MOLECULAR AND CELLULAR BIOLOGY, June 1993, p. 3660–3674
0270-7306/93/063660-15$02.00/0

Vol. 13, No. 6

The Cardiac Troponin T Alternative Exon Contains a Novel Purine-Rich Positive Splicing Element

RUISHU XU,[1] JAMES TENG,[1] AND THOMAS A. COOPER[1,2]*

Departments of Pathology[1] and Cell Biology,[2] Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030

Received 14 January 1993/Returned for modification 8 March 1993/Accepted 23 March 1993

We have characterized a novel positive-acting splicing element within the developmentally regulated alternative exon (exon 5) of the cardiac troponin T (cTNT) gene. The exon splicing element (ESE) is internal to the exon portions of the splice sites and is required for splicing to the 3′ splice site but not the 5′ splice site flanking the exon. Sequence comparisons between cTNT exon 5 and other exons that contain regions required for splicing reveal a common purine-rich motif. Sequence within cTNT exon 5 or a synthetic purine-rich motif facilitates splicing of heterologous alternative and constitutive splice sites in vivo. Interestingly, the ESE is not required for the preferential inclusion of cTNT exon 5 observed in primary skeletal muscle cultures. Our results strongly suggest that the purine-rich ESE serves as a general splicing element that is recognized by the constitutive splicing machinery.

Most metazoan genes express pre-mRNAs that contain multiple exons which must be precisely joined to form mature mRNA. Extensive sequence comparisons have identified conserved pre-mRNA *cis* elements that flank exons and mutation analyses have demonstrated the importance of these elements for pre-mRNA splicing (30). Consensus sequences have been derived for four splicing elements: (i) the 5′ splice site which includes the last 3 nucleotides of the exon and the first 6 nucleotides of the intron, (ii) the 3′ splice site which includes the last 4 nucleotides of the intron and the first nucleotide of the exon, (iii) a pyrimidine-rich tract of variable length generally located immediately upstream from the 3′ splice site, and (iv) a variable heptanucleotide at the site of lariat formation called the branch site (for simplicity, elements ii to iv will be collectively referred to as the 3′ splice site). These sequences direct assembly of the splicing complex, or spliceosome, by sequentially binding small nuclear ribonucleoproteins (U1, U2, U4, U5, and U6) and several auxiliary factors (for reviews, see references 22 and 56). Recognition of the 5′ splice site involves direct base pairing with U1 small nuclear RNA, while recognition of the 3′ splice site requires several complex interactions including direct base pairing of U2 small nuclear RNA with the branch site preceded by binding of U2 auxiliary factor to the polypyrimidine tract. Several other factors have been demonstrated to interact with the 3′ splice site and are important for its recognition.

While it is clear that the consensus signals are necessary for splicing, these elements are not sufficient for splice site selection. Because of their small size and loosely conserved sequence, the consensus splicing elements do not contain sufficient information to allow authentic splice sites to be distinguished from unused or cryptic splice sites that are commonly found within introns and exons (46, 51). Indeed, computer-based efforts to identify splice sites in the nucleotide sequences of uncharacterized genes are plagued by both false negatives and false positives (3, 51). The lack of understanding of how splice sites are selected is further underscored by the ability of some pre-mRNAs to undergo alternative splicing in which splicing to specific splice sites is optional. In many cases, the relative levels of alternative spliced mRNAs are regulated according to cell type or developmental stage (53).

A complete understanding of the mechanism for splice site selection requires the identification of all *cis*-acting elements that contribute to the efficiency with which splice sites are recognized by the splicing machinery. The lack of sequence specificity in the consensus splicing elements suggests the possibility that additional *cis* elements are involved in the selection of some splice sites. Specifically, internal exon sequence has been demonstrated to play a role in both alternative and constitutive splice site selection in several experimental systems (internal exon sequence refers to exon sequence that is internal to the exon portions of the consensus splice sites) (4, 12–14, 23, 25, 28, 29, 34, 37, 42, 48, 52, 54, 57, 61). In addition, some genetic diseases have been found to be caused by spontaneous mutations within internal exon sequence that lead to missplicing (38, 65). The best-characterized exon splicing element (ESE) is in the female-specific exon of the *Drosophila doublesex* (*dsx*) gene in which a repeated 13-nucleotide sequence in the downstream exon located more than 200 nucleotides from the 3′ splice site is required for 3′ splice site activation (24, 26, 42). *dsx* 3′ splice site activation is mediated by binding of the *transformer* (*tra*) and *transformer-2* (*tra-2*) gene products to the exon repeats (60). In vertebrates, specific internal exon sequences required for splicing have yet to be defined and characterized.

We have been using the avian cardiac troponin T (cTNT) gene as a model gene to investigate the role of internal exon sequence in splice site selection. Expression of the cTNT gene is restricted to cardiac muscle and embryonic skeletal muscle (10, 11). The cTNT gene expresses a single pre-mRNA that contains 18 exons and is alternatively spliced such that one exon, exon 5, is either included or excluded from the mature mRNA (11). Developmental regulation of alternative splicing results in a transition from >90% of mRNAs that include the exon in early embryonic heart and skeletal muscle to >95% of mRNAs that exclude the exon in the adult (11). Using transfection analysis of cTNT minigenes, we have previously shown that mutations within the

* Corresponding author.

alternative exon resulted in exon skipping in vivo (12). In vitro, these mutations specifically blocked splicing of exons 4 and 5, but not exons 5 and 6, suggesting that internal exon mutations specifically disrupt 3′ splice site recognition (8). To investigate the role of internal exon sequence in regulated splice site selection, we have established a transient transfection system using two cell cultures that differ in their regulation of exon 5 alternative splicing. In this report, we demonstrate that cTNT exon 5 contains an ESE that has the following characteristics: (i) it is required specifically for splicing to the flanking 3′ splice site, while splicing to the flanking 5′ splice site is independent of internal exon sequence; (ii) it contains a purine-rich motif that is found in several heterologous exons within regions shown to be required for splicing; (iii) it facilitates splicing of both alternative and constitutive exons from heterologous genes; and (iv) it is not required for the regulated splicing observed in primary skeletal muscle cultures. Our results suggest that ESEs may be a general feature of some constitutive and alternative exons.

MATERIALS AND METHODS

DNA constructs. Standard cloning protocols were used (50). Oligonucleotides used for primer extension and polymerase chain reaction (PCR) cloning were synthesized by the Institute for Molecular Genetics, Baylor College of Medicine.

The ΔPB construct was derived from SM-1 (9) by replacing the cTNT genomic fragment downstream of exon 6 with a skeletal α-actin genomic fragment containing the last exon (exon 7) and 377 bp of 3′-flanking genomic sequence (18). The two exons were fused at an *Spe*I site created by 3-nucleotide substitutions in cTNT exon 6 (leaving 55 nucleotides of the 57-nucleotide exon) and by the addition of an *Spe*I linker to a *Bcl*I site in α-actin exon 7 located 261 nucleotides from its 3′ end. S1 nuclease analysis of ΔPB mRNAs demonstrate efficient use of the correct polyadenylation site in both muscle and nonmuscle cells (data not shown). In addition, Northern (RNA) blot analysis indicates that minigene pre-mRNAs are efficiently spliced (data not shown).

The E51 and E53 mutations (12) were introduced into ΔPB by replacing a *Cla*I-*Xho*I fragment containing exon 5. To replace cTNT exon 5 with skeletal troponin I (sTNI) exon 2 and derivatives, the first 2 and last 3 nucleotides of exon 5 were changed to create *Pvu*II and *Sna*BI sites, respectively. These restriction sites cut at the intron-exon junctions of exon 5, allowing the precise removal of the exon. Three 30-bp DNAs containing the ΔPSR, ΔE25, and ΔE2CON replacement exons were synthesized directly or by PCR and cloned blunt ended into the *Pvu*II-*Sna*BI vector. All clones were confirmed by sequencing of the exon and at least 50 nucleotides of both flanking introns.

ΔDMR (see Fig. 2B) was constructed by using PCR-generated restriction sites to replace cTNT exon 5 and immediately flanking introns with the comparable region of sTNI exon 2 without introducing significant changes in intron sizes.

The ΔE5C vector (see Fig. 3A) was constructed by cloning a *Cla*I-*Xho*I fragment from E53 into the *Cla*I-*Xho*I sites of E52 (12). To expand the size of the exon, both orientations of the following fragments from pBR322 were cloned into the *Cla*I site: 15-, 26-, 34-, 76-, and 110-bp *Hpa*II; 30- and 40-bp *Hin*PI; and 48- and 107-bp *Hpa*II-*Taq*. The integrity and orientation of inserts and the integrity of splice sites were confirmed by sequencing.

ΔBBC (see Fig. 4A) was constructed by a collapse from the second nucleotide of intron 5 (*Bbs*I site) to a *Dde*I site located 86 nucleotides upstream of exon 6. The intron sequence immediately flanking exon 5 in ΔBBC (GT|TCAGT [vertical line denotes *Bbs*I-*Dde*I fusion site]) does not recreate a functional 5′ splice site. ΔBCE51 and ΔBCPSR were constructed from inserts generated by PCR using an oligonucleotide that primed 5′-ward from the 3′ end of the E51 or PSR exon. PCR fragments were treated with T4 DNA polymerase to remove A residues added to the 3′ end, cut with *Cla*I (upstream site), and cloned into ΔBBC *Cla*I-*Bbs*I fragments (blunt ended).

To construct the μBRB series (see Fig. 5A), ΔPB was modified such that the 5′ end of exon 5 was fused to the 3′ end of the Rous sarcoma virus long terminal repeat, and a *Bam*HI site was introduced 24 nucleotides downstream of exon 5. PCR-generated fragments from ΔPB, ΔE51, ΔPSR, and ΔE25 that included the first nucleotide of the exon to 3′ of an *Xho*I site in the downstream intron were used to replace a *Bam*HI-*Xho*I fragment of the modified plasmid. The resulting plasmids were identical except for the substitutions shown in Fig. 2A.

μAVWT.BSC (see Fig. 6A) was constructed using a truncated avian sarcoma virus (ASV) *env* construct generously provided by R. Katz and A. M. Skalka (Fox Chase Cancer Center). The intron contains a deletion of about 4,450 nucleotides and removes the *gag* and *pol* sequences (29). This deletion leaves intact about 100 nucleotides upstream of the 3′ splice site and 125 nucleotides downstream of the 5′ splice site. PCR primers containing the appropriate restriction enzyme sites were used to amplify and clone the ASV *env* fragment shown in Fig. 6A between the Rous sarcoma virus long terminal repeat and skeletal α-actin exon 7, removing all cTNT sequence from ΔPB. DNA fragments to be tested for splicing-enhancing activity were blunt end cloned between *Bgl*II and *Spe*I sites (see Fig. 6A). cTNT exon 5 and sTNI exon 2 were derived from synthetic double-stranded oligonucleotides. GAR and MGAR concatamers (see Fig. 8B) were derived from 9-nucleotide oligonucleotides and their complements that base paired to give single-nucleotide 5′ overhangs which ensured concatamerization in a single orientation. Following concatamerization, ends were filled in and double-stranded concatamers were isolated on gels for cloning. All clones were confirmed by sequencing.

μIHBG.D (see Fig. 7A) was constructed by PCR-assisted cloning in a modified design of Dominski and Kole (15). The entire intron 1 of the human β-globin gene was duplicated to create an 8-nucleotide middle exon (**GACGTCAG**) containing an *Aat*II site (bold print) incorporated into the first nucleotide of exon 1 (G) and the last 3 nucleotides of exon 2 (CAG). Cleavage with *Aat*II and removal of the ACGT 3′ overhangs (not underlined) using T4 DNA polymerase left the exon portions of the β-globin splice sites intact (underlined). Fragments to be tested for splicing-enhancing activity were blunt end cloned into this site, as described for μAVWT.BSC above. All clones were confirmed by sequencing.

