Splice-Junction Elements and Intronic Sequences Regulate Alternative Splicing of the Drosophila Myosin Heavy Chain Gene Transcript

David M. Standiford,[1] Mary Beth Davis,[1] Weitao Sun and Charles P. Emerson, Jr.

Department of Cell and Developmental Biology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6058

Manuscript received April 30, 1997
Accepted for publication July 7, 1997

ABSTRACT

The Drosophila muscle myosin heavy chain (*Mhc*) gene primary transcript contains five alternatively spliced exon groups (exons 3, 7, 9, 11 and 15), each of which contains two to five mutually exclusive members. Individual muscles typically select a specific alternative exon from each group for incorporation into the processed message. We report here on the *cis*-regulatory mechanisms that direct the processing of alternative exons in *Mhc* exon 11 in individual muscles using transgenic reporter constructs, RT-PCR and directed mutagenesis. The 6.0-kilobase exon 11 domain is sufficient to direct the correct processing of exon 11 alternatives, demonstrating that the alternative splicing *cis*-regulatory elements are local to *Mhc* exon 11. Mutational analysis of *Mhc* exon 11 reveals that the alternative exon nonconsensus 5′-splice donors are essential for alternative splicing regulation in general, but do not specify alternative exons for inclusion in individual muscles. Rather, we show, through exon substitutions and deletion analyses, that a 360-nucleotide intronic domain precisely directs the normal processing of one exon, *Mhc* exon 11e, in the indirect flight muscle. These and other data indicate that alternative exons are regulated in appropriate muscles through interactions between intronic alternative splice-specificity elements, nonconsensus exon 11 splice donors and, likely, novel exon-specific alternative splicing factors.

THE development of skeletal muscle follows a multistep pathway beginning with the specification of the mesoderm and ending with the elaboration of the cellular components of the contractile apparatus. The myogenic regulatory factors (MRFs) that function during the early part of this process have been described in some detail and their activity is clearly required for the correct development of all skeletal muscle (EMERSON 1993; ABMAYR *et al.* 1995). Further, muscle consists of many different fiber types, each with differing contractile and physiological properties (FYRBERG and BEALL 1990; BANDMAN 1992). Such functional differences among fiber types are determined largely by the particular isoforms of the contractile proteins expressed in that fiber type and this expression is often influenced developmentally or physiologically. Thus, the contractile behavior of muscle can be dynamically regulated through the modification of contractile protein isoform expression. Despite its central role in governing the function of muscle, little is yet known about the mechanisms that define, coordinate and enforce the expression of the appropriate contractile protein isoforms in specific muscles. It is clear, however, that one general mechanism for fiber-type isoform production involves alternative pre-mRNA processing, which regulates the isoform expression of many of the major contractile proteins, including myosin, troponin and tropomyosin (EPSTEIN and BERNSTEIN 1992; BERNSTEIN *et al.* 1993; HODGES and BERNSTEIN 1994; SCHIAFFINO and REGGIANI 1996). An understanding of alternative splicing in muscle, then, can lead to a fuller picture of the mechanisms through which different fiber types are specified.

Several investigations have examined alternative splicing in muscle and have focused largely on the α- and β-tropomyosin and troponin T transcripts in vertebrates. These studies, using splicing extracts or transfected cell lines, have identified several mechanisms through which splicing regulation can occur. For instance, the exclusion of the human α-tropomyosin skeletal muscle alternative exon SK in nonmuscle cells is negatively regulated through the presence of a weak 3′ splice acceptor and an exonic repressive element (GRAHAM *et al.* 1992). Similarly, the chicken β-tropomyosin exon 6B is also repressed in nonmuscle cells, but this is mediated through RNA secondary structure (CLOUET D'ORVAL *et al.* 1991; LIBRI *et al.* 1991). In contrast, the rat cardiac troponin-T alternative exon 5 is positively regulated for inclusion into embryonic transcripts through the activity of several intronic sequences in addition to a general exonic enhancer (XU *et al.* 1993; RYAN and COOPER 1996). Biochemical descriptions of these processes have depended on the components of the assay system and, while they have been successful in identifying *trans*-acting factors involved in

Corresponding author: Charles P. Emerson, Jr., Department of Cell and Developmental Biology, University of Pennsylvania School of Medicine, 245 Anatomy-Chemistry Building, Philadelphia, PA 19104-6058. E-mail: emersonc@mail.med.upenn.edu

[1] These authors have contributed equally to this work.

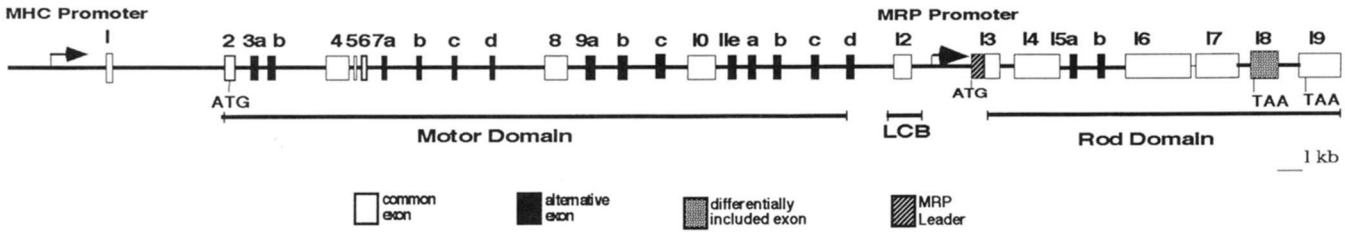

FIGURE 1.—Diagram of the 36B myosin heavy chain (*Mhc*) gene in *D. melanogaster* showing the distribution of the five alternatively spliced exon groups (exons 3, 7, 9, 11 and 15) and the differentially included exon 18. Also shown are the exons coding for the MHC motor, light chain binding, the rod domains and the location of the myosin rod protein promoter element.

muscle-specific alternative splicing regulation (RAMCHATESINGH *et al.* 1995; ROBERTS *et al.* 1996), these studies have been largely limited to the description of factors and or elements that can be distinguished through the comparison of muscle and nonmuscle tissues. Thus, although alternative splicing is widely used to increase protein isoform diversity in muscle (BANDMAN 1992; LOWEY *et al.* 1993; HARRIDGE *et al.* 1996; SCHIAFFINO and REGGIANI 1996), the mechanisms that function to regulate alternative splicing in different muscle types have been largely unaddressed. The focus in this paper is to establish an *in vivo* system in which to study alternative splicing in individual muscle types and to do so in a genetically compliant organism, Drosophila, so that a complete picture of alternative splicing regulation can be generated.

