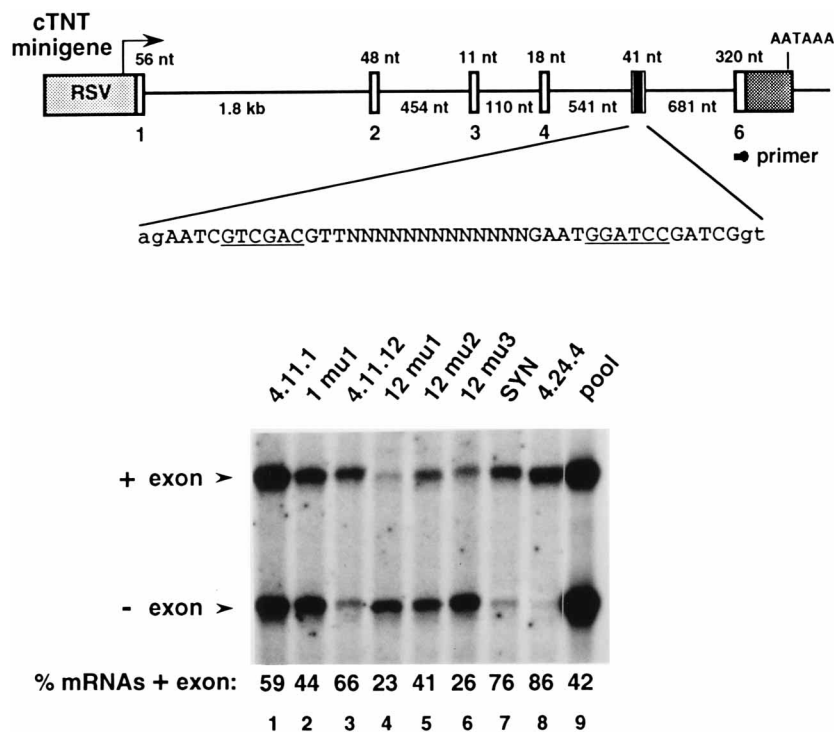


FIG. 4. Selected exons enhance splicing in a different minigene. The cTNT minigene is shown at the top. Below the minigene is the sequence of the modified exon 5 in which the natural purine-rich enhancer was replaced by the selectable *SalI/BamHI* cassettes. The intron sequence is in lowercase. The *SalI* and *BamHI* sites are underlined. The primer extension products are 169 and 210 nucleotides. RSV, Rous sarcoma virus; nt, nucleotides.

results presented in Fig. 7 demonstrate that the improved (UP) and natural (WT) cTNT purine-rich enhancers titrate factors required for ACE-dependent splicing. Furthermore, the levels of competition by the three enhancers directly correlate with their binding affinities for SR proteins (32). Competition is not due to general titration of SR proteins or other general splicing factors, since cTNT exon 5 does not compete splicing of three precursor RNAs that lack purine-rich enhancers or ACEs (data not shown). Therefore, we conclude that at least one of the factors that binds to the cTNT purine-rich enhancer is required for splicing activity of the ACE-containing exon.

RNA elements rich in adenosines and cytosines have recently been shown to affect RNA processing in several experimental systems (Table 2). All but two of the elements shown in Table 2 are naturally occurring splicing enhancers (*Drosophila dsx*, cTNT exon 16, and calcitonin/CGRP exon 4), while human  $\beta$ -globin splicing was stimulated by a mutation of six uracils to six cytosines in a heterologous exon sequence ( $\beta$ -globin/URA3 chimera). The herpes simplex virus thymidine kinase fragment allows cytoplasmic accumulation of an otherwise-unstable intronless human  $\beta$ -globin mRNA by an unknown mechanism (24). Identification of the ACE-like motif in the *Drosophila dsx* enhancer is particularly intriguing. To analyze *dsx* enhancer activity in vertebrate cells, a sequence that is found in three of the six *dsx* repeats was inserted into the selectable exon cassette and was tested in both the globin and cTNT minigenes in QT35 cells (Fig. 8). In both minigenes, the *dsx* element was one of the strongest enhancers tested (Fig. 8, lanes 1 and 4). To determine the sequence requirements for *dsx* enhancer activity in vertebrate cells, we introduced separate mutations into the T/C-rich 5' region and the A/C-rich 3' region. A 3-nucleotide substitution in the A/C-rich segment of