Cell culture, transfection, RNA extraction, and analysis. Media and medium components were obtained from GIBCO/BRL. The preparation of embryonic day 11 chicken breast muscle cultures was as described by Mar et al. (36). Transient transfections were done by the method of Chen and Okayama (5) with the following modifications. Cells were

plated at 10^6 (QT35) and 1.8×10^6 (primary skeletal muscle) per 60-mm-diameter dish. QT35 and muscle cultures were transfected with 5 to 15 μg of DNA per plate using aliquots from a single cocktail. Plates were incubated in 5% CO_2. Muscle cultures were washed after 4 h, and QT35 cultures were washed after incubation overnight. Total RNA was extracted 48 h after the start of transfection by the guanidinium thiocyanate procedure of Chomczynski and Sacchi (7) as modified by Xie and Rothblum (67). Primer extension and S1 nuclease analysis were performed as described previously with approximately 25 μg of total RNA from transfected cultures (half of a 60-mm-diameter plate) or 5 μg of total RNA from tissue (9).

RESULTS

Alternative splicing of cTNT minigene pre-mRNAs is regulated in primary skeletal muscle cultures. To define the role of internal exon sequence in exon recognition and in the regulation of alternative splicing, we have established a transient transfection system which demonstrates regulated alternative splicing of cTNT minigene pre-mRNAs. The results from transfection analysis of cTNT minigenes have been described previously (9, 12). Here we demonstrate that in primary skeletal muscle cultures, mRNAs from both the endogenous cTNT gene and transiently transfected minigenes preferentially include exon 5, while a nonmuscle cell line displays a default splicing pattern for minigene pre-mRNAs in which the exon is predominantly skipped.

The cTNT minigenes used in this study were derived from the ΔPB construct, which is diagrammed in Fig. 1A. The splicing patterns of both the endogenous (primary skeletal muscle only) and transfected gene products were assayed simultaneously by primer extension on total cellular RNA with an oligonucleotide complementary to cTNT exon 6. The sizes expected from primer extension on endogenous and minigene mRNAs are indicated in Fig. 1B.

Primary skeletal muscle cultures were prepared from embryonic day 11 breast muscle tissue in which >90% of the endogenous cTNT mRNAs include exon 5 (Fig. 1B, lane 5). Like breast muscle tissue, primary breast muscle cultures express endogenous mRNAs that predominantly include exon 5 (Fig. 1B, lane 4). Following transfection into these muscle cultures, cTNT minigenes also express predominantly mRNAs that include the exon (Fig. 1B, lane 2). The percentage of ΔPB mRNAs that include exon 5 in these cultures varies from 70 to 85%. Therefore, primary skeletal muscle cultures retain the splicing pattern that is characteristic of the embryonic tissue from which they were derived and presumably express the relevant *trans*-acting regulatory factors. The ΔPB pre-mRNA appears to contain the *cis* elements required to respond to these factors.

The transition to exon 5 skipping that occurs during muscle development in vivo does not occur in primary skeletal muscle cultures even after extended culture times (data not shown). These cultures also fail to express other protein isoforms characteristic of late developmental stages which may reflect a requirement for extrinsic cues (58). Since we were unable to observe a transition of splicing patterns in muscle cultures, cTNT minigenes were transiently transfected into the QT35 avian fibroblast cell line for purposes of comparison. Because QT35 cells do not express muscle-specific genes, these cells presumably lack factors that regulate splicing of muscle-specific pre-mRNAs. In QT35 cultures, only 15 to 25% of minigene mRNAs include exon 5 (Fig. 1B, lane 1). Preferential exon skipping in these cells is taken as the default splicing pathway. The same splicing pattern is observed in other nonmuscle cell types, such as HeLa and Chinese hamster oocytes (data not shown).

Using forced expression of cTNT cDNAs, we determined that mRNAs that include or exclude exon 5 do not differ in their relative stability in either muscle or QT35 cells (data not shown). Therefore, the steady-state ratio of the alternatively spliced mRNAs is determined primarily by splice site selection.

To determine whether our results in primary skeletal muscle cultures were affected by contaminating nonmuscle cells (10 to 20% of the nuclei are in nonmuscle cells), we constructed a minigene that was expressed only in differentiated myogenic cells by replacing the Rous sarcoma virus long terminal repeat of the ΔPB minigene with the muscle-specific promoter of the cTNT gene (36). We found that this minigene and ΔPB expressed the same percentage of mRNAs that include exon 5 in primary skeletal muscle cultures (data not shown). Therefore, our results from primary skeletal muscle cultures are not affected by contaminating nonmuscle cells.

The splice sites flanking exon 5 are not sufficient for splicing of a heterologous constitutive exon. We have previously reported that internal exon mutations in cTNT exon 5 lead to skipping of the exon (12). The effects of two previously reported exon mutations, ΔE51 and ΔE53, are shown in Fig. 2C to demonstrate that their effects are reproducible in different constructs and different cell types. The nucleotide substitutions for these mutations are shown in Fig. 2A. The E51 mutation results in almost complete skipping of the exon in both muscle and nonmuscle cultures (Fig. 2C, lanes 3 and 4), while the E53 mutation reduces but does not eliminate exon inclusion (lanes 5 and 6). These mutations have been demonstrated to affect splicing of exon 5 rather than mRNA stability (12 and data not shown). These results initially suggested the possibility that mutations within exon 5 inactivated a positive-acting element that is required for exon recognition.

We next tested whether the splice sites that flank exon 5 are sufficient for splicing of a heterologous exon of the same size in the exon substitution construct, ΔPSR. In this minigene, cTNT exon 5 was cleanly removed from the flanking introns and replaced with a 30-nucleotide constitutive exon (exon 2) from the chicken sTNI gene (45) (see Methods and Materials for details of cloning). It is important to note that ΔPSR is identical to ΔPB, except that exon 5 is precisely replaced by sTNI exon 2 without modification of the adjacent introns (Fig. 2A). In both muscle and QT35 cultures, the substituted exon is completely skipped (Fig. 2C, lanes 7 and 8). The failure to detect ΔPSR mRNAs that include sTNI exon 2 is not due to mRNA instability, since the identical mRNA is expressed when sTNI exon 2 is flanked by its natural splice sites (construct ΔDMR [Fig. 2B and C, lanes 11 and 12]).

It was possible that the substituted sTNI exon 2 was completely skipped because of the changes in the exon portions of the splice sites, rather than the changes in internal exon sequence. The first nucleotide of sTNI exon 2 is a consensus G and is unlikely to have a negative effect on splicing. However, the last 3 nucleotides of sTNI exon 2 in ΔPSR altered the flanking 5′ splice site such that potential base pairing with U1 snRNA was less than that with cTNT exon 5 (compare ΔPB and ΔPSR in Fig. 2A). To restore the complete natural 5′ splice site in the exon substitution construct, the last 3 nucleotides of sTNI exon 2 were

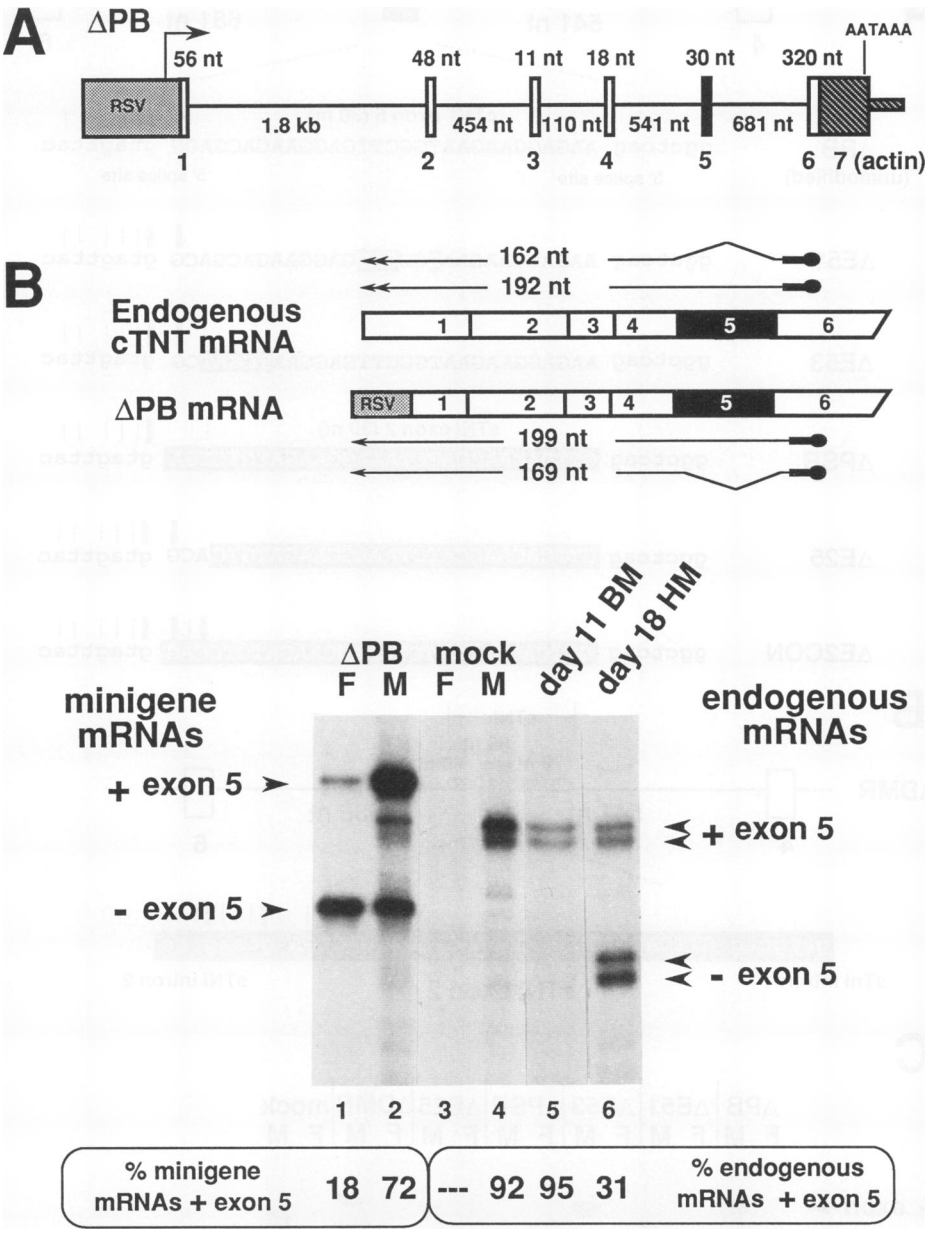

FIG. 1. Transient transfection analysis of cTNT minigenes in primary skeletal muscle and QT35 fibroblast cultures. (A) Diagram of the cTNT minigene, ΔPB. (B) Primer extension analysis. (Top) Diagrams of 5′ portions of endogenous and ΔPB mRNAs. The sizes of primer extension products from mRNAs that include or exclude exon 5 are indicated. (Bottom) Primer extension analysis. RNA from embryonic day 18 heart muscle (lane 6) shows the products from endogenous cTNT mRNAs. These mRNAs produce a doublet, because transcription initiation occurs at two sites separated by three nucleotides (11). Bands were quantified directly from the gel with an Ambis Radioanalytic Imaging System, and the results are expressed as the percentage of spliced mRNA that contains the alternative exon. Lanes: F, QT35; M, primary skeletal muscle. The diagrams are not drawn to scale. Abbreviations: RSV, Rous sarcoma virus long terminal repeat; BM, breast muscle; HM, heart muscle.

replaced with the last 3 nucleotides of cTNT exon 5 (construct ΔE25 [Fig. 2A]). However, even with the complete 5′ splice site that normally flanks cTNT exon 5, sTNI exon 2 was predominantly skipped (Fig. 2C, lanes 9 and 10).