In Drosophila, potentially 480 different isoforms of myosin heavy chain protein (MHC) can be generated through the alternative splicing of the pre-mRNA transcript arising from the single *Mhc* gene (BERNSTEIN *et al.* 1983; GEORGE *et al.* 1989). This diversity is remarkable not only for the array of different MHC protein isoforms available to participate in the structure and function of muscle in this organism, but also for the complexity and precision of the splicing reactions that give rise to the mRNAs that code for the different protein products. The *Mhc* gene consists of 19 exons, of which five (exons 3, 7, 9, 11 and 15; Figure 1) exist as alternatively spliced exon groups that can themselves contain two to five members (GEORGE *et al.* 1989). Further, the *Mhc* gene contains a differentially included exon, exon 18, which provides an early stop codon and polyadenylation signal. Each group of alternative exons encodes the same domain of the MHC protein and, as such, they are included into the message in a mutually exclusive fashion. Previous studies have examined the overall pattern of exon use during development and demonstrated that exon expression can differ markedly in larval and adult animals (reviewed in BERNSTEIN *et al.* 1993). For example, exons e and b of the exon 11 alternative group are not expressed in larvae, but are used in specific adult muscles. The pattern of exon 18 inclusion also differs dramatically in a stage-dependent fashion and is excluded from the larval messages, while being included in most adult transcripts (GEORGE *et al.* 1989). Further, a detailed analysis of alternative exon usage in adult thoracic muscles by HASTINGS and EMERSON (1991) showed that different muscles express very specific sets of alternatives, indicating that exon splice choice is tightly regulated and specific to individual muscles. Thus, alternative splicing of the *Mhc* pre-mRNA transcript is under both developmental and muscle-type specific regulation that results in the expression of the correct MHC isoform in the appropriate muscle.

The mechanisms that ensure the fidelity of the *Mhc* alternative splicing reactions are currently not well understood. In an earlier examination of *Mhc* exon 18, HODGES and BERNSTEIN (1992) showed that the exclusion of this exon from the processed message could be deregulated by the replacement of the native nonconsensus splice donor and nonconsensus splice acceptor with consensus sequences, indicating that these sites play a role in the differential inclusion of this exon. Less is known concerning the mechanisms that regulate the usage of the alternatively spliced exons. While certain combinations of alternatives from different exon groups are not known to occur, suggesting that regulation of splice choice can be coordinated across different alternative exon groups, analysis from *Mhc* mutants indicates that alternative splicing is regulated independently for each exon. The Mhc^{11} mutation is a defective 5′ splice donor in *Mhc* exon 9a that eliminates the ability of this exon to be spliced (KRONERT *et al.* 1991). In the indirect flight muscle (IFM), where exon 9a is normally incorporated into the processed message, a cryptic splice donor in the 9c intron is activated, but no alternative exon 9 is selected. Alternative exons 5′ and 3′ to the mutated exon 9 appear to be processed normally in the Mhc^{11} mutants. Thus, aberrant processing of exon 9 appears not to affect the processing of other exons, indicating that alternative splicing regulation is local to each exon. If alternative processing is regulated by sequences local to the alternative exon domains themselves, then it may be possible to study single exon groups in *Mhc* minigenes for *in vivo* molecular analyses of *cis* regulatory elements that define the splicing regulatory mechanisms.

Mhc exon 11 is of particular interest for study because, with five alternatives, it provides an example of

alternative exon complexity not available in other systems used for the study of alternative splicing. Further, *Mhc* exon 11 includes one alternative, exon 11e, whose expression is restricted to the IFM (HASTINGS and EMERSON 1991). The specificity of the mechanism that identifies and enforces the exon 11e splice choice offers an opportunity to study a precisely regulated muscle-type specific splicing event. In addition, the exclusive use of exon 11e in the IFM indicates that the protein coding capacity of this exon is uniquely important for the function of myosin in this specialized muscle. Therefore, mutations that interfere with the IFM-specific splicing of exon 11e would be expected to result in a muscle incompatible with flight. Thus, genetic screens for flightless mutants should be feasible to identify the transacting factors that participate in the regulation of alternative splicing.

In this article, we demonstrate that transcripts arising from a minigene containing *Mhc* exon 11 are properly and specifically alternatively spliced in muscles of transgenic flies. This transgenic assay system has been used to investigate the *cis*-regulatory elements that participate in muscle-type specific alternative splicing. Our findings show that splicing regulation of *Mhc* exon 11 depends on a multicomponent mechanism that includes the 5′ splice donors in conjunction with intronic sequences that act to specify exon inclusion in particular muscles.

MATERIALS AND METHODS

Drosophila cultures and procedures: All Drosophila cultures were maintained at 25° on standard cornmeal molasses medium (ASHBURNER 1989). A stock of *Drosophila virilis* was obtained from the National Stock Center, Bloomington, Indiana. Transgenic flies were generated by standard methods. The strain $y^1\ w^{67c23(2)}$ was used as the injection strain and as a "wild-type" reference strain in all of the experiments. A minimum of three independently derived transgenic lines was generated for each construct. To examine the expression of the LacZ reporter gene, newly emerged adult flies were embedded in 4% low-melting-point agarose (Sigma) in 1× Buffer A (ASHBURNER 1989), 125-μm sections were taken using a Vibratome 2000 and histochemical staining for β-galactosidase (β-gal) activity was performed as described by ASHBURNER (1989).

Molecular analysis of the *D. virilis Mhc* gene: A genomic library of *D. virilis* in λEMBL3 (constructed by MARY PROUT and JOHN TAMKUN, University of Colorado-Boulder) was screened with a cDNA clone from the *D. melanogaster Mhc* gene, cD301 (GEORGE *et al.* 1989) following the protocols outlined in BRAY and HIRSH (1986). Genomic restriction fragments corresponding to the exon 10 to exon 13 interval were subcloned into the vector pSPORT1 (GIBCO BRL), and deletions were generated for DNA sequencing with the Erase-a-base kit (Promega Corp.). Subclones were sequenced with S-^{35}dATP (NEN) using the Sequenase kit (United States Biochemicals) and the sequence data (GenBank accession number AF019137) were analyzed with the Wisconsin Package Version 9.0 software [Genetics Computer Group (GCG) Madison, WI] and the Staden software package (DEAR and STADEN 1991).