the motif significantly decreased enhancer activity (Fig. 8, lanes 3 and 6). Mutation of the first two nucleotides from TC to GA (Fig. 8, lanes 2 and 5) also decreased enhancer activity. This mutation has been shown to prevent Tra2 binding and to inactivate Tra/Tra2-dependent splicing in transfected *Drosophila* Kc cells (21). Our results demonstrate that a single copy of the *Drosophila dsx* repeat element strongly enhances splicing of vertebrate splice sites in vertebrate cells. Furthermore, the sequence requirements for enhancer activity in vertebrate cells include but are not limited to the ACE-like region.

## DISCUSSION

We describe a novel approach to enrich for an RNA-processing signal in vertebrate cells. In vivo randomization-selection approaches have been used for *Saccharomyces cerevisiae* to demonstrate RNA-RNA interactions that are required for splicing. Madhani and Guthrie (28) used this approach to demonstrate intermolecular base pairing of U2 and U6 small nuclear RNAs, and Libri et al. (23) demonstrated a role for pre-mRNA intramolecular base pairing in splicing of a large yeast intron. These procedures used a single-step selection based on cell survival. In contrast, our approach is more like performing SELEX (46) in living cells: a pool of pre-mRNAs containing randomized sequence is spliced, and the product of one of two splicing pathways is selected by the choice of PCR priming sites, amplified, and cycled through multiple rounds. The advantage of this approach is the ability to enrich for sequences that are optimally active in vivo. RNA-processing signals in vertebrate genes rarely fit a strict consensus, making it difficult to identify them on the basis of sequence alone. An iterative procedure allowed us to enrich for on the basis of