Taken together, these results demonstrated that the 5′ and 3′ splice sites that flank cTNT exon 5 were not sufficient for recognition of a heterologous 30-nucleotide constitutive exon. These results strongly indicate that a positive-acting *cis* element or ESE within cTNT exon 5 facilitates exon recognition. The fact that this element was required for exon recognition in both muscle and nonmuscle cell types suggested that it serves as a general splicing element that is recognized by the constitutive splicing machinery (see below).

The ESE is not required for regulation of alternative splicing in primary skeletal muscle cultures. To determine whether the ESE was required for the regulation of alternative splicing, we sought to induce alternative splicing of ESE-minus exons that replaced cTNT exon 5. If the ESE were required for regulated splicing, we expect the ratio of

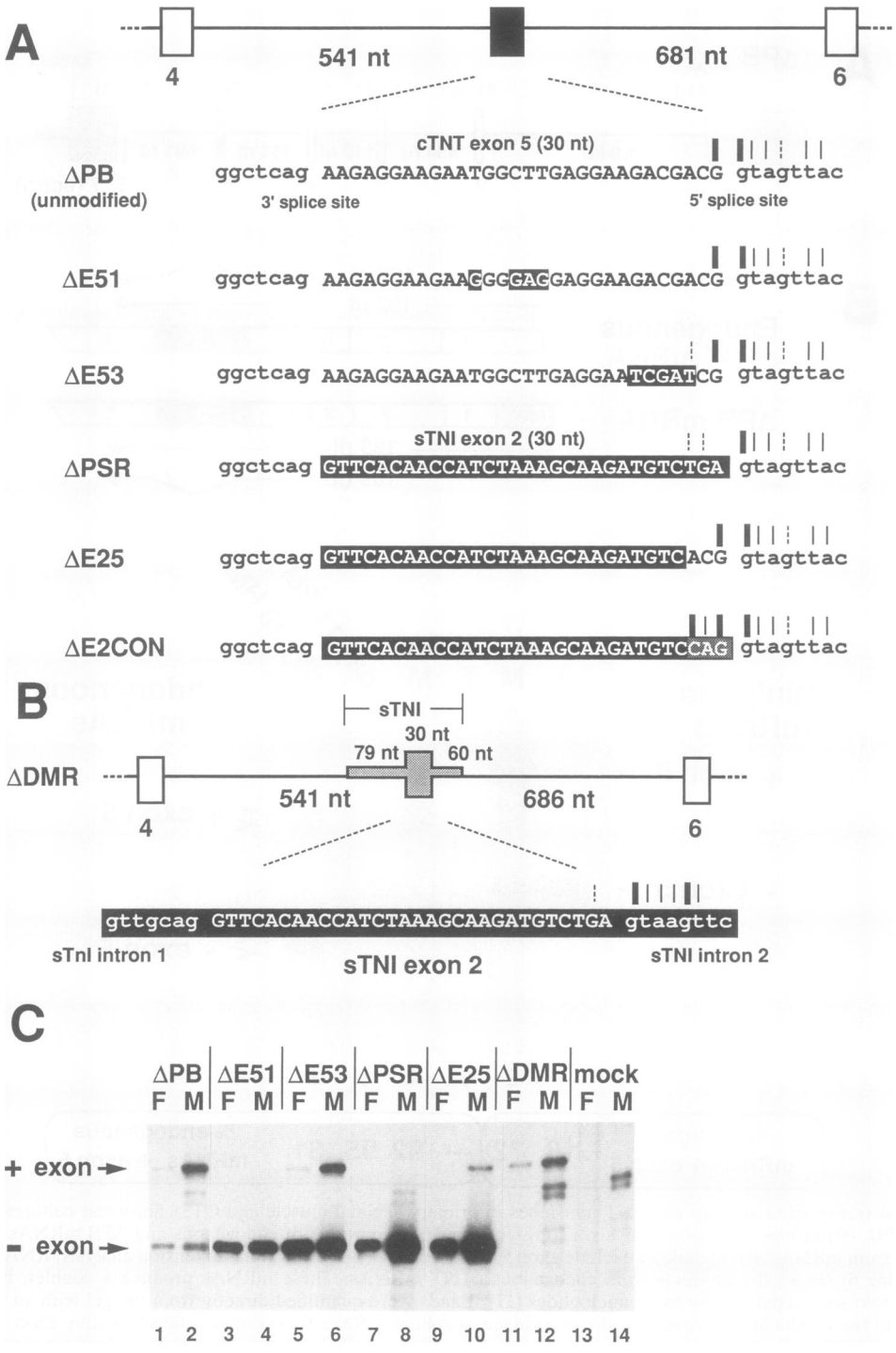

FIG. 2. Transfection analysis of exon 5 mutations and substitutions. (A) Nucleotide sequences of cTNT exon 5 and five modified exons. Exon sequence is shown in uppercase, and intron sequence is shown in lowercase. All constructs are identical to ΔPB except for the nucleotide changes indicated by black or shaded background. The potential base pairing of the 5′ splice sites with U1 snRNA is indicated as follows: G-C (‖), A-U (|), G-U (¦). The ΔPSR mutation represents a complete substitution of cTNT exon 5 with a heterologous constitutive exon (sTNI exon 2) of the identical size (see text). The ΔE25 mutation replaces the last 3 nucleotides of sTNI exon 2 with those of cTNT exon 5 to reestablish the complete natural 5′ splice site. The ΔE2CON mutation replaces the last 3 nucleotides of sTNI exon 2 with the exon portion of the consensus 5′ splice site. ΔPB retains the correct cTNT open reading frame from the natural AUG codon in exon 2 to a termination codon immediately 3′ to the cTNT exon 6–α-actin exon 7 fusion. None of these mutations introduce stop codons into this open reading frame. (B) Diagram of ΔDMR. The shaded region in the diagram represents a segment of the sTNI gene that replaced a comparable genomic region containing cTNT exon 5. This substitution maintains the size of intron 4 and increases intron 5 by only 5 nucleotides. (C) Primer extension analysis of cTNT exon 5 mutants. Primer extension was performed as described in the legend to Fig. 1, and products from minigene mRNAs that include (+) or exclude (−) exon 5 are indicated. A shorter exposure from the same gel is shown for ΔE51, ΔE53, ΔPSR, and ΔE25. Differences in minigene mRNA levels primarily reflect differences in transfection efficiency. We have determined that high levels of minigene mRNAs do not significantly affect the ratio of alternatively spliced mRNAs in either muscle or QT35 cultures (data not shown). Lanes: F, QT35; M, primary skeletal muscle.

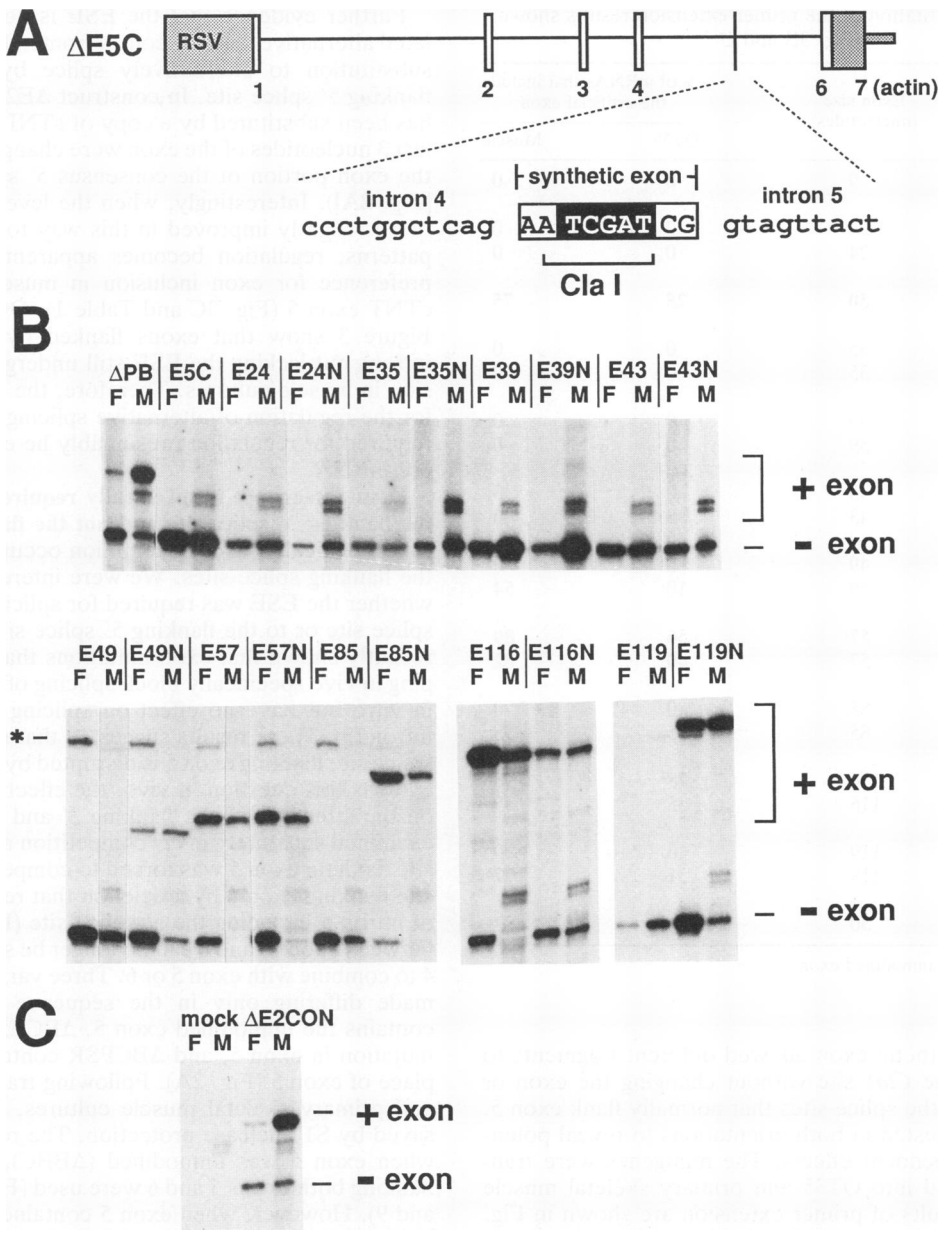

FIG. 3. The ESE is not required for regulation of alternative splicing in muscle cultures. (A) Diagram of the ΔE5C construct in which cTNT exon 5 has been replaced by a 9-bp synthetic exon containing a *Cla*I restriction site flanked by the first and last 2 nucleotides of exon 5. Exon size was increased by inserting plasmid DNA fragments into the *Cla*I site (see Materials and Methods for details). Both orientations for each fragment were tested. The names of the constructs indicate the size of the exon, and N indicates an opposite orientation of the inserted fragment. RSV, Rous sarcoma virus long terminal repeat. (B) Primer extension analysis of the ΔE5C series in QT35 and primary skeletal muscle cultures. Primer extension was performed described in the legend to Fig. 1. Each construct has been tested at least twice. Representative results from different transfections are shown. Differences in mRNA levels reflect variability in transfection efficiency. The asterisk marks a band representing RNAs that retain the 110-nucleotide intron 3. (C) Primer extension analysis of ΔE2CON. The ΔE2CON substitution of cTNT exon 5 is shown in Fig. 2A. Lanes: F, QT35; M, primary skeletal muscle. Bands were quantified directly from the gel, and the results are presented in Table 1.