***In situ* hybridization to *D. virilis* thoracic muscles:** Antisense RNA probes from subclones of individual *D. melanogaster* alternative exon 11s (HASTINGS and EMERSON 1991) were labeled with digoxygenin dUTP using the Genius 4 RNA labeling kit (Boehringer Mannheim Biochemicals) and the probe quality was assayed by dot-blot hybridization using the Genius 3 Nucleic Acid Detection Kit (BMB). Thoracic sections of newly eclosed *D. virilis* flies were prepared as previously described (HASTINGS and EMERSON 1991) through the proteinase K step. Using the BMB Genius Kits, the procedures for prehybridization, hybridization, posthybridization wash and probe detection were followed as outlined in the Non-Radioactive *In Situ* Hybridization Manual (BMB).

Plasmid constructions: In general each *Mhc-lacZ* minigene (shown in Figure 11) was the product of a series of subcloning steps using standard molecular techniques (SAMBROOK *et al.* 1989) to create the final fusion genes in the Drosophila vector CasPerβgal (THUMMEL *et al.* 1988). Details of the final structures are given below for each *Mhc-lacZ* minigene; the cloning strategies used to construct these minigenes are available from the authors upon request. Nucleotide position numbers refer to the *Mhc* genomic sequence in GenBank (M61229) (GEORGE *et al.* 1989).

gD1048: gD1048 is the wild-type *Mhc-lacZ* minigene and it served as the parental vector for insertion of *in vitro* mutagenized exon 11 segments. The promoter through exon 2 segment consists of 2.3-kilobase (kb) genomic DNA 5′ of the *Xba*I site that demarcates the beginning of the *Mhc* gene sequence in GenBank file M61229 and ends at the 'C' at position 2569 in exon 2. This C is followed by 10 nucleotides from the pGEM5 (Promega) polylinker, which provide an in-frame bridge sequence between exons 2 and 10. After the 10th nucleotide in the bridge sequence comes the *Mhc* exon 10–13 interval, from nucleotide 10540 to 15972. Thus, the sequence between exons 2 (pos 2569) and 10 (10540) in gD1048 is ATCCCGCGGC and the *Mhc* reading frame is maintained from exon 2 through the bridge sequence into exon 10 and throughout the remainder of the *Mhc* sequences. This clone containing the basic *Mhc* genomic portion for the reporter genes is gD1041. To fuse exon 13 in frame to the *lacZ* structural gene the CasPerβgal vector was digested with *Eco*RI, the *Eco*RI ends were made blunt with Klenow (New England Biolabs) and this vector was ligated to a blunted restriction fragment containing the *Mhc* portion of the minigene. The nucleotide sequence at the *Mhc* exon 13 to CasPerβgal junction is CTCGA (*Mhc* exon 13)–AATTC (*lacZ* gene). All cloning junctions were checked by DNA sequencing at each step in the cloning scheme.

gD1043: The normal *Mhc* reading frame was altered in gD1043 such that only RNAs containing the *Mhc* exon 2–10 spliced to exon 11e then to exon 12 then to exon 13-*lacZ* have a continuous open reading frame. Using the Altered Sites Mutagenesis System (Promega), one nucleotide, an A, was inserted into a subclone containing exon 11e between nucleotides number 11426 and 11427, and, in an exon 12 subclone, one nucleotide, base number 14130, was deleted from exon 12. These altered exons were cloned into gD1041 (see above). This mutated *Mhc* segment was digested with *Xho*I, the *Xho*I ends were partially filled in with nucleotides T and C and the fragment was ligated to the CasPerRβgal vector to which *Bam*HI-digested ends had been partially filled in with nucleotides A and G. The resulting sequence at the *Mhc* exon 13-*lacZ* junction is CTCGATCCC.

gD1090, gD1177, gD1105, gD1120, gD1168 and gD1060: A two-step PCR method (ZARET *et al.* 1990) was used to generate mutations in various segments of the wild-type minigene gD1048. Briefly, the segment to be changed was subcloned into a Bluescript plasmid vector (Stratagene), subjected to

the two-step PCR *in vitro* mutagenesis and the mutant products were cloned back into the parental gD1048 construct. A description of each mutated minigene is given below and the sequence of the sense strand mutagenic oligonucleotide is listed.

gD1090: The 5′ splice site of exon 11e was changed to a consensus 5′ splice site. The sense strand mutagenic primer was oD098 (5′ CACACCAAGGTAAGTACGCAAAAAAATG 3′).

gD1177: The 5′ splice site of exon 11b was changed to a consensus 5′ splice site. The sense strand mutagenic primer was oD116 (5′ GGACACACCAAGGTAAGTAAACGTTTC 3′).

gD1105: The 5′ splice site of exon 11e was converted into the 5′ splice site of exon 11b using mutagenic primer oD100 (5′ CACACCAAGGCATATACGCAAAAAAAT 3′). The 5′ splice site of exon 11b was converted into the 5′ splice of exon 11e using mutagenic primer oD096 (5′ GGA CACACC AAGGCACGCAAACGTTTC 3′).

gD1120: The ICR (no. 13870–13949)was removed using the two-step PCR and the mutagenic primers oD110(5′ TTG GTACAAACAAGAACAGAACAGAAAGG) and oD111 (5′ CTTGTTTGTACCAAAATGAAAAAATGAATGAACC).

gD1168: The *Afl*III restriction fragment corresponding to positions 11291–11539 from the *Mhc* gene sequence was deleted from the exon 10 to 13 domain in gD1048. This deletion removed exon 11e, 27 base pairs (bp) 5′ and 103 bp 3′ of exon 11e.

gD1060: The *Afl*II restriction fragment corresponding to positions 11622 and 13927 from the *Mhc* gene sequence was deleted from the exon 10–13 domain in gD1048. This deletion removed all of the alternative exon 11s 3′ to exon 11e.

gD1222: The conversion of exon 11e into exon 11b and the conversion of exon 11b into exon 11e in the gD1222 minigene was accomplished using modifications to published protocols for PCR mutagenesis using megaprimers (CLACKSON *et al.* 1989) and our standard two-step PCR mutagenesis protocol described above. Hybrid oligonucleotides containing both exon 11e and exon 11b sequences were used in two parallel PCR reactions to generate megaprimers containing coding sequence of one alternative exon 11 flanked by the intron sequence of the other alternative exon 11. Single-stranded megaprimers were prepared for use in the following round of amplification by phosphorylating appropriate oligonucleotides prior to the first round of amplification, then digesting these products with the Strandase enzyme (Novagene), which selectively degrades the phosphorylated strand. These megaprimers were purified and used as mutagenic oligonucleotides in the two-step PCR reaction. For the construction of a subclone containing exon 11e flanked by the intron of exon 11e, the following hybrid oligonucleotides were synthesized.

oD156: Sense strand oligonucleotide in which bases 1–24 correspond to exon 11e intron and 3′ splice site acceptor and bases 25–49 correspond to the exon 11b coding sequence. (5′–3′ GTGAATTGTAATACTCAACTATAGCTATCAAAT CCTGAACCCAGCTGG).

oD155: Antisense strand oligonucleotide in which bases 1–23 correspond to exon 11e intron sequence and 5′ splice site and bases 24–46 correspond to exon 11b coding sequence. (5′–3′ GGAACATTTTTTTGCGTGCGTGCCTTGGTGTGT CCAATGCGATAC).