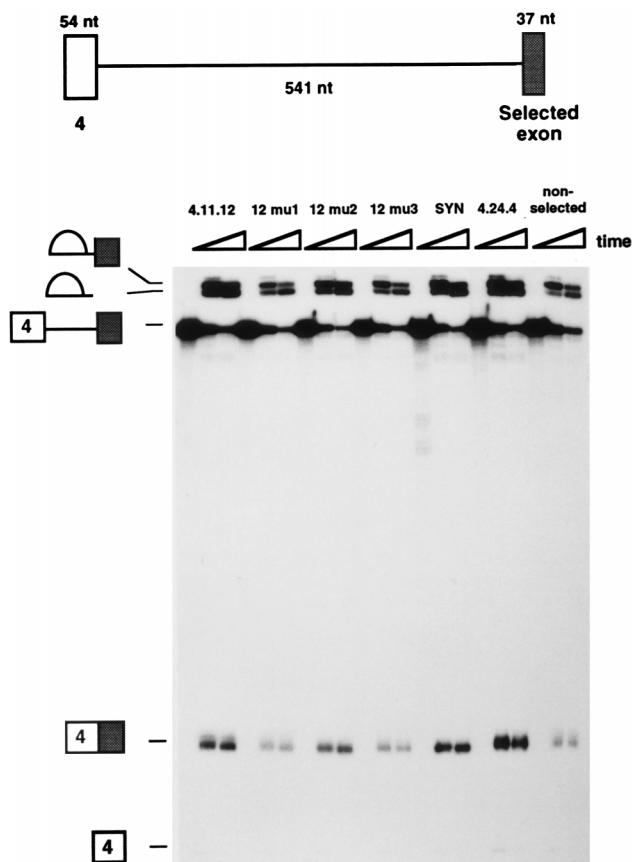


FIG. 5. In vitro analysis of selected and nonselected exons. The splicing substrate was derived from intron 4 of the cTNT minigene shown in Fig. 4 (see Materials and Methods). All splicing substrates are identical except for the sequences shown in Fig. 3. All seven RNAs were synthesized from the same cocktail to ensure identical specific activities. Reactions were performed in HeLa cell nuclear extracts for 0, 45, and 90 min. Splice products and intermediates are identified to the left of the gel. The nonselected exon contains the sequence CACTCCTTACCGA and expresses 7% exon inclusion in the SXN13 minigene in vivo. nt, nucleotides.

activity and identify a general sequence motif even though no clear consensus sequence could be discerned. It should be possible to use this approach to identify cell-specific splicing enhancers by performing the selection procedure in differentiated cell types. It may also be useful to perform directed "in vivo SELEX" in cells that overexpress individual splicing factors.

Interestingly, the ACE motifs described in this report were not identified in a similar in vitro splicing selection scheme that involved HeLa cell nuclear extracts (42). This may reflect differences in the splicing enhancers preferred in HeLa compared to QT35 cells or differences between in vivo and in vitro selection approaches. As in our study, these authors identified a number of enhancer sequences that were represented only once in the set of sequenced clones. We could not find common motifs among this group of non-ACE, non-purine-rich enhancers isolated in the two studies. In transfection studies, we find that enhancer activities of non-ACE, non-purine-rich selected exons are as strong as or stronger than those of the purine-rich elements or ACEs (data not shown). Therefore, our results are consistent with those of Tian and Kole (42) in suggesting that additional classes of splicing enhancers remain to be identified.

ACEs can be found in elements shown to affect RNA processing in several systems (Table 2). Our results identify this motif as a class of splicing enhancers. Exon 16 of cTNT was found to contain an element that enhanced splicing of heterologous transcripts in vitro and promoted binding of U2AF65 to the upstream 3' splice site (47). U2AF65 binding was competed by the 3' one-third of cTNT exon 16, which contains ACEs and purine-rich motifs. Point mutations in the ACE but not the purine-rich motif prevented competition of U2AF65 binding. Therefore, an ACE in cTNT exon 16 promotes binding of U2AF65 to the upstream 3' splice site. Exon-facilitated binding of U2AF65 to the 3' splice site required nuclear components in addition to U2AF65. Presumably these factors include those that make direct contact with the splicing enhancer.

The specific factors that interact with the ACE and mediate enhancer activity remain to be determined. We show that the cTNT purine-rich enhancer titrates factors required for ACE activity (Fig. 7), suggesting that at least one of the SR proteins required for cTNT enhancer activity is also required for ACE activity. It is not yet clear whether SR proteins bind directly to the ACE. We previously demonstrated that the cTNT enhancer binds four of the essential SR proteins and that the binding affinity of the SR proteins directly correlated with enhancer strength in vivo and in vitro (32). The ACE identified in cTNT exon 16 was shown to immunoprecipitate with anti-