RNAs with and without an ESE-minus exon to be the same in muscle and nonmuscle cells. If the ESE were not required for regulated splicing, ESE-minus exons should continue to be preferentially included in muscle cultures compared to nonmuscle cultures.

Several laboratories have demonstrated that increased exon size results in increased splicing efficiency (1, 2, 15), suggesting the possibility that the requirement for an ESE could be overcome by increased exon size. The ΔPB minigene was modified such that exon 5 was replaced by a 9-nucleotide synthetic exon that contains an internal *Cla*I site flanked by the first 2 and last 2 nucleotides of exon 5 (ΔE5C [Fig. 3A]). A series of minigenes was constructed in which the size of the ΔE5C synthetic exon was increased by inserting fragments from plasmid DNA digested with restriction enzymes that generated ends compatible with *Cla*I. The

TABLE 1. Quantitation of the primer extension results shown in Fig. 3B and C

Clone	Exon size (nucleotides)	% of mRNAs that include the artificial exon:	
		QT35	Muscle
ΔE5C	9	0	0
ΔE24	24	0	0
ΔE24N	24	0	0
ΔPB[a]	**30**	**25**	**75**
ΔE35	35	0	0
ΔE35N	35	0	0
ΔE39	39	0	0
ΔE39N	39	0	0
ΔE43	43	0	0
ΔE43N	43	0	0
ΔE49	49	0	0
ΔE49N	49	10	54
ΔE57	57	56	89
ΔE57N	57	40	82
ΔE85	85	0	0
ΔE85N	85	85	>98
ΔE116	116	78	>98
ΔE116N	116	35	45
ΔE119	119	0	0
ΔE119N	119	31	77
ΔE2CON	30	21	83

[a] ΔPB contains the unmodified exon.

design of the synthetic exon allowed different fragments to be cloned into the *Cla*I site without changing the exon or intron portion of the splice sites that normally flank exon 5. All inserts were tested in both orientations to reveal potential sequence-dependent effects. The minigenes were transiently transfected into QT35 and primary skeletal muscle cultures. The results of primer extension are shown in Fig. 3B. The bands were quantified directly from the gel with an Ambis Radioanalytic Imaging System, and these results are shown in Table 1.

The results from the exon expansion series demonstrate three points. First, artificial exons that are smaller than 49 nucleotides and are flanked by the natural splice sites of cTNT exon 5 are not recognized in the absence of the ESE. Second, artificial exons 49 to 119 nucleotides are alternatively spliced, yielding mRNAs that include or exclude the exon (with the exception of ΔE49, ΔE85, and ΔE119). Third, all alternatively spliced exons show a higher level of exon inclusion in primary skeletal muscle cultures than in QT35 cultures. From these results, we can draw two conclusions. First, in the absence of the ESE, an exon flanked by the natural splice sites of exon 5 must have a minimum size of 49 nucleotides to be recognized by the splicing machinery. This result strongly suggests that the ESE within exon 5 serves to compensate for the relatively small size of the exon. Second, the ESE is not required for the preferential exon inclusion in primary skeletal muscle cultures.

Further evidence that the ESE is not required for regulated alternative splicing comes from inducing a sTNI exon 2 substitution to alternatively splice by strengthening the flanking 5′ splice site. In construct ΔE2CON, cTNT exon 5 has been substituted by a copy of sTNI exon 2 in which the last 3 nucleotides of the exon were changed to CAG to match the exon portion of the consensus 5′ splice site (ΔE2CON [Fig. 2A]). Interestingly, when the level of exon 2 recognition is slightly improved in this way to allow both splicing patterns, regulation becomes apparent: there is a strong preference for exon inclusion in muscle cells just as for cTNT exon 5 (Fig. 3C and Table 1). Overall, the results in Figure 3 show that exons flanked by the natural cTNT introns but lacking the ESE still undergo preferential inclusion in muscle cultures. Therefore, the ESE is not required for the regulation of alternative splicing. The *cis* element(s) required for regulation presumably lie elsewhere within the pre-mRNA.

Exon 5 sequence is specifically required for recognition of the flanking 3′ splice site and not the flanking 5′ splice site. Mechanistically, exon recognition occurs via recognition of the flanking splice sites. We were interested in determining whether the ESE was required for splicing to the flanking 3′ splice site or to the flanking 5′ splice site in vivo. We have recently shown that exon mutations that lead to exon skipping in vivo specifically block splicing of the upstream intron in vitro but have no effect on splicing of the downstream intron (8). These results suggested that recognition of the 3′ splice site flanking exon 5 is disrupted by exon mutations. To address this question in vivo, the effects of exon mutations on the strengths of the flanking 3′ and 5′ splice sites were examined separately in *cis* competition assays. The 3′ splice site flanking exon 5 was forced to compete with the 3′ splice site flanking exon 6 by a deletion that removed the majority of intron 5 including the 5′ splice site (Fig. 4A). The intron between exons 5 and 6 can no longer be spliced, forcing exon 4 to combine with exon 5 or 6. Three variant constructs were made differing only in the sequence of exon 5: ΔBBC contains the unmodified exon 5, ΔBCE51 contains the E51 mutation in exon 5, and ΔBCPSR contains sTNI exon 2 in place of exon 5 (Fig. 2A). Following transfection into QT35 and primary skeletal muscle cultures, total RNA was assayed by S1 nuclease protection. The results indicated that when exon 5 was unmodified (ΔBBC), the 3′ splice sites flanking both exons 5 and 6 were used (Fig. 4B, lanes 2, 3, 8, and 9). However, when exon 5 contained the E51 mutation (ΔBCE51), exon 4 combined only with exon 6 (lanes 4 and 5). When exon 5 was replaced by sTNI exon 2 (ΔBCPSR), the flanking 3′ splice site was predominantly skipped (lanes 10 and 11). Therefore, when exon 5 contained either the E51 mutation or was replaced by a heterologous exon, the flanking 3′ splice site was predominantly skipped. These results are consistent with our previous in vitro results and demonstrate that the sequence within exon 5 plays a major role in the strength of the flanking 3′ splice site in vivo.

A similar approach was used to determine the role of exon 5 sequence in the strength of the flanking 5′ splice site. To use competing 5′ splice sites of equivalent strength, exon 5 and its flanking 5′ splice site were duplicated and placed in self-competition. By PCR cloning, a genomic fragment containing exon 5 and 24 bp of the downstream intron replaced exons 1 to 4 in ΔPB to construct μBRB (Fig. 5A). The intervening sequence between the duplicated exons cannot be spliced because of its small size and deletion of the 3′ end of the intron. Four minigenes were constructed that were identical except for the following modifications of the down-

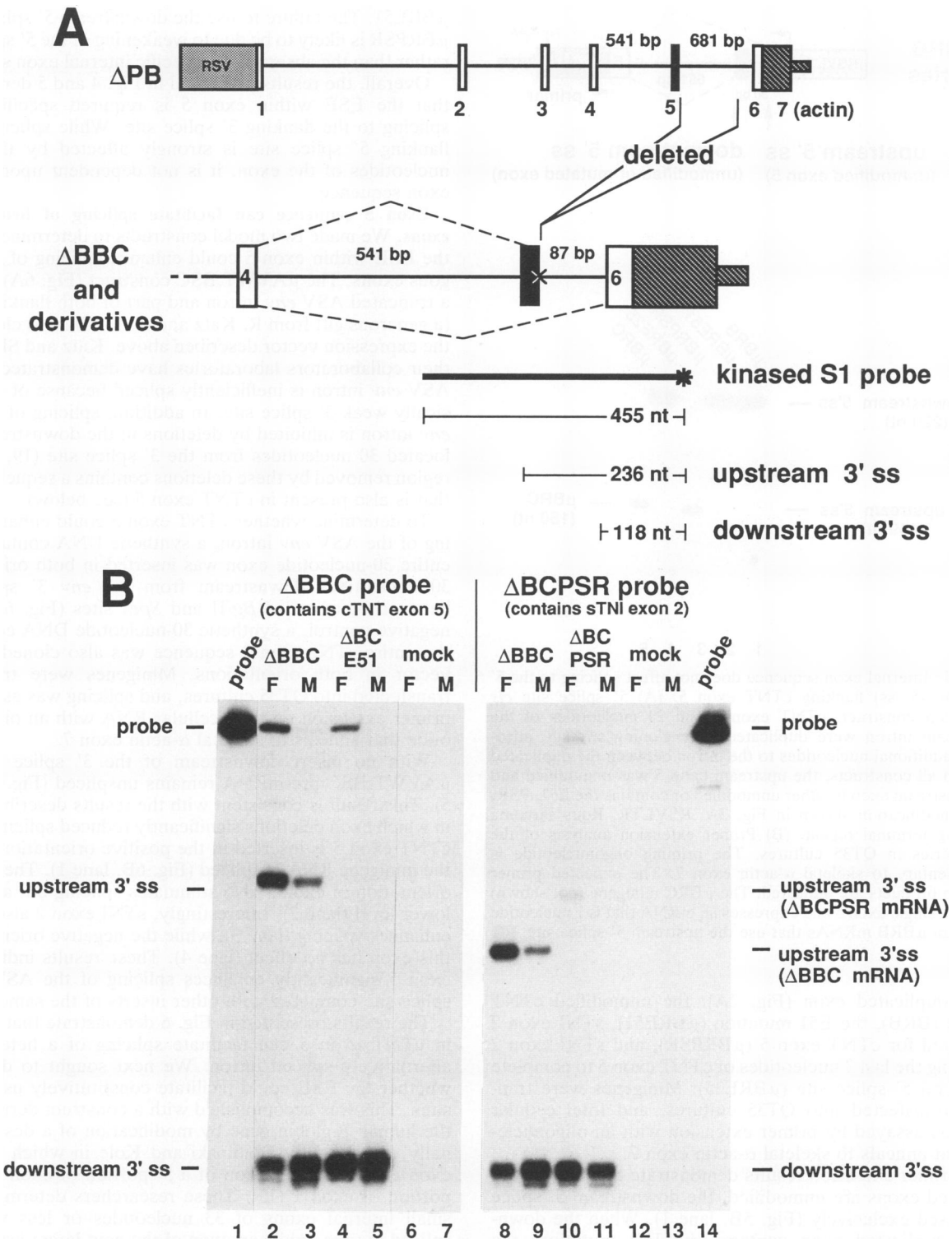

FIG. 4. Internal exon mutations disrupt splicing to the 3′ splice site (3′ ss) flanking cTNT exon 5. (A) Constructs forcing *cis* competition between the 3′ splice sites of exons 5 and 6. A 594-bp deletion in intron 5 removed the 5′ splice site and prevented splicing of exons 5 and 6. The three test constructs are identical except for the following modifications: ΔBBC contains the unmodified exon 5, ΔBCE51 contains the E51 mutation, and ΔBCPSR contains sTNI exon 2 in place of cTNT exon 5 (Fig. 2A). The probes used for S1 nuclease analysis were ^{32}P 5′-end-labeled *Ava*II fragments derived from ΔBBC (lanes 1 to 7) or ΔBCPSR (lanes 8 to 14). The labeled end is indicated by the asterisk. The *Ava*II sites are located 219 nucleotides upstream of exon 5 and 60 nucleotides downstream of the cTNT exon 6–α-actin exon 7 fusion. The diagrams are not drawn to scale. The expected protection products are indicated. RSV, Rous sarcoma virus. (B) S1 nuclease analysis. The protected products for the ΔBBC probe are indicated on the left and those for the ΔBCPSR probe are indicated on the right.