For the construction of a subclone containing exon 11e flanked by the intron of exon 11b the following hybrid oligonucleotides were synthesized.

oD135: Sense strand oligonucleotide in which bases 1–19 correspond to exon 11b intron and 3′ splice site acceptor and bases 20–37 correspond to exon 11e coding sequences. (5′–3′ GGAACATTTTTTTGCGTGCGTGCCTTGGTGTGTCCA ATGCGATAC).

oD133: Antisense strand oligonucleotide in which bases 1–20 correspond to exon 11b intron and 5′ splice site and bases 21–39 correspond to exon 11e coding sequence. (5′–3′ GGA TTGAAACGTTTATATGCCTTGGTGTGACCAATGCGG).

Westerns and immunoblot analysis: Newly eclosed flies with unexpanded wings were collected and processed as described in STANDIFORD *et al.* (1997) for protein gels and immunoblot analysis. Three flies were prepared together for each genotype and the protein equivalent of a single fly was loaded onto each lane. Anti β-gal antibodies (5′–3′) were used at a 1:5000 dilution. Blots were detected with 0.5× Super-CL chemiluminescent reagent (Pierce), recorded onto X-omat film (Kodak) and the bands were scanned and quantitated using IPLab Gel H (Signal Analytics) image analysis software.

Reverse-transcriptase-dependent PCR (RT-PCR): Newly eclosed flies were collected and then soaked for a few minutes in petri plates containing 100% ethanol at room temperature. After this treatment the IFM, transtergal depressor of the trochanter (TDT) and other muscles become stiff and can be easily dissected away from each other. Separated IFM and TDT muscles were placed into 1.5-ml microfuge tubes containing 100% ethanol on ice until 10–15 flies per genotype were dissected. The tubes containing the muscles were spun at 2000 rpm, 4° in an Eppendorf microfuge, and the ethanol was removed. The muscles were resuspended in 100 ml of H_2O, and 0.5 ml of Trizol reagent (GIBCO-BRL) was immediately added along with 20 μg of yeast transfer RNA (Sigma) as carrier. Total RNA was isolated with the Trizol reagent using the manufacturer's instructions. Following the isopropanol precipitation the RNA pellet was resuspended in 100 μl reverse transcriptase buffer with 100 μM dNTPs, 40 units RNAsin (Promega), 200 units SuperScript II Reverse Transcriptase (GIBCO-BRL) and an antisense oligonucleotide primer to *Mhc* exon 12. Reverse transcription reactions were at 42° for 1 hr. For the primary PCR reaction, 5 μl of the reverse transcriptase reaction was used in a 100-μl reaction under the following conditions: one cycle at 95° for 5 min, 30 cycles of 95° for 30 sec, 55° for 1 min, 72° for 1 min and one cycle of 74° for 4 min. Five microliters of the primary PCR reaction were used for the secondary PCR reaction and the same PCR cycle profile was followed. The sequences for all of the oligonucleotide primers used in these experiments are presented in Table 1.

RESULTS

***Mhc* alternative splice regulation is exon specific:** To define sequences that control the muscle-type specific alternative splicing of exon 11 *in vivo*, a *Mhc* exon 11 minigene reporter construct gD1048 was developed for analysis in transgenic animals (Figure 2). This construct contains the entire *Mhc* exon 11 domain and is flanked 5′ by exon 10 and 3′ by exon 12, intron 12 and the first 30 nucleotides of exon 13, which are directly followed by an in-frame *lacZ* gene. Sequences 5′ of exon 10 include the *Mhc* promoter domain, *Mhc* exon 1 and *Mhc* exons 2 and 10, which are fused together in-frame. An important feature of this construct is the natural split codon that exists between exon 10 and each of the exon 11s. Thus, production of a functional 175 kD MHC::β-gal protein from transcripts arising from the *Mhc* promoter in gD1048 requires mRNA processing and inclusion of a single alternative exon 11: skipping of exon 11 or the inclusion of multiple alternatives would gener-

TABLE 1

Oligonucleotides used in the RT-PCR assay

Oligonucleotide no.	Sequence	Position
364	CGTGTGTCGAAGCTGTTCGG	Exon 8 outer -sense
363	GGTGTACTGGATATTGCTC	Exon 8 inner -sense
361	CGAAATTTAAGCAAGATGC	Exon 2 outer -sense
590	GACTCGAAGAAGTCTTGCTG	Exon 2 inner -sense
365	GAATGCTGCCGGAAAGTGCTTG	Exon 11e -sense
301	AAAGTGTGGATCCATCATTTTG	Exon 11b -sense
362	CCAAATGATGGATCCACAC	Exon 11b -antisense
380	TCAATCAGAACCTTGGAGCCT	Exon 11a -antisense
379	TCGATCAGAACCTTCGTGCAT	Exon 11c -antisense
378	TTAATGATAATTTCAGTGGCC	Exon 11d -antisense
360	CAAGCACTTTCCGGCAGCA	Exon 11e -antisense
367	CTGACCCAGGACACCGGCGCG	Exon 12 inner -antisense
366	GGACATGATCTTGCCCAGACGC	Exon 12 outer -antisense
302	CACGACGTTGTAAAACGACGGCC	LacZ inner -antisense
303	GCTGCAAGGCGATTAAGTTGGG	LacZ outer -antisense

ate an out-of-frame message, resulting in the absence of β-gal activity. The gD1048 construct also contains the promoter for the myosin rod protein gene (MIEDEMA *et al.* 1995; STANDIFORD *et al.* 1997), which is active in a subset of muscles, but importantly it is not expressed in the IFM or the TDT muscles. In this study, we have used the presence or absence of β-gal in the IFM or TDT as an indicator of correct splicing of transcripts arising from the gD1048 transgene and its derivatives.