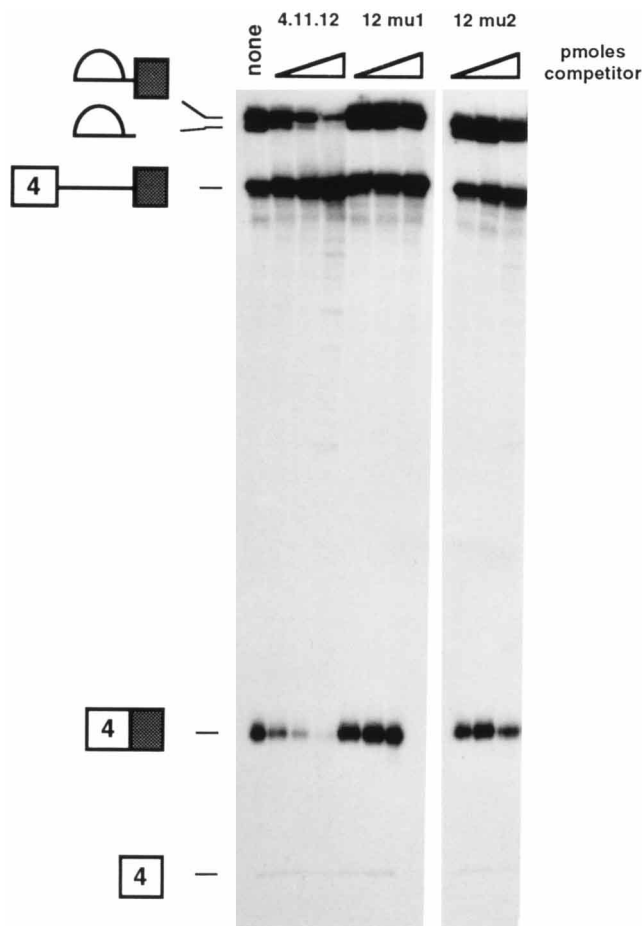


FIG. 6. Splicing of the 4.11.12 substrate shown in Fig. 5 was challenged with 4, 10, and 20 pmol of competitor RNA. Competitor RNAs contain the last 50 nucleotides of cTNT intron 4 linked with the exon shown in Fig. 4.

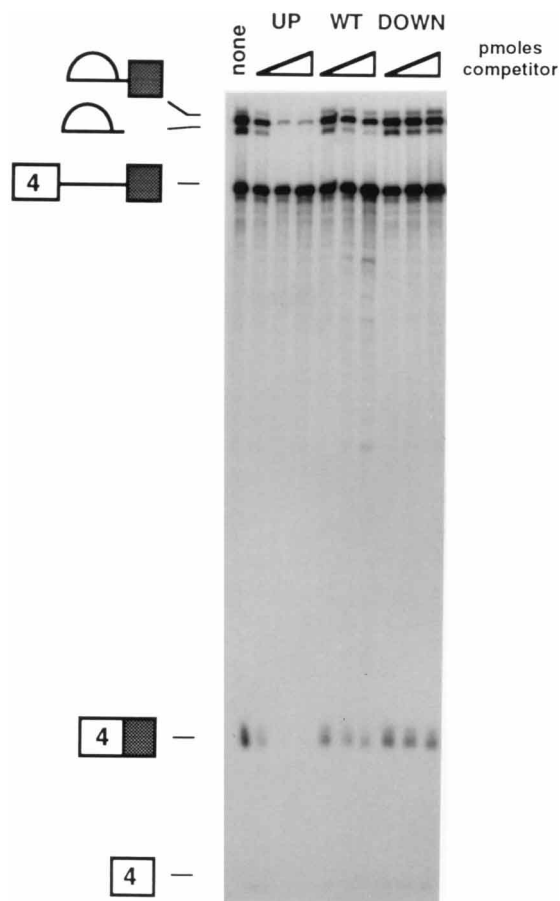


FIG. 7. Splicing of the 4.11.12 substrate was challenged with 10, 30, and 50 pmol of RNAs containing the cTNT exon 5 purine-rich splicing enhancer. The UP, WT, and DOWN RNAs contain, respectively, a mutation that increases enhancer activity in vivo, the unmodified exon, and a mutation that nearly eliminates activity (32).