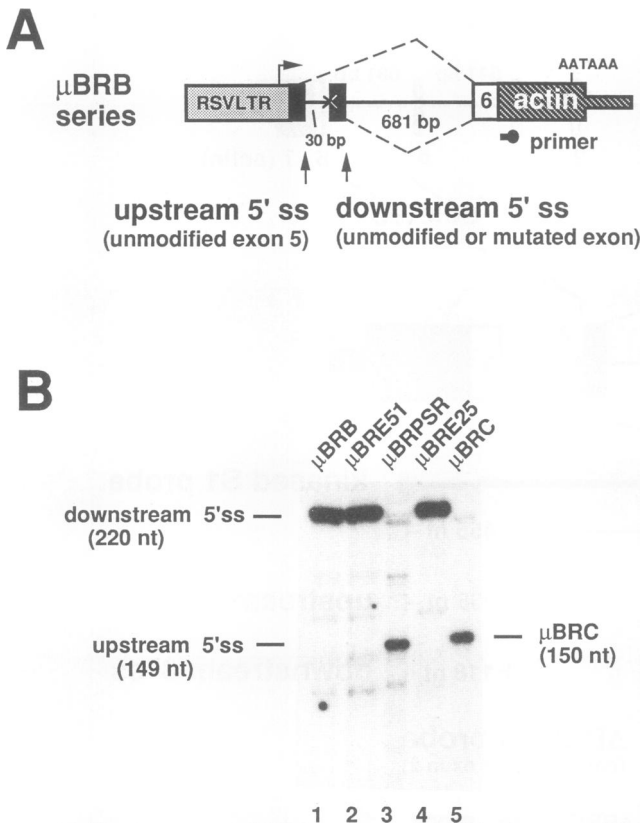

FIG. 5. Internal exon sequence does not affect splicing to the 5′ splice site (5′ ss) flanking cTNT exon 5. (A) 5′ splice site *cis* competition construct. cTNT exon 5 and 24 nucleotides of the downstream intron were duplicated. The cloning strategy introduced 6 additional nucleotides to the intron between the duplicated exons. In all constructs, the upstream exon 5 was unmodified and the downstream exon is either unmodified or contains the E51, PSR, or E25 modification shown in Fig. 2A. RSVLTR, Rous sarcoma virus long terminal repeat. (B) Primer extension analysis of the μBRB series in QT35 cultures. The priming oligonucleotide is complementary to skeletal α-actin exon 7. The expected primer extension products are indicated. The μBRC minigene (not shown) contains a single exon 5 and expresses an mRNA that is 1 nucleotide larger than μBRB mRNAs that use the upstream 5′ splice site.

stream duplicated exon (Fig. 2A): the unmodified cTNT exon 5 (μBRB), the E51 mutation (μBRE51), sTNI exon 2 substituted for cTNT exon 5 (μBRPSR), and sTNI exon 2 containing the last 3 nucleotides of cTNT exon 5 to complete the natural 5′ splice site (μBRE25). Minigenes were transiently transfected into QT35 cultures, and total cellular RNA was assayed by primer extension with an oligonucleotide that anneals to skeletal α-actin exon 7.

The primer extension results demonstrate that when both duplicated exons are unmodified, the downstream 5′ splice site is used exclusively (Fig. 5B, lane 1). When the downstream duplicated exon contains the E51 mutation, the downstream 5′ splice site is still used exclusively (μBRE51 [lane 2]). Substitution of the downstream exon 5 with sTNI exon 2 (μBRPSR) results in almost exclusive use of the upstream 5′ splice site (lane 3). However, when the last 3 nucleotides of sTNI exon 2 are changed to reestablish the complete downstream 5′ splice site, this site is used exclusively once again (μBRE25, lane 5). Therefore, the downstream 5′ splice site is used despite the absence of internal exon 5 sequence in μBRE25 and the E51 mutation in μBRE51. The failure to use the downstream 5′ splice site in μBRPSR is likely to be due to weakening of the 5′ splice site, rather than the absence of a specific internal exon sequence.

Overall, the results presented in Fig. 4 and 5 demonstrate that the ESE within exon 5 is required specifically for splicing to the flanking 3′ splice site. While splicing to the flanking 5′ splice site is strongly affected by the last 3 nucleotides of the exon, it is not dependent upon internal exon sequence.

Exon 5 sequence can facilitate splicing of heterologous exons. We made two model constructs to determine whether the ESE within exon 5 could enhance splicing of heterologous exons. The μAVWT.BSC construct (Fig. 6A) contains a truncated ASV *env* intron and part of both flanking exons (a generous gift from R. Katz and A. M. Skalka) cloned into the expression vector described above. Katz and Skalka and their collaborators laboratories have demonstrated that the ASV *env* intron is inefficiently spliced because of an intrinsically weak 3′ splice site. In addition, splicing of the ASV *env* intron is inhibited by deletions in the downstream exon located 30 nucleotides from the 3′ splice site (19, 29). The region removed by these deletions contains a sequence motif that is also present in cTNT exon 5 (see below).

To determine whether cTNT exon 5 could enhance splicing of the ASV *env* intron, a synthetic DNA containing the entire 30-nucleotide exon was inserted in both orientations 30 nucleotides downstream from the *env* 3′ splice site between blunt-ended *Bgl*II and *Spe*I sites (Fig. 6A). As a negative control, a synthetic 30-nucleotide DNA containing the entire sTNI exon 2 sequence was also cloned into the vector in both orientations. Minigenes were transiently transfected into QT35 cultures, and splicing was assayed by primer extension on total cellular RNA with an oligonucleotide that anneals to skeletal α-actin exon 7.

With no insert downstream of the 3′ splice site, the μAVWT.BSC pre-mRNA remains unspliced (Fig. 6B, lane 5). This result is consistent with the results described above in which exon deletions significantly reduced splicing. When cTNT exon 5 is inserted in the positive orientation, 40% of the minigene RNA is spliced (Fig. 6B, lane 1). The negative orientation of exon 5 also stimulates splicing but at a much lower level (lane 2). Interestingly, sTNI exon 2 also slightly enhances splicing (lane 3), while the negative orientation of this exon has no effect (lane 4). These results indicate that exon 5 significantly enhances splicing of the ASV *env* 3′ splice site compared with other inserts of the same size.

The results presented in Fig. 6 demonstrate that the ESE in cTNT exon 5 can facilitate splicing of a heterologous alternatively spliced intron. We next sought to determine whether the ESE could facilitate constitutively used splice sites. This was accomplished with a construct derived from the human β-globin gene by modification of a design originally described by Dominski and Kole in which a middle exon is created by fusion of a 5′ portion of exon 2 to a 3′ portion of exon 1 (15). These researchers determined that small internal exons of 33 nucleotides or less were not spliced in vivo, while an exon of the next larger size tested, 51 nucleotides, was spliced. The minimum exon size for exon recognition depends in part on the strength of the adjacent splice sites, since improving the pyrimidine tract, branch site, or 5′ splice site increases the level of inclusion of the 33-nucleotide internal exon (15, 16). Our experimental series was designed to determine whether the size limitation defined by these researchers could be overcome by incorporating the ESE of cTNT exon 5.

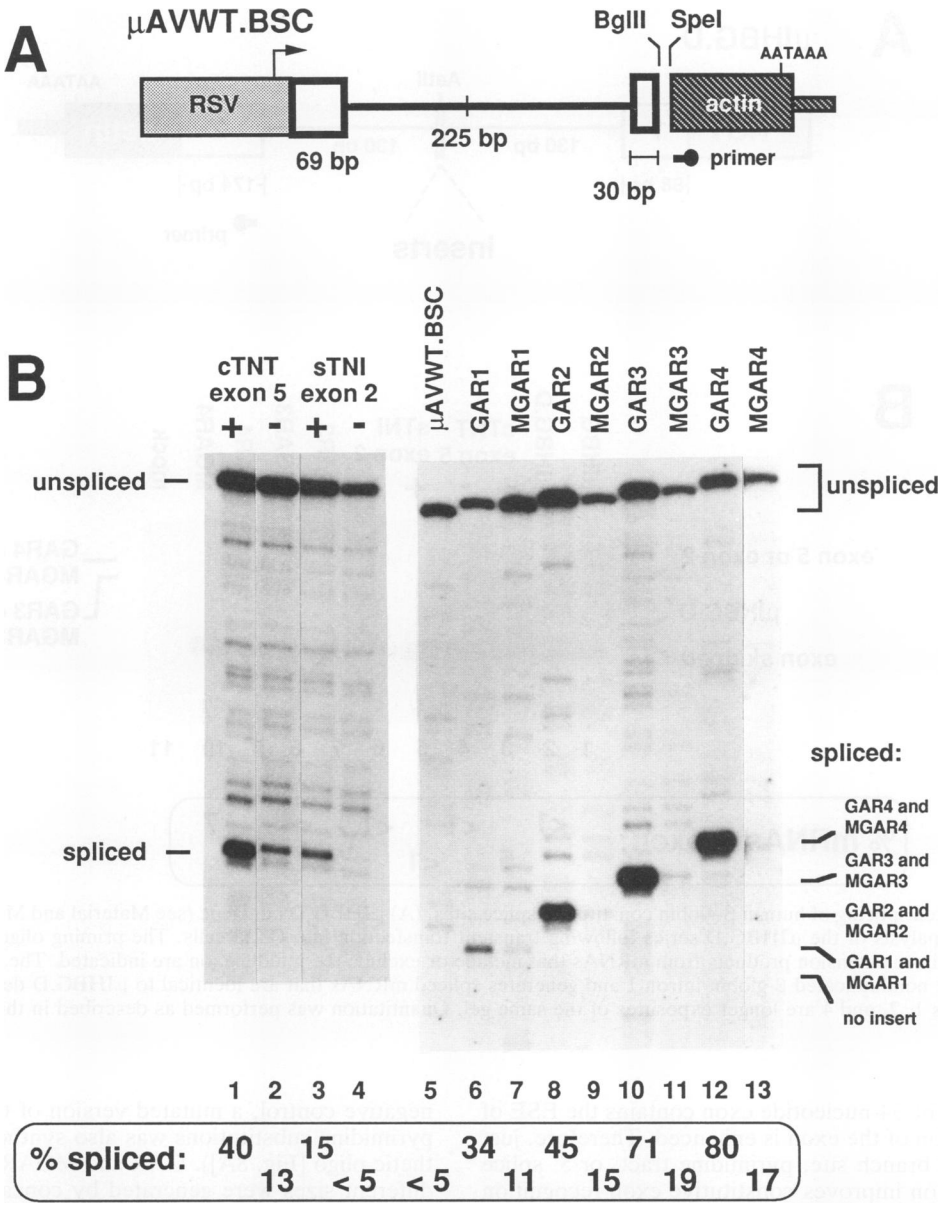

FIG. 6. Facilitated splicing of ASV *env* intron 3′ splice site. (A) Diagram of the ASWT.BSC test construct. A truncated ASV *env* intron and flanking exons (heavy lines) was cloned into a Rous sarcoma virus (RSV)-driven minigene (see Materials and Methods). Inserts to be tested for splicing-enhancing activity were cloned between *Bgl*II and *Spe*I sites located 30 nucleotides from the 3′ splice site. The diagram is not drawn to scale. (B) Primer extension analysis of μAVWT.BSC series following transient transfection into QT35 cells. The GAR and MGAR inserts are described in the text and shown in Fig. 8. The priming oligonucleotide anneals to α-actin exon 7. Extension products from unspliced and spliced RNAs are indicated. + and − indicate the orientations of cTNT exon 5 and sTNI exon 2. The cTNT exon 5 + construct is missing a single nucleotide because of incomplete filling in of the *Bgl*II site, making the primer extension product 1 nucleotide shorter than expected. Constructs that contain this nucleotide give the same result (data not shown). Lane 2 is a shorter exposure form the same gel, and lane 13 is a longer exposure of the same gel. Band intensities of unspliced and spliced RNA were quantified directly from the gel, and the percent spliced RNA is calculated as the ratio of spliced RNA to spliced plus unspliced RNA × 100.