When flies transformed with the gD1048 transgene were tested using β-gal staining, the IFM and TDT stained positively, indicating that the *Mhc* minigene is expressed and processed into an mRNA that contains only a single exon 11 (Figure 3A). The fidelity of exon usage from the minigene was tested by the gD1043 minigene (Figure 3B). This construct has a single base addition in exon 11e and a single base deletion in exon 12, so that only processed messages containing exon 11e are capable of β-gal expression. In flies transformed with gD1043, the IFM was positive for β-gal, while the TDT does not stain, indicating that exon 11e utilization in the minigene is restricted to the IFM. Thus, the gD1048 minigene has sufficient *cis*-information to direct the proper splicing of *Mhc* exon 11.

To detect changes in the pattern of alternative exon use in specific muscles that might result during the mutational analysis of the gD1048 construct, a RT-PCR assay was developed (Figure 4). In this assay, IFM or TDT muscles are dissected and used to prepare total RNA. Messenger RNAs arising from both the endogenous and minigene *Mhc* promoters are reverse transcribed using antisense primers specific for *Mhc* exon 12 (Figure 4A). Messages arising from either the endogenous *Mhc* or the *Mhc* minigene are then specifically amplified and distinguished with the use of PCR primers for exon 8 or exon 2, respectively, in combination with exon 12 primers placed 5′ (inner) to the exon 12 primers used in the RT reaction. Individual exon 11s are detectable in these amplified products through the use of primers specific for each exon 11 when used in combination with either exon 2 or exon 8 primers that are 3′ (inner) to those primers used in the first round of PCR. The products from the second round of PCR representing the correctly processed endogenous messages are 870 bp in size, while the correctly processed minigene transcripts are 768 bp. Thus, the IFM or TDT can be specifically assayed for their alternative splice utilization of the exon 11 alternatives arising from the

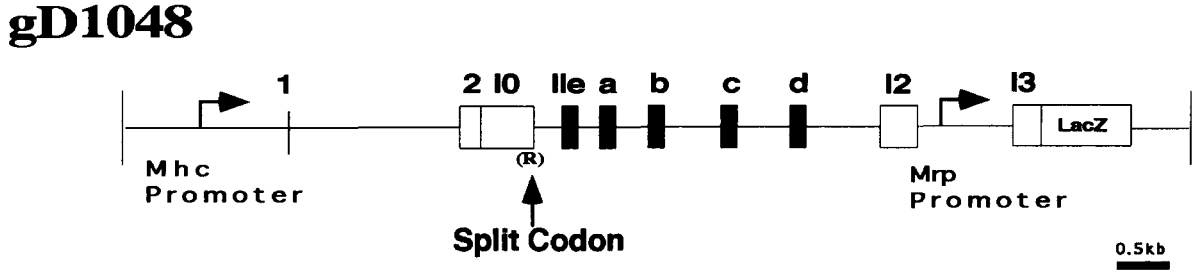

FIGURE 2.—Diagram of the gD1048 *Mhc* exon 11 minigene designed to study the *cis*-regulation of alternative splicing in transgenic flies. The gD1048 construct contains the *Mhc* promoter, exon 2 fused in frame to exon 10, all exon 11 alternatives followed by exon 12, intron 12 and exon 13 fused in frame with the *lacZ* reporter gene. Exon 10 contains a split codon such that only processed messages that contain one exon 11 will produce a functional *lacZ* product.

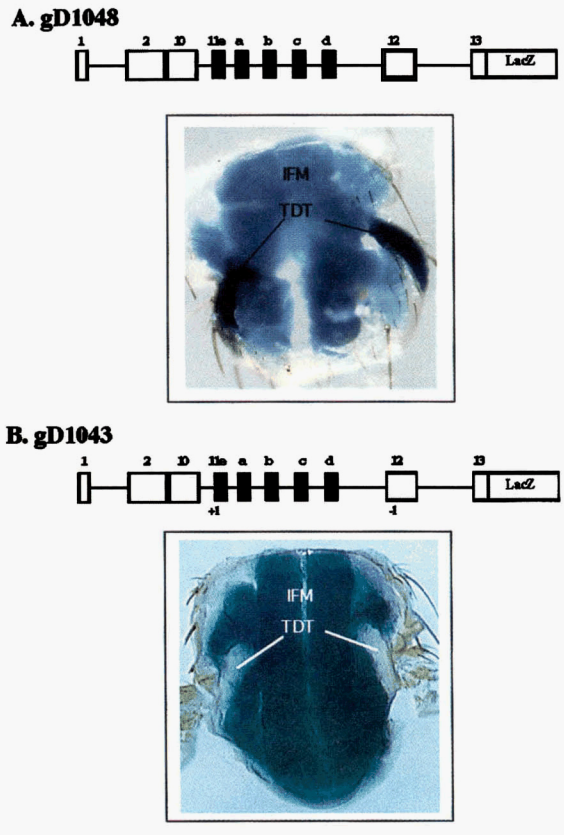

FIGURE 3.—(A) Flies transformed with the gD1048 construct stain positively for β-gal in all muscles. Shown here are the IFM and TDT in thoracic sections of flies containing the gD1048 transgene, which are both positive for β-gal. (B) The gD1043 construct contains an insertion in exon 11e and a deletion in exon 12 and expresses β-gal in the IFM but not the TDT.

minigene, and this can be concurrently compared to the splice utilization of alternative exon 11s arising from the endogenous *Mhc* gene. Further, this assay can be used to examine the overall expression of exon 11 alternatives in larval muscles, where exons 11a, 11c and 11d are normally the only exon 11 alternatives expressed.

When the RT-PCR assay was used to determine the muscle-type specific use of gD1048 alternative exon 11s in the IFM, only exon 11e was detected, which matches the pattern for the endogenous gene. In the TDT, only exon 11b was detected, which is the normally utilized alternative in this muscle. These results establish that correct muscle-type specific splicing of exon 11 alternatives is properly directed by the gD1048 minigene. To further test the fidelity of the gD1048 minigene, RNAs collected from larvae were assayed for exon 11 expression with RT-PCR. This analysis shows the overall pattern of exon use from the minigene to be identical to that of the endogenous gene (Figure 4C) Specifically, exons 11a, 11c and 11d are detected in larval RNAs arising from both the endogenous *Mhc* and the gD1048 transgene, while exons 11e and 11b are not detected, which is consistent with the known expression patterns for the different exon 11s in larvae (GEORGE *et al.* 1989). These data show, therefore, that the sequence information required to specify and direct the selection and use of alternative exons in individual muscles is entirely contained within the 6-kb exon 11 domain and that this process does not require the presence of any other *Mhc* alternative exon group.