SC35 antibodies (47). However, since SC35 neither binds nor activates the cTNT enhancer (32), this is not likely to be the factor titrated by an excess of cTNT exon 5 RNA. We expect that ACEs are ubiquitously recognized. We have demonstrated ACE activity in QT35 cells and HeLa cell nuclear extracts. We have not tested ACE activity in other cells; however, identi-

TABLE 2. RNA-processing elements containing ACEs<sup>a</sup>

Gene	Sequence <sup>a</sup>	Reference
CT/CGRP exon 4	<b>ACUUC</b> <u>CAACA</u> AGUU	46
cTNT exon 16	<b>CCACC</b> GAGAGGU	47
$\beta$ -globin/URA3 chimera	<b>CCAAGUACA</b> <u>CCCCCACU</u>	15
<i>dsx</i> (1)	<u>UCUUCAAU</u> <b>CAACA</b> AG	31
<i>dsx</i> (2)	<u>UCUUCAAU</u> <b>CAACA</b> UU	31
<i>dsx</i> (3)	<u>UCUACA</u> <b>AAUACA</b> AUU	31
<i>dsx</i> (4)	<u>UCUUCAAU</u> <b>CAACA</b> AU	31
<i>dsx</i> (5)	<u>UCAACA</u> <b>AAUACA</b> AUA	31
HSV <sup>b</sup> thymidine kinase	UCGCG <b>AA</b> CAUUC <b>ACACCACACAAC</b> CCGCCUCGA	24

<sup>a</sup> ACE-like elements are in boldface. The underlined nucleotides in the  $\beta$ -globin/URA3 chimera sequence is a substitution of six cytosines for six uracils that activated splicing of  $\beta$ -globin splice sites. Five of the six repeats within the *Drosophila dsx* exon 4 enhancer are shown, with the 13-nucleotide consensus sequences underlined.

<sup>b</sup> HSV, herpes simplex virus.

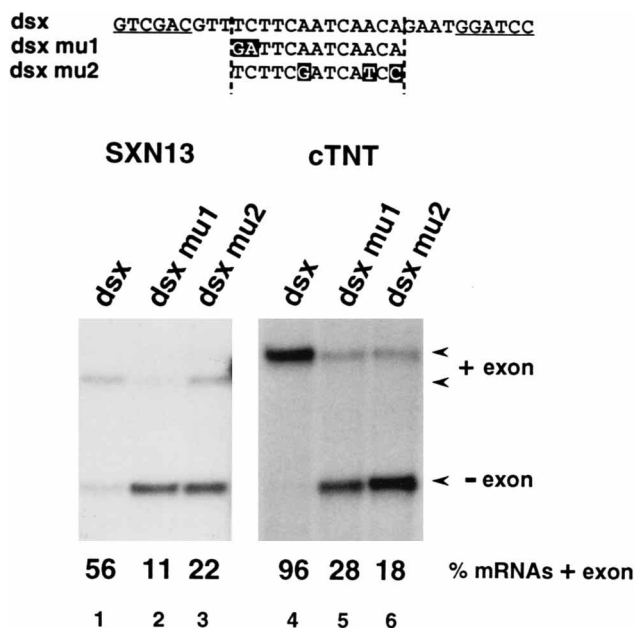


FIG. 8. A single copy of the *Drosophila dsx* repeat enhances splicing of vertebrate splice sites in vertebrate cells. The sequences of the unmodified and mutant *dsx* elements within the 24-nucleotide *SalI/BamHI* selection cassette are shown at the top. Nucleotide substitutions are boxed in black. The *SalI* and *BamHI* restriction sites are underlined. The cassette was tested in the SXN13 and cTNT minigenes shown in Fig. 2 and 4, respectively. The results from transient transfection and primer extension analysis are shown.

cation of a natural ACE in the muscle-specific cTNT gene strongly suggests that ACEs are active in striated muscle.