The middle exon of the μIHBG.D minigene (Fig. 7A) was designed such that DNA fragments introduced into the middle exon would be flanked by the complete β-globin splice sites including the exon components (see Materials and Methods for details of cloning). Synthetic 30-nucleotide fragments containing the entire cTNT exon 5 or sTNI exon 2 were cloned into μIHBG.D in both orientations, generating four constructs each containing a 34-nucleotide middle exon. The constructs were transfected into QT35 cultures and assayed by primer extension with an oligonucleotide that anneals to human β-globin exon 2. Of the four test constructs, only the construct containing cTNT exon 5 in the positive orientation expressed significant levels of mRNAs that included the middle exon (Fig. 7B, lane 3). The level of these mRNAs is relatively low (5%), but this result has been obtained in repeated experiments. Our results are consistent with those of Dominski and Kole (15) in that three small exons were completely skipped by the splicing machinery.

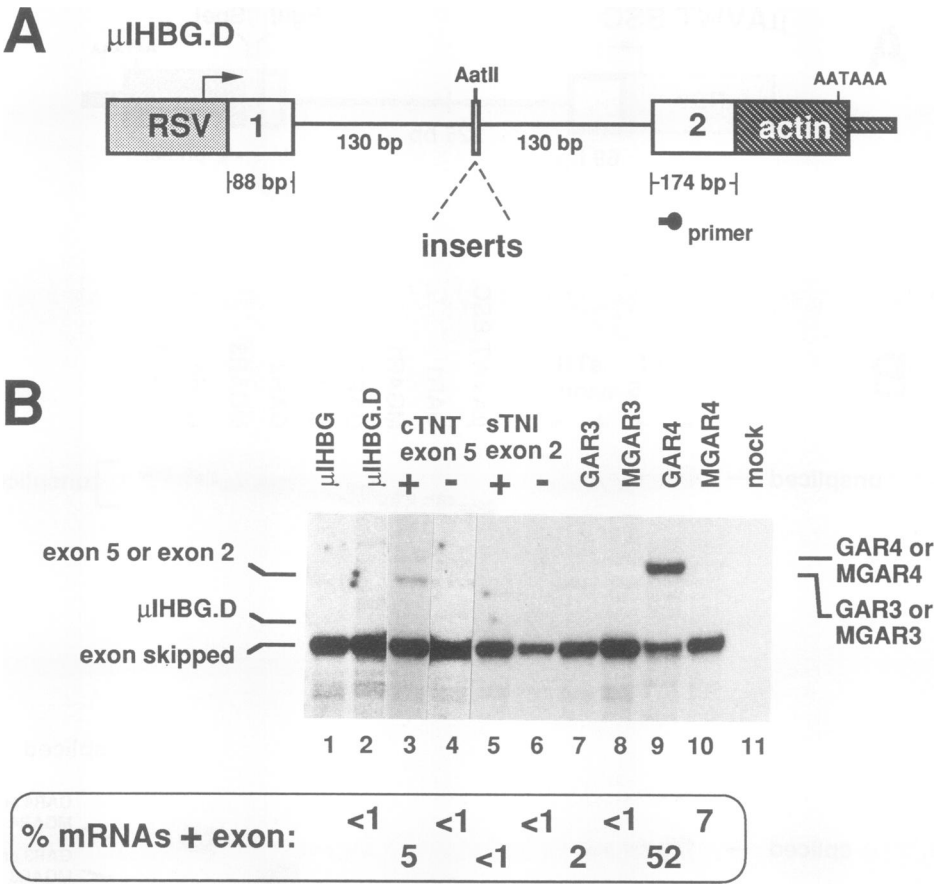

FIG. 7. Facilitated splicing of human β-globin constitutive splice sites. (A) μIHBG.D minigene (see Material and Methods for details). (B) Primer extension analysis of the μIHBG.D series following transient transfection into QT35 cells. The priming oligonucleotide anneals to β-globin exon 2. Primer extension products from mRNAs that include or exclude the middle exon are indicated. The μIHBG minigene (not shown) contains an nonduplicated β-globin intron 1 and generates spliced mRNAs that are identical to μIHBG.D derivatives that skip the middle exon. Lanes 1, 2, and 4 are longer exposures of the same gel. Quantitation was performed as described in the legend to Fig. 1B.

However, when the 34-nucleotide exon contains the ESE of exon 5, recognition of the exon is enhanced. Therefore, just as improving the branch site, pyrimidine tract, or 5′ splice site of a small exon improves constitutive exon recognition (15, 16), specific internal exon sequence can also improve constitutive exon recognition. These results demonstrate that the ESE within cTNT exon 5 can facilitate splicing to constitutive as well as alternative splice sites.

cTNT exon 5 contains a purine-rich motif that is found in heterologous exons within regions demonstrated to be required for splicing. Mutations of internal exon sequences within many different genes have a negative effect on splicing (see introduction). Figure 8A depicts the nucleotide sequences from five different exons that have been shown to contain regions required for efficient splicing. Mutations that disrupt splicing are underlined. A comparison of the sequences within these regions reveals a common purine-rich motif (shown in bold print). A common (but not universal) feature of this motif is the repeated 3-nucleotide sequence GAR where R is an A or G giving the variable common motif, GARGARGAR. In general, the exons contain more than one motif and motifs are variable in length.

To determine whether the purine-rich motif can enhance splicing, a synthetic motif (GAR synthetic oligo [Fig. 8A]) was cloned into μAVWT.BSC and μIHBG.D vectors. As a negative control, a mutated version of the motif containing pyrimidine substitutions was also synthesized (MGAR synthetic oligo [Fig. 8A]). GAR and MGAR DNA fragments of different sizes were generated by concatamerization of the synthetic DNA fragments (Fig. 8B). Double-stranded DNAs with one to four repeats of the synthetic oligonucleotide were isolated on gels and ligated into the μAVWT.BSC and μIHBG.D vectors. In designing the GAR and MGAR sequences, care was taken not to introduce sequence features that are known to affect splicing such as cryptic splice sites, pseudo splice sites, or termination codons (6, 41, 44, 52, 62).

μAVWT.BSC constructs containing one to four copies of the GAR or MGAR oligonucleotides were transfected into QT35 cultures and assayed by primer extension as described above. The results, shown in Fig. 6B, clearly demonstrated that the GAR motif significantly facilitated splicing of the ASV *env* intron (lanes 6, 8, 10, and 12), while the MGAR element had only a slight effect on splicing (lanes 7, 9, 11, and 13). Furthermore, the level of splicing increased as the number of GAR repeats increases. Therefore, the GAR motif enhances splicing to the flanking 3′ splice site in a sequence-specific and additive manner.

GAR and MGAR concatamers containing three and four copies of the synthetic motif were cloned into the blunt-ended *Aat*II site of μIHBG.D as described above giving

A

cTNT exon 5	ag AAGAGGAAGAATGGCTTGAGGAAGACGACG
ASV *env*	ag GCAGTTCTGATCGGATACCCTGGGGAGACGAGCAAGAAGGACTCCAAGAAGAAGCCGCCAGCAACAAGCAAGAAGGACC
β-tropomyosin exon 8	ag TATTCCACCAAAGAGGACAAATACGAAGAAGAGATCAAACTTCTGGAGGAGAAGCTGAAGGAGGC
fibronectin EDIIIA	ag -163 nt- CTACTCGAGCCCTGAGGATGGAATCCATGAGCTATTCCCTGCACCTGATGGTGAAGAAGACAC
growth hormone exon 5	ag -115 nt- CTTCCGGAAGGACCTGCATAAGACGGAGACGTACCTGAGGG

common motif	**GARGARGAR**
GAR synthetic oligo	**GAAGAGGAG**
MGAR synthetic oligo	**CAATACGAG**

B

No. of Repeats	Insert Size	GAR	MGAR
1	10 nt	GAAGAGGAGG	CAATACGAGC
2	19 nt	GAAGAGGAGGAAGAGGAGG	CAATACGAGCAATACGAGC
3	28 nt	GAAGAGGAGGAAGAGGAGGAAGAGGAGG	CAATACGAGCAATACGAGCAATACGAGC
4	37 nt	GAAGAGGAGGAAGAGGAGGAAGAGGAGGAAGAGGAGG	CAATACGAGCAATACGAGCAATACGAGCAATACGAGC

FIG. 8. A purine-rich motif within exon regions that are required for splicing. (A) Sequence comparison of five exons demonstrated to contain regions that are required for splicing. The sequence begins with the AG dinucleotide of the upstream intron. For fibronectin EDIIIA and growth hormone exon 5, the distances from the 3′ splice sites are indicated. The underlined nucleotides indicate substitutions or deletions that disrupt splicing (cTNT exon 5 [12 and this report], ASV *env* [29], β-tropomyosin exon 8 [25], fibronectin exon EDIIIA [37], and bovine growth hormone [23]). A common purine-rich motif found within these regions is shown in bold print. Synthetic DNAs containing a 9-bp copy of a purine-rich motif (GAR) or a mutated version (MGAR) were synthesized. The underlined nucleotides in the MGAR oligonucleotide (oligo) indicate substitutions relative to GAR. (B) Double-stranded concatamers containing one to four copies of the GAR or MGAR synthetic oligonucleotides were isolated for cloning. The sequences (sense strand only) and sizes of the inserts are shown.

middle exons of 32 and 41 nucleotides, respectively. DNAs were transfected into QT35 cultures, and total cellular RNA were isolated for primer extension analysis. Figure 7B shows that exon recognition is not significantly facilitated by either GAR3 or MGAR3. In particularly good transfections, however, a low level of GAR3 but not MGAR3 exon inclusion can be detected (data not shown). Overall, facilitated splicing by the GAR3 insert is weak at best. In contrast, exons containing the GAR4 insert was included in 52% of the mRNAs (Fig. 7B, lane 9), while exons containing the MGAR insert were predominantly skipped (7% inclusion [lane 10]). Therefore, the GAR motif facilitated use of the constitutive splice sites in this construct in a sequence-specific manner. It is not clear whether the different results from GAR3 and GAR4 are due to differences in exon size or the number of GAR motifs or both.

DISCUSSION

A purine-rich element within cTNT exon 5 serves as a general splicing signal. During the course of investigating pre-mRNA *cis* elements involved in regulated alternative splicing of cTNT exon 5, we have identified a novel splicing element located within the exon. We refer to this internal exon sequence as an ESE. The experimental evidence for the ESE is as follows. (i) Nucleotide substitutions within the exon that are separate from the exon portions of the splice sites result in exon skipping in a variety of cell types. (ii) The splice sites that normally flank cTNT exon 5 are not sufficient for splicing of a heterologous constitutive exon of the same size. (iii) Exon 5 contains a purine-rich motif found in several exons within regions that have been shown to be required for splicing. (iv) cTNT exon 5 sequence or a synthetic element containing the purine-rich motif facilitates splicing in two different test genes.