Sequence conservation in *Mhc* exon 11: To identify conserved sequence elements that might be important for exon 11 splicing regulation, the *Mhc* exon 11 domain from the distantly related (40–60 myr divergent) drosophilid species (BEVERLEY and WILSON 1984) *D. virilis* was cloned and sequenced (Figure 5). This analysis revealed that, like *D. melanogaster*, the *D. virilis* exon 11 domain also contains five alternative exons. Sequence alignments (GCG; Pileup) of exon 11 alternatives from *D. melanogaster* and *D. virilis* pairs all equivalently positioned alternative exon 11s (*i.e.*, the *D. melanogaster* exon 11e was most similar to the most 5′ alternative exon 11 from *D. virilis*), indicating that order of alternative exons is a conserved feature. Interestingly, alternative exon 11e is the most divergent between the two species, both at the nucleotide (77% similar; Figure 5) and protein level (not shown), which perhaps reflects a functional difference in the IFM isoform of myosin in the considerably larger *D. virilis*. The exon 11 alternatives each contain a unique, nonconsensus 5′ splice donor and these are conserved in *D. virilis*. The 3′ splice acceptors from *D. virilis* and *D. melanogaster* exon 11e and 11c are nonconsensus (TAG), while the remainder are consensus (CAG). All exon 11s from *D. virilis* and *D. melanogaster* have relatively poor polypyrimidine tracts (MOUNT *et al.* 1992), containing ~50% purine nucleotides, although the exact sequence of these domains is not well conserved (not shown).

Within the introns of the *Mhc* exon 11 domain, which are similar in overall composition to other Drosophila introns (CSANK *et al.* 1990), there are several short segments of conservation between *D. virilis* and *D. melanogaster* and a single large, 78 nucleotide (nt) intronic conserved region (ICR) residing between exon 11d and exon 12 (*D. melanogaster* sequence no. 13861–13949; Figure 5C). The ICR is generally A/T rich, (60%), contains three repeated ATCC sequence elements and is not reiterated elsewhere in the *D. melanogaster Mhc* gene. Intronic elements that are AT rich have been shown to influence the selection of 5′ splice sites in Drosophila (MCCULLOUGH and SCHULER 1993), suggesting a role in splice regulation for the ICR. Perhaps also significant is the fact that, while the ICR is not overall purine rich (38 G/A), there are several clusters of purine rich sequence in this domain (Figure 5C). Splicing enhancers in the alternatively spliced troponin T (RAMCHATESINGH *et al.* 1995), *doublesex* (LYNCH and MANIATIS 1996) and bovine growth hormone (SUN *et al.* 1993) transcripts are composed of purine-rich sequences that have been shown to interact with various