Our results suggest the possibility that the *dsx* enhancer is a member of a previously unrecognized class of conserved ACEs. The *dsx* repeat appears to contain two sequence motifs: a T/C-rich motif at the 5' end and an A/C-rich motif at the 3' end. Our results indicate that both motifs are required for its function (Fig. 8). Consistent with a bipartite structure, these two motifs have recently been shown to bind different sets of splicing factors by site-specific UV cross-linking (27). The T/C-rich 5' half binds to an SR protein (RBP1 in *Drosophila* and 9G8 in mammals), and the A/C-rich component binds *Drosophila* Tra2. In these experiments, Tra2 was added exogenously to HeLa cell nuclear extracts. However, previous results suggested that HeLa cell extracts contain Tra2-like activity, as addition of Tra alone to HeLa cell nuclear extracts activated the *dsx* female-specific 3' splice site (43). Interestingly, two human homologs of Tra2 have recently been identified (8). Taken together, these results suggest the interesting possibility that the ACEs selected in this procedure are targets for human Tra2.

In summary, our results demonstrate that it is feasible to perform iterative selection for RNA-processing signals in vivo. Active motifs were enriched within three rounds, suggesting that in vivo selection conditions are inherently stringent. This is a powerful approach for identifying nucleic acid sequences that are directly relevant to RNA processing in living cells.

#### ACKNOWLEDGMENTS

We thank Sue Berget and Bill Mattox for advice during the course of the project and Gil Cote, Bill Mattox, Miles Wilkinson, and members of the Cooper laboratory for comments on the manuscript.

This project was supported by the American Cancer Society and the Moran Foundation. T.A.C. is an Established Investigator of the American Heart Association.