Although the purine-rich ESE was identified in an developmentally regulated alternative exon in a striated muscle-specific gene product, we propose that this element is not limited to alternative exons or muscle-specific genes on the basis of the following observations. (i) The ESE is required for recognition of exon 5 in both muscle and nonmuscle cells. (ii) cTNT exon 5 sequence or a synthetic purine-rich motif both facilitates splicing in nonmuscle cells and is capable of facilitating splicing of the constitutive splice sites of human β-globin intron 1. (iii) The putative ESEs shown in Fig. 8 are within exons of genes that are expressed in a variety of cell types. (iv) cTNT exon 6, a constitutive exon, also contains a purine-rich region in a 32-nucleotide segment of the exon that is required for splicing to the flanking 3′ splice site (67a). (v) The ESE within cTNT exon 5 is not required for regulation of alternative splicing, suggesting that it serves a more general splicing function. Therefore, we propose that the ESE in cTNT exon 5 serves as a general

splicing *cis* element that is recognized by the constitutive splicing machinery.

An important feature of the ESE described here is that it is internal to the first and last 3 nucleotides of the exon and therefore is distinct from the exon components of the 3′ and 5′ splice sites. The terminal exon nucleotides have been demonstrated to contribute to splice site strength (17, 33, 39), and spontaneous mutations in these nucleotides can disrupt splicing, generally leading to exon skipping in vertebrates (21, 40, 64, 66). Presumably these nucleotides are recognized as components of the adjacent splice sites by the splicing machinery. Indeed, 5′ splice site strength was increased when the last 3 nucleotides of an exon were modified to increase potential base pairing with U1 small nuclear RNA (compare splicing of ΔPSR and ΔE25 to ΔE2CON [Figs. 2 and 3] and μBRPSR to μBRE25 [Fig. 5]). It should also be noted that the exon 5 ESE does not contain sequences known to affect splicing, such as cryptic splice sites, pseudo splice sites, or stop codons (6, 41, 44, 52, 62). Therefore, the ESE within cTNT exon 5 represents a previously undescribed *cis* splicing element.

Several pieces of evidence suggest that positive-acting (and perhaps negative-acting) ESEs may not be uncommon. The deficiencies of the known splicing elements and the large number of studies that support a role for exon sequence in splice site selection have been discussed in the introduction. The demonstration that selection of a splice site depends upon its context (44) is also consistent with a role for exon sequence in splice site selection. In addition, characterization of spontaneous or mutagen-induced mutations in the endogenous hypoxanthine-guanine phosphoribosyltransferase (*hprt*) gene have revealed a surprising number of internal exon mutations that resulted in partial or complete exon skipping (47, 55, 68). The effects of many of these mutations cannot be explained by the introduction of cryptic splice sites, pseudo splice sites, or termination codons (6, 41, 44, 52, 62). These results and the examples shown in Fig. 8 suggest the possibility that a number of alternative and constitutive exons contain internal *cis* elements that affect splice site strength. We expect that the exons most likely to contain positive-acting splicing elements are those that are intrinsically poor splicing substrates because of small exon size or weak flanking splice sites.

Initially, sTNI exon 2 was thought to lack an ESE, since this exon was completely skipped when substituted for cTNT exon 5 (Fig. 2C). The sTNI exon is flanked by relatively strong splice sites, which accounts for the constitutive splicing of this exon in the endogenous gene (45) and in construct ΔDMR (Fig. 2B). Interestingly, this exon weakly enhanced splicing of the ASV *env* intron in the positive orientation but not in the negative orientation (Fig. 6B). Whether this result is indicative of a weak ESE is not clear.

By what mechanism does the cTNT exon 5 ESE facilitate splicing? While the specific mechanism remains unknown, the target for ESE activity has been localized to the flanking 3′ splice site by the in vivo *cis* competition experiments reported here (Figs. 4 and 5) and our previous in vitro results (8). In addition, mutations in three of the other four putative ESEs shown in Fig. 8A (ASV *env*, bovine growth hormone, and β-tropomyosin [23, 25, 29]) decreased the splicing efficiency of an upstream intron, suggesting a potential link between the purine-rich ESE and facilitated splicing to the flanking 3′ splice site. We are currently investigating the role of the ESE in 3′ splice site recognition in vitro. Preliminary results suggest that the ESE within exon 5 serves as a binding site for *trans*-acting factors and binding is required for assembly of the prespliceosome complex on the flanking 3′ splice site (46a).

We expect that the mechanism of ESE-mediated facilitation of splicing differs from that promoting cooperation between back-to-back 3′ splice sites described in three different genes (20, 35, 69). In all three reported cases, the downstream AG is part of a bona fide 3′ splice site containing a consensus YAG and an associated polypyrimidine tract. In contrast, the sequence of the ESE does not resemble a 3′ splice site, since no AG dinucleotides are preceded by a pyrimidine and there is only one polypyrimidine tract which is upstream of the authentic 3′ splice site.

cTNT exon 5 contains two purine-rich regions separated by a few central pyrimidines. Nucleotide substitutions within the purine-rich 5′ or 3′ region of the exon resulted in complete or partial exon skipping, respectively (mutants E52 and E53 [12] [Fig. 2]). Interestingly, the E51 mutation changes only the central pyrimidines (Fig. 2A). While the results with this mutation (Figs. 2C and 4B) raise the possibility that the central pyrimidines perform a role in exon 5 ESE function, synthetic elements containing only purines are capable of facilitating splicing (Fig. 6 and 7). It is currently unclear whether the E51 mutation replaces nucleotides that are required for optimal function or substitutes nonessential pyrimidines with a sequence (such as four consecutive G's) that somehow inactivates the ESE. Clearly, the optimal sequence of the ESE remains to be defined.

cTNT exon 5 is flanked by a nonconsensus 5′ splice site that is functionally weak (12). In contrast, the flanking 3′ splice site appears to contain strong splicing elements: the branch site maps to an adenosine located 38 nucleotides from the 3′ splice site in a 6- of 7-nucleotide match to the consensus sequence, and immediately upstream of the 3′ splice site are 22 uninterrupted pyrimidines (8). Why, then, does this 3′ splice site require a facilitating element located within the exon? One possibility is that the 3′ splice site is not as strong as its sequence suggests. Another possibility is that the ESE contributes to exon definition rather than to 3′ splice site definition. The exon definition model proposes that the factors that recognize splice sites serve to define exons prior to intron removal (49). Several laboratories have demonstrated that the strength of a 5′ splice site affects recognition of the upstream intron consistent with a model proposing interactions across the exon (31, 43, 59). The ESE within exon 5 could serve to compensate for the weak 5′ splice site by stabilizing the interactions across the exon. Alternatively, the ESE could substitute for the 5′ splice site during exon definition. Once exon 5 is defined, the 5′ splice site would be selected for cleavage.

As mentioned above, mutations within the first nucleotide or last 3 nucleotides of an exon can disrupt splicing. Therefore, these coding positions are somewhat restricted by the requirements for pre-mRNA splicing. Our results indicate that the overlap between information for RNA processing and protein coding can extend to internal portions of exons. It is not uncommon to find overlapping information within a single RNA that codes for different processes (see, for example, references 27, 32, and 63). The identification of a splicing element within a coding exon has interesting implications regarding the potential restrictions that RNA processing puts on protein evolution.

We initially studied the ESE within exon 5 with the expectation that this novel splicing element would be involved in the regulation of alternative splicing. However,

our results strongly suggest that the ESE serves as a general splicing element that facilitates exon recognition but is not required for modulating the ratio of the two alternative splicing pathways. The *cis* elements that serve as the primary targets for regulatory factors must therefore lie outside of the exon.

ACKNOWLEDGMENTS

We thank Sue Berget and members of her laboratory for helpful discussions and Sue Berget and Janet Mar for critical reviews of the manuscript. We thank Reena Kannan and Jackie Ramchatesingh for important contributions to this project.

J.T. was supported by the SMART program at Baylor College of Medicine. This work was supported by a Basil O'Conner March of Dimes Starter Scholar Research Award and an NIH First Award (T.A.C.). T.A.C. is an Established Investigator of the American Heart Association.

ADDENDUM IN PROOF

While this article was in review, the characterization of a purine-rich exon element in the M2 exon of the IgM gene was reported (A. Watakabe, K. Tanaka, and Y. Shimura, Genes Dev. 7:407–418, 1993).