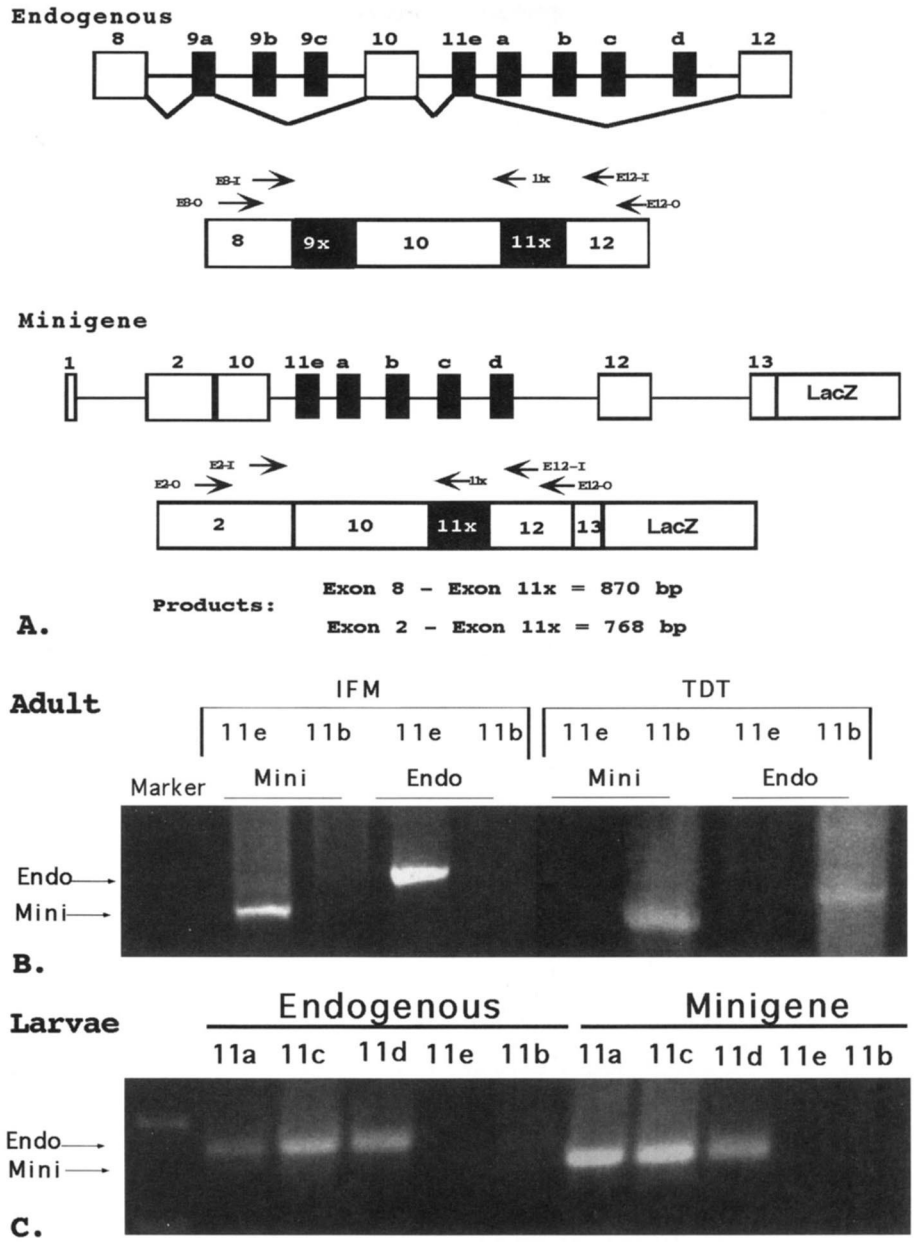

FIGURE 4.—(A) Diagram of RT-PCR assay designed to determine the pattern of exon usage in processed messages arising from either the endogenous *Mhc* gene or the *Mhc* exon 11 minigene. Total RNA collected from the IFM or TDT is reverse transcribed with a primer common to both endogenous and minigene transcripts. A first round of PCR uses a common exon 12 primer and a primer that is specific to exon 8 or exon 2 to differentiate endogenous or minigene transcripts, respectively. A second round of PCR uses primers specific for the exon 11 of interest in conjunction with primers specific to either exon 8 or exon 2 to determine the pattern of alternative exon 11 use from the endogenous or minigene, respectively. Thus, this assay determines the alternative exon 11 use from the *Mhc* minigene in specific muscles and this can be compared directly to the known pattern of exon use from the endogenous gene. (B) When applied against RNAs collected from the IFM of flies transformed with the gD1048 construct, a 768-bp minigene product is seen when the exon 11e primer and not the exon 11b primer is used in conjunction with exon 2 primers. An 870-bp band is detected in the IFM when an exon 8 primer is used with an exon 11e primer and not the exon 11b primer. The minigene product is only detected in the TDT with the exon 11b primer in conjunction with either exon 2 or exon 8 primers. These data show the muscle-specific pattern of exon use is identical for both the endogenous gene and the minigene exon 11 alternatives. (C) Total RNA collected from gD1048 larvae was assayed using RT-PCR and the overall pattern of alternative exon 11 usage was assayed for both the endogenous and minigene transcripts and found to be identical. Marker bands are 1.0 or 0.5 kb.

SR proteins directly and a similar role might exist for the *Mhc* exon 11 ICR.

The pattern of exon usage also is conserved between *D. virilis* and *D. melanogaster*. *In situ* hybridization experiments using exon 11 probes from *D. melanogaster* against thoracic sections from *D. virilis* show that exon 11e is expressed exclusively in the IFM in *D. virilis* thoracic muscles (Figure 5D) and exon 11b is expressed in the

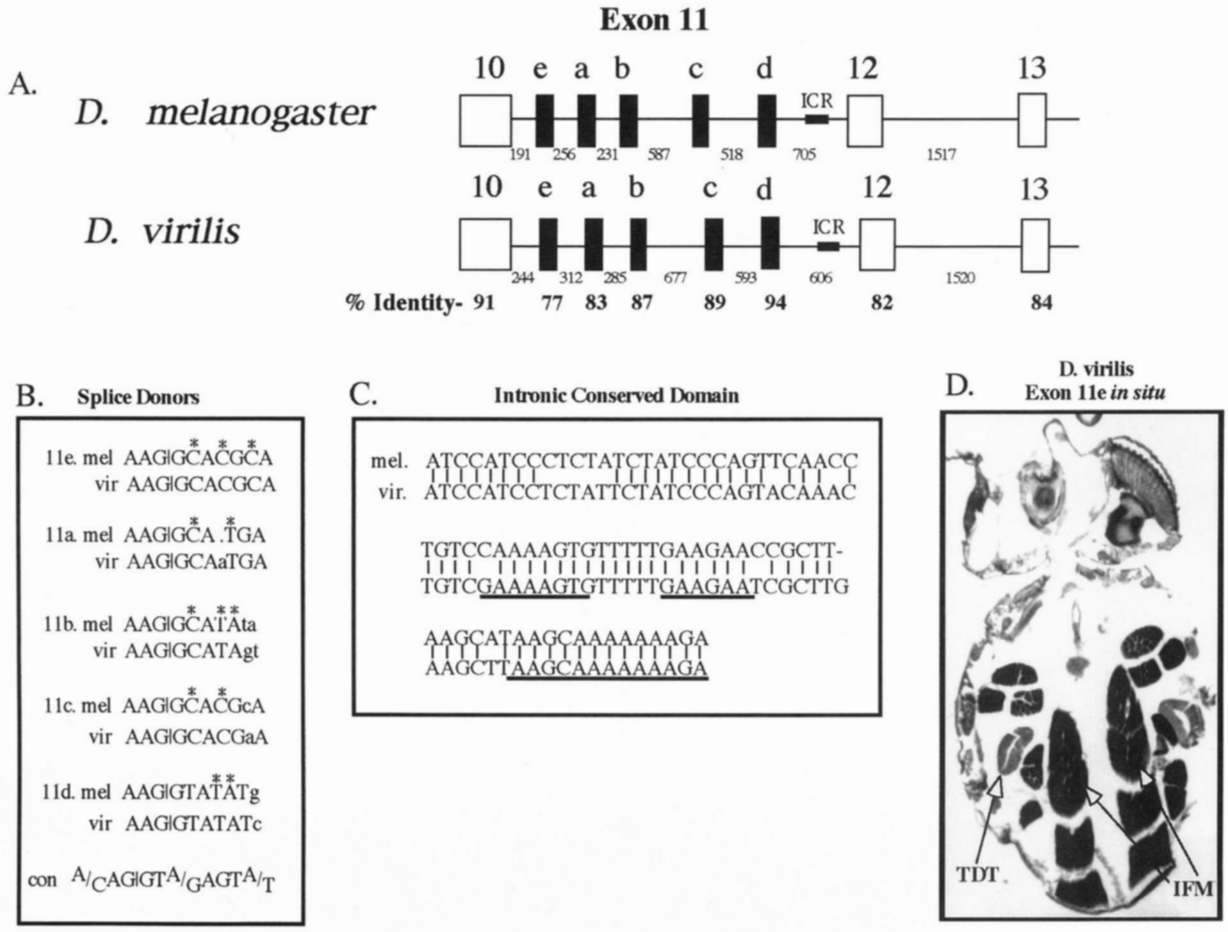

FIGURE 5.—Summary of the sequence comparison between *D. melanogaster* and *D. virilis*. (A) Overall, the structure is the same and exon 11 from *D. virilis* contains five alternatives that are, based on similarity, in the same order as the *D. melanogaster* alternatives. Percent identity between corresponding alternatives is shown below each exon. Exon size is conserved as is the distance between each alternative. (B) Exon 11 alternative exon splice donors are conserved and all are nonconsensus (*, nonconsensus nucleotides). (C) A large ICR between exon 11d and 12 is present, is 89% identical between *D. virilis* and *D. melanogaster* and represents the largest site of intronic conservation. Several domains of purine-rich sequence are underlined. (D) The regulation of the alternative exons in *D. virilis* was tested through *in situ* analysis using *D. virilis* thoracic sections and *D. melanogaster* exon 11 probes. Shown is the result of exon 11e probe, which specifically hybridizes to the IFM of *D. virilis*.