## REFERENCES

1. **Amendt, B. A., D. Hesslein, L.-J. Chang, and C. M. Stoltzfus.** 1994. Presence of negative and positive *cis*-acting RNA splicing elements within and flanking the first *tat* coding exon of human immunodeficiency virus type 1. *Mol. Cell. Biol.* **14**:3960–3970.
2. **Berget, S. M.** 1995. Exon recognition in vertebrate splicing. *J. Biol. Chem.* **270**:2411–2414.
3. **Black, D.** 1995. Finding splice sites within a wilderness of RNA. *RNA* **1**:763–771.
4. **Black, D. L.** 1992. Activation of *c-src* neuron-specific splicing by an unusual RNA element in vivo and in vitro. *Cell* **69**:795–807.
5. **Caputi, M., G. Casari, S. Guenzi, R. Tagliabue, A. Sidoli, C. A. Melo, and F. E. Baralle.** 1994. A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. *Nucleic Acids Res.* **22**:1018–1022.
6. **Carlo, T., D. A. Sterner, and S. M. Berget.** 1996. An intron splicing enhancer containing a G-rich repeat facilitates inclusion of a vertebrate micro-exon. *RNA* **2**:342–353.
7. **Cooper, T. A.** 1992. In vitro splicing of cardiac troponin-T precursors—exon mutations disrupt splicing of the upstream intron. *J. Biol. Chem.* **267**:5330–5338.
8. **Dauwalder, B., F. Amayanzanaires, and W. Mattox.** 1996. A human homologue of the *Drosophila* sex determination factor transformer-2 has conserved splicing regulatory functions. *Proc. Natl. Acad. Sci. USA* **93**:9004–9009.
9. **Del Gatto, F., and R. Breathnach.** 1995. Exon and intron sequences, respectively, repress and activate splicing of a fibroblast growth factor receptor 2 alternative exon. *Mol. Cell. Biol.* **15**:4825–4834.
10. **Dietz, H. C., D. Valle, C. A. Francomano, R. J. Kendzior, R. E. Pyeritz, and G. R. Cutting.** 1993. The skipping of constitutive exons in vivo induced by nonsense mutations. *Science* **269**:680–683.
11. **Dirksen, W. P., R. K. Hampson, Q. Sun, and F. M. Rottman.** 1994. A purine-rich exon sequence enhances alternative splicing of bovine growth hormone pre-mRNA. *J. Biol. Chem.* **269**:6431–6436.
12. **Dirksen, W. P., Q. A. Sun, and F. M. Rottman.** 1995. Multiple splicing signals control alternative intron retention of bovine growth hormone pre-mRNA. *J. Biol. Chem.* **270**:5346–5352.
13. **Dominski, Z., and R. Kole.** 1991. Selection of splice sites in pre-mRNAs with short internal exons. *Mol. Cell. Biol.* **11**:6075–6083.
14. **Dominski, Z., and R. Kole.** 1992. Cooperation of pre-RNA sequence elements in splice site selection. *Mol. Cell. Biol.* **12**:2108–2114.
15. **Dominski, Z., and R. Kole.** 1994. Identification of exon sequences involved in splice site selection. *J. Biol. Chem.* **269**:23590–23596.
16. **Fu, X. D.** 1995. The superfamily of arginine/serine-rich splicing factors. *RNA* **1**:663–680.
17. **Gontarek, R. R., and D. Derse.** 1996. Interactions among SR proteins, an exonic splicing enhancer, and a lentivirus Rev protein regulate alternative splicing. *Mol. Cell. Biol.* **16**:2325–2331.
18. **Heinrichs, V., and B. S. Baker.** 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.
19. **Huh, G. S., and R. O. Hynes.** 1994. Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. *Genes Dev.* **8**:1561–1574.
20. **Humphrey, M. B., J. Bryan, T. A. Cooper, and S. M. Berget.** 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.
21. **Inoue, K., K. Hoshijima, I. Higuchi, H. Sakamoto, and Y. Shimura.** 1992. Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of doublesex primary transcript for sex-specific RNA processing. *Proc. Natl. Acad. Sci. USA* **89**:8092–8096.
22. **Lavigne, A., H. La Branche, A. R. Kornbliht, and B. Chabot.** 1993. A splicing enhancer in the human fibronectin alternate ED1-exon interacts with SR-proteins and stimulates U2-snRNP binding. *Genes Dev.* **7**:2405–2417.
23. **Libri, D., F. Stutz, T. McCarthy, and M. Rosbash.** 1995. RNA structural patterns and splicing: molecular basis for an RNA-based enhancer. *RNA* **1**:425–436.
24. **Liu, X. D., and J. E. Mertz.** 1995. HnRNP L binds a *cis*-acting RNA sequence element that enables intron-independent gene expression. *Genes Dev.* **9**:1766–1780.
25. **Lou, H., Y. Yang, G. J. Cote, S. M. Berget, and R. F. Gagel.** 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.
26. **Lynch, K. W., and T. Maniatis.** 1995. Synergistic interactions between two distinct elements of a regulated splicing enhancer. *Genes Dev.* **9**:284–293.