REFERENCES

1. **Adami, G. R., and G. G. Carmichael.** 1987. The length but not the sequence of the polyoma virus later leader exon is important for both late RNA splicing and stability. Nucleic Acids Res. **15**:2593–2610.
2. **Black, D. L.** 1991. Does steric interference between splice sites block the splicing of a short c-src neuron-specific exon in non-neuronal cells? Genes Dev. **5**:389–402.
3. **Brunak, S., J. Engelbrecht, and S. Knudsen.** 1991. Prediction of human mRNA donor and acceptor sites from the DNA sequence. J. Mol. Biol. **220**:49–65.
4. **Chain, A. C., S. Zollman, J. C. Tseng, and F. A. Laski.** 1991. Identification of a *cis*-acting sequence required for germ line-specific splicing of the P element ORF2-ORF3 intron. Mol. Cell. Biol. **11**:1538–1546.
5. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7**:2745–2752.
6. **Cheng, J., P. M. Fogel, and L. E. Maquat.** 1990. Translation to near the distal end of the penultimate exon is required for normal levels of spliced triosephosphate isomerase mRNA. Mol. Cell. Biol. **10**:5215–5225.
7. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-choroform extraction. Anal. Biochem. **162**:156–159.
8. **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.
9. **Cooper, T. A., M. H. Cardone, and C. P. Ordahl.** 1988. Cis requirements for alternative splicing of the cardiac troponin T pre-mRNA. Nucleic Acids Res. **16**:8443–8465.
10. **Cooper, T. A., and C. P. Ordahl.** 1984. A single cardiac troponin T gene regulated by different programs in cardiac and skeletal muscle development. Science **226**:979–982.
11. **Cooper, T. A., and C. P. Ordahl.** 1985. A single cardiac troponin T gene generates embryonic and adult isoforms via developmentally regulated alternate splicing. J. Biol. Chem. **260**:11140–11148.
12. **Cooper, T. A., and C. P. Ordahl.** 1989. Nucleotide substitutions within the cardiac troponin T alternative exon disrupt pre-mRNA alternative splicing. Nucleic Acids Res. **17**:7905–7921.
13. **Cote, G. J., D. T. Stolow, S. Peleg, S. M. Berget, and R. F. Gagel.** 1992. Identification of exon sequences and an exon binding protein involved in alternative RNA splicing of calcitonin/CGRP. Nucleic Acids Res. **20**:2361–2366.
14. **Domenjoud, L., H. Gallinaro, L. Kister, S. Meyer, and M. Jacob.** 1991. Identification of a specific exon sequence that is a major determinant in the selection between a natural and a cryptic 5′ splice site. Mol. Cell. Biol. **11**:4581–4590.
15. **Dominski, Z., and R. Kole.** 1991. Selection of splice sites in pre-mRNAs with short internal exons. Mol. Cell. Biol. **11**:6075–6083.
16. **Dominski, Z., and R. Kole.** 1992. Cooperation of pre-mRNA sequence elements in splice site selection. Mol. Cell. Biol. **12**:2108–2114.
17. **Eperon, L. P., J. P. Estibeiro, and I. C. Eperon.** 1986. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. Nature (London) **324**:280–282.
18. **Fornwald, J. A., G. Kuncio, I. Peng, and C. P. Ordahl.** 1982. The complete nucleotide sequence of the chick α-actin gene and its evolutionary relationship to the actin gene family. Nucleic Acids Res. **10**:3861–3876.
19. **Fu, X. D., R. A. Katz, A. M. Skalka, and T. Maniatis.** 1991. The role of branchpoint and 3′-exon sequences in the control of balanced splicing of avian retrovirus RNA. Genes Dev. **5**:211–220.
20. **Ge, H., J. Noble, J. Colgan, and J. L. Manley.** 1990. Polyoma virus small tumor antigen pre-mRNA splicing requires cooperation between two 3′ splice sites. Proc. Natl. Acad. Sci. USA **87**:3338–3342.
21. **Grandchamp, B., C. Picat, F. de Rooij, C. Beaumont, P. Wilson, J. C. Deybach, and Y. Nordmann.** 1989. A point mutation G→A in exon 12 of the porphobilinogen deaminase gene results in exon skipping and is responsible for acute intermittent porphyria. Nucleic Acids Res. **17**:6637–6649.
22. **Guthrie, C.** 1991. Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. Science **253**:157–163.
23. **Hampson, R. K., L. L. Follette, and F. M. Rottman.** 1989. Alternative processing of bovine growth hormone mRNA is influenced by downstream exon sequence. Mol. Cell. Biol. **9**:1604–1610.
24. **Hedley, M. L., and T. Maniatis.** 1991. Sex-specific splicing and polyadenylation of dsx pre-mRNA requires a sequence that binds specifically to tra-2 protein in vitro. Cell **65**:579–586.
25. **Helfman, D. M., W. M. Ricci, and L. A. Finn.** 1988. Alternative splicing of tropomyosin pre-mRNAs in vitro and in vivo. Genes Dev. **2**:1627–1628.
26. **Hoshijima, K., K. Inoue, I. Higuchi, H. Sakamoto, and Y. Shimura.** 1991. Control of doublesex alternative splicing by transformer and transformer-2 in Drosophila. Science **252**:833–836.
27. **Jacks, T., H. D. Madhani, F. R. Masiarz, and H. E. Varmus.** 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. Cell **55**:447–458.
28. **Kakizuka, A., T. Ingi, T. Murai, and S. Nakanishi.** 1990. A set of U1 snRNA-complementary sequences involved in governing alternative RNA splicing of the kininogen genes. J. Biol. Chem. **265**:10102–10108.
29. **Katz, R. A., and A. M. Skalka.** 1990. Control of retroviral RNA splicing through maintenance of suboptimal processing signals. Mol. Cell. Biol. **10**:696–704.
30. **Krainer, A. R., and T. Maniatis.** 1988. RNA splicing, p. 131–207. *In* B. D. Hames and D. M. Glover (ed.), Frontiers in molecular biology: transcription and splicing. IRL Press, Oxford.
31. **Kreivi, J. P., K. Zefrivitz, and G. Akusjarvi.** 1991. A U1 snRNA binding site improves the efficiency of in vitro pre-messenger RNA splicing. Nucleic Acids Res. **19**:6956.
32. **Lamb, R. A., and C. M. Horvath.** 1991. Diversity of coding strategies in influenza virus. Trends Genet. **7**:261–266.
33. **Lear, A. L., L. P. Eperon, I. M. Wheatley, and I. C. Eperon.** 1990. Hierarchy for 5′ splice site preference determined in vivo. J. Mol. Biol. **211**:103–115.
34. **Libri, D., M. Goux-Pelletan, E. Brody, and M. Y. Fiszman.** 1990. Exon as well as intron sequences are *cis*-regulating elements for the mutually exclusive alternative splicing of the β tropomyosin gene. Mol. Cell. Biol. **10**:5036–5046.
35. **Ligtenberg, M. J., A. M. Gennissen, H. L. Vos, and J. Hilkens.**

1990. A single nucleotide polymorphism in an exon dictates allele dependent differential splicing of episialin mRNA. Nucleic Acids Res. **19**:297–301.
36. **Mar, J. H., P. B. Antin, T. A. Cooper, and C. P. Ordahl.** 1988. Analysis of the upstream regions governing expression of the chicken cardiac troponin T gene in embryonic cardiac and skeletal muscle cells. J. Cell Biol. **107**:573–585.
37. **Mardon, H. J., G. Sebastio, and F. E. Barralle.** 1987. A role for exon sequence in alternative splicing of the human fibronectin gene. Nucleic Acids Res. **15**:7725–7733.
38. **Matsuo, M., T. Masumura, H. Nishio, T. Nakajima, Y. Kitoh, T. Takumi, J. Koga, and H. Nakamura.** 1991. Exon skipping during splicing of dystrophin mRNA precursor due to an intraexon deletion in the dystrophin gene of Duchenne muscular dystrophy kobe. J. Clin. Invest. **87**:2127–2131.
39. **Mineo, I., and E. W. Holmes.** 1991. Exon recognition and nucleocytoplasmic partitioning determine AMPD1 alternative transcript production. Mol. Cell. Biol. **11**:5356–5363.
40. **Mitchell, P. J., G. Urlaub, and L. Chasin.** 1986. Spontaneous splicing mutations at the dihydrofolate reductase locus in Chinese hamster ovary cells. Mol. Cell. Biol. **6**:1926–1935.
41. **Naeger, L. K., R. V. Schoborg, Q. H. Zhao, G. E. Tullis, and D. J. Pintel.** 1992. Nonsense mutations inhibit splicing of MVM RNA in cis when they interrupt the reading frame of either exon of the final spliced product. Genes Dev. **6**:1107–1119.
42. **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.
43. **Nasim, F. H., P. A. Spears, H. M. Hoffmann, H. C. Kuo, and P. J. Grabowski.** 1990. A sequential splicing mechanism promotes selection of an optimal exon by repositioning a downstream 5′ splice site in preprotachykinin pre-mRNA. Genes Dev. **4**:1172–1184.
44. **Nelson, K. K., and M. R. Green.** 1988. Splice site selection and ribonucleoprotein complex assembly during in vitro pre-mRNA splicing. Genes Dev. **2**:319–329.
45. **Nikovits, W., G. Kuncio, and C. P. Ordahl.** 1986. The chicken fast skeletal troponin I gene: exon organization and sequence. Nucleic Acids Res. **14**:3377–3390.
46. **Ohshima, Y., and Y. Gotoh.** 1987. Signals for the selection of a splice site in pre-mRNA. J. Mol. Biol. **195**:247–259.
46a. **Ramchatesingh, J., and T. Cooper.** Unpublished data.
47. **Recio, L., D. Simpson, J. Cochrane, H. Liber, and T. R. Skopek.** 1990. Molecular analysis of hprt mutants induced by 2-cyanoethylene oxide in human lymphoblastoid cells. Mutat. Res. **242**:195–208.
48. **Reed, R., and T. Maniatis.** 1986. A role for exon sequence and splice-site proximity in splice-site selection. Cell **46**:681–690.
49. **Robberson, B. L., G. J. Cote, and S. M. Berget.** 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell. Biol. **10**:84–94.
50. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
51. **Senapathy, P., M. B. Shapiro, and N. L. Harris.** 1990. Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. Methods Enzymol. **183**:252–278.
52. **Siebel, C. W., and D. C. Rio.** 1990. Regulated splicing of the Drosophila P transposable element third intron in vitro: somatic repression. Science **248**:1200–1208.
53. **Smith, C. W. J., J. G. Patton, and B. N. Ginard.** 1989. Alternative splicing in the control of gene expression. Annu. Rev. Genet. **23**:527–577.
54. **Somasekhar, M. B., and J. E. Mertz.** 1985. Exon mutations that affect the choice of splice sites used in processing the SV40 late transcripts. Nucleic Acids Res. **13**:5591–5609.
55. **Steingrimsdottir, H., G. Rowley, G. Dorado, J. Cole, and A. R. Lehmann.** 1992. Mutations which alter splicing in the human hypoxanthine-guanine phosphoribosyltransferase gene. Nucleic Acids Res. **20**:1201–1208.
56. **Steitz, J. A., D. L. Black, V. Gerke, K. A. Parker, A. Kramer, D. Frendwey, and W. Keller.** 1988. Function of the abundant U-snRNPs, p. 115–153. *In* M. L. Birnstiel (ed.), Structure and function of the major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Berlin.
57. **Streuli, M., and H. Saito.** 1989. Regulation of tissue-specific alternative splicing: exon-specific cis-elements govern the splicing of leukocyte common antigen pre-mRNA. EMBO J. **8**:787–796.
58. **Strohman, R. C., E. Bayne, D. Spector, T. Obinata, J. Micou-Eastwood, and A. Maniotis.** 1990. Myogensis and histogenesis of skeletal muscle on flexible membranes in vitro. In Vitro Cell. Dev. Biol. **26**:201–208.
59. **Talerico, M., and S. M. Berget.** 1990. Effect of 5′ splice site mutations on splicing of the preceding intron. Mol. Cell. Biol. **10**:6299–6305.
60. **Tian, M., and T. Maniatis.** 1992. Positive control of pre-messenger-RNA splicing in vitro. Science **256**:237–240.
61. **Tsai, A. Y., M. Streuli, and H. Saito.** 1989. Integrity of the exon 6 sequence is essential for tissue-specific alternative splicing of human leukocyte common antigen pre-mRNA. Mol. Cell. Biol. **9**:4550–4555.
62. **Urlaub, G., P. J. Mitchell, C. J. Ciudad, and L. A. Chasin.** 1989. Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. Mol. Cell. Biol. **9**:2868–2880.
63. **Vellard, M., A. Sureau, J. Soret, C. Martinerie, and B. Perbal.** 1992. A potential splicing factor is encoded by the opposite strand of the trans-spliced c-myb exon. Proc. Natl. Acad. Sci. USA **89**:2511–2515.
64. **Vidaud, M., R. Gattoni, J. Stevenin, D. Vidaud, S. Amselem, J. Chibani, J. Rosa, and M. Goossens.** 1989. A 5′ splice-region G→C mutation in exon 1 of the human β-globin gene inhibits pre-mRNA splicing: a mechanism for β+ thalassemia. Proc. Natl. Acad. Sci. USA **86**:1041–1045.
65. **Wakamatsu, N., H. Kobayashi, T. Miyatake, and S. Tsuji.** 1992. A novel exon mutation in the human beta-hexosaminidase beta subunit gene affects 3′ splice site selection. J. Biol. Chem. **267**:2406–2413.
66. **Weil, D., M. D'Alessio, F. Ramirez, W. de Wet, W. G. Cole, D. Chan, and J. F. Bateman.** 1989. A base substitution in the exon of a collagen gene causes alternative splicing and generates a structurally abnormal polypeptide in a patient with Ehlers-Danlos syndrome type VII. EMBO J. **8**:1705–1710.
67. **Xie, W., and L. I. Rothblum.** 1991. Rapid, small-scale RNA isolation from tissue culture cells. BioTechniques **11**:325–327.
67a. **Xu, R., and T. A. Cooper.** Unpublished data.
68. **Zhang, L. H., H. Vrieling, Z. A. A. van Zeeland, and D. Jenssen.** 1992. Spectrum of spontaneously occurring mutations in the hprt gene of V79 Chinese hamster cells. J. Mol. Biol. **223**:627–635.
69. **Zhuang, Y., and A. M. Weiner.** 1990. The conserved dinucleotide AG of the 3′ splice site may be recognized twice during in vitro splicing of mammalian mRNA precursors. Gene **90**:263–269.