27. **Lynch, K. W., and T. Maniatis.** 1996. Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosophila* doublesex splicing enhancer. *Genes Dev.* **10**:2089–2101.
28. **Madhani, H. D., and C. Guthrie.** 1994. Randomization-selection analysis of snRNAs in vivo: evidence for a tertiary interaction in the spliceosome. *Genes Dev.* **8**:1071–1086.
29. **Manley, J. L., and R. Tacke.** 1996. SR proteins and splicing control. *Genes Dev.* **10**:1569–1579.
30. **Moore, M. J., C. C. Query, and P. A. Sharp.** 1993. Splicing of precursors to mRNAs by the spliceosome, p. 305–358. *In* R. F. Gesteland and J. F. Atkins (ed.), *RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Nagoshi, R. N., and B. S. Baker.** 1990. Regulation of sex-specific RNA splicing at the *Drosophila* doublesex gene: *cis*-acting mutations in exon sequences alter sex-specific RNA splicing patterns. *Genes Dev.* **4**:89–97.
32. **Ramchatesingh, J., A. M. Zahler, K. M. Neugebauer, M. B. Roth, and T. A. Cooper.** 1995. A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol. Cell. Biol.* **15**:4898–4907.
33. **Reed, R.** 1996. Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr. Opin. Genet. Dev.* **6**:215–220.
34. **Rio, D.** 1993. Splicing of pre-mRNA: mechanism, regulation and role in development. *Curr. Opin. Genet. Dev.* **3**:574–584.
35. **Ryan, K. J., and T. A. Cooper.** 1996. Muscle-specific splicing enhancers regulate inclusion of the cardiac troponin T alternative exon in embryonic skeletal muscle. *Mol. Cell. Biol.* **16**:4014–4023.
36. **Santisteban, I., F. X. Arredondovega, S. Kelly, M. Loubser, N. Meydan, C. Roifman, P. L. Howell, T. Bowen, K. I. Weinberg, M. L. Schroeder, and M. S. Hersfield.** 1995. Three new adenosine deaminase mutations that define a splicing enhancer and cause severe and partial phenotypes: implications for evolution of a CpG hotspot and expression of a transduced ADA cDNA. *Hum. Mol. Genet.* **4**:2081–2087.
37. **Sirand-Pugnet, P., P. Durosay, E. Brody, and J. Marie.** 1995. An intronic (A/U)GGG repeat enhances the splicing of an alternative intron of the chicken  $\beta$ -tropomyosin pre-mRNA. *Nucleic Acids Res.* **23**:3501–3507.
38. **Staffa, A., and A. Cochrane.** 1995. Identification of positive and negative splicing regulatory elements within the terminal *tat-rev* exon of human immunodeficiency virus type 1. *Mol. Cell. Biol.* **15**:4597–4605.
39. **Staknis, D., and R. Reed.** 1994. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Mol. Cell. Biol.* **14**:7670–7682.
40. **Sun, Q., A. Mayeda, R. K. Hampson, A. R. Krainer, and F. M. Rottman.** 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. *Genes Dev.* **7**:2598–2608.
41. **Tanaka, K., A. Watakabe, and Y. Shimura.** 1994. Polypurine sequences within a downstream exon function as a splicing enhancer. *Mol. Cell. Biol.* **14**:1347–1354.
42. **Tian, H. C., and R. Kole.** 1995. Selection of novel exon recognition elements from a pool of random sequences. *Mol. Cell. Biol.* **15**:6291–6298.
43. **Tian, M., and T. Maniatis.** 1992. Positive control of pre-messenger-RNA splicing in vitro. *Science* **256**:237–240.
44. **Tian, M., and T. Maniatis.** 1993. A splicing enhancer complex controls alternative splicing of doublesex pre-messenger RNA. *Cell* **74**:105–114.
45. **Tian, M., and T. Maniatis.** 1994. A splicing enhancer exhibits both constitutive and regulated activities. *Genes Dev.* **8**:1703–1712.
46. **van Oers, C. C. M., G. J. Adema, H. Zandberg, T. C. Moen, and P. D. Baas.** 1994. Two different sequence elements within exon 4 are necessary for calcitonin-specific splicing of human calcitonin/calcitonin gene-related peptide I pre-mRNA. *Mol. Cell. Biol.* **14**:951–960.
47. **Wang, Z., H. M. Hoffmann, and P. J. Grabowski.** 1995. Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA* **1**:21–35.
48. **Watakabe, A., K. Tanaka, and Y. Shimura.** 1993. The role of exon sequences in splice site selection. *Genes Dev.* **7**:407–418.
49. **Xie, W., and L. I. Rothblum.** 1991. Rapid, small-scale RNA isolation from tissue culture cells. *BioTechniques* **11**:325–327.
50. **Xu, R., J. Teng, and T. A. Cooper.** 1993. The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. *Mol. Cell. Biol.* **13**:3660–3674.
51. **Yeakley, J. M., F. Hedjran, J.-P. Morfin, N. Merillat, M. G. Rosenfeld, and R. B. Emeson.** 1993. Control of calcitonin/calcitonin gene-related peptide pre-mRNA processing by constitutive intron and exon elements. *Mol. Cell. Biol.* **13**:5999–6011.