TDT (not shown), which is the pattern of expression for these two exons in *D. melanogaster*. The conservation of alternative exon use in *D. virilis* indicates that the splicing mechanism is also conserved, thereby suggesting that the conserved elements in this domain have a role in splicing regulation. Our strategy to identify the important regulatory elements involved in alternative splicing is, therefore, focused on the conserved 5′ splice donors, the exons themselves, the position and spacing of the exons and the conserved intronic sequences.

Nonconsensus 5′ splice donors are essential for muscle-specific splicing regulation: Each exon 11 contains a unique 5′ splice-donor that is nonconsensus and whose sequence is conserved. This suggests a simple model that alternative exons are specified through a mechanism that identifies a particular donor in the appropriate muscle type, thus effecting inclusion of only that exon. To test this, the gD1105 clone was constructed in which the 5′ splice donors from exon 11e and 11b were exactly switched. Transformed flies were found to express β-gal in all muscles of the thorax (not shown), indicating that this donor swap does not disrupt overall splicing. When the IFM and TDT mRNAs were assayed for exon 11 expression using RT-PCR (Figure 6A), the pattern of exon usage was found to be identical to that of the endogenous gene: exon 11e is still expressed in the IFM and exon 11b is expressed in the TDT. No effect on the pattern of exon use was seen in larvae containing the gD1105 construct (not shown). These data indicate that muscle specificity of exon choice is not determined by the nonconsensus 5′ splice donors.

The nonconsensus nature and conservation of exon 11 donors suggest a more general role in splicing regulation. The requirement for nonconsensus donor sequences was examined through the replacement of the donor of exon 11e with a consensus donor in the gD1090 construct (Figure 6B). Flies transformed with

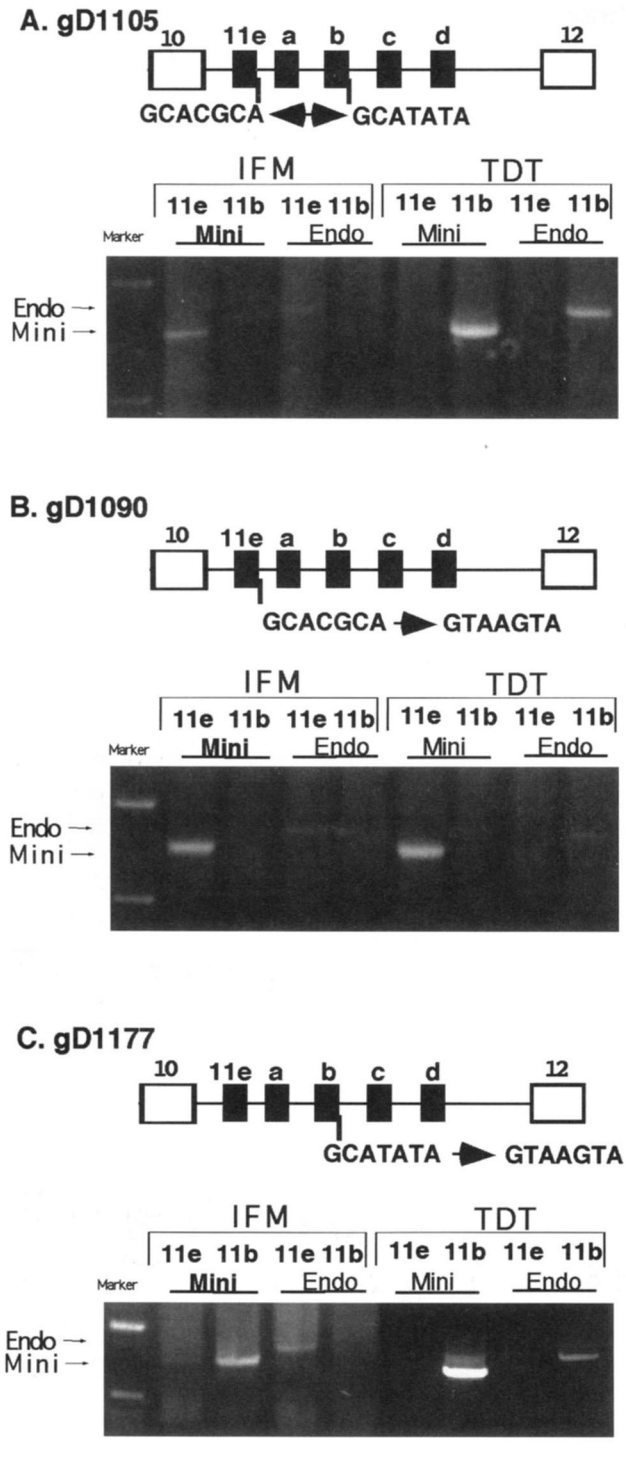

FIGURE 6.—Constructs designed to test the function of the exon 11 nonconsensus 5′ splice donors in splice regulation. (A) The 5′ splice donors from exons 11e and 11b were exactly swapped in the gD1105 construct. The usage of minigene exons 11e and 11b in the IFM and TDT were determined with RT-PCR and found to be identical to the wild-type usage. (B) The nonconsensus donor of minigene exon 11e was replaced with a consensus donor in the gD1090 construct. RT-PCR analysis of RNA from isolated muscles revealed that the minigene exon 11e is still used in the IFM, but is now also processed into the messages arising from the minigene in the TDT. No exon 11b minigene products were detected in either the IFM or TDT. (C) The nonconsensus 5′ splice donor of this construct express β-gal in all muscles, showing that modification of the exon 11e 5′ splice donor does not promote the inclusion of multiple alternative exons or the skipping of all alternative exons during processing. However, the substitution of a consensus donor in exon 11e does result in the dominant expression of this alternative, which, as demonstrated by RT-PCR, is now aberrantly expressed in the TDT in addition to the IFM. Further, no minigene transcripts containing exon 11b were detectable in the TDT, showing that the activation of exon 11e completely represses the splicing of exon 11b in the this muscle. Finally, the gD1090 exon 11e was also exclusively used in larval minigene transcripts (not shown), demonstrating that a consensus donor, when placed in an alternative exon, can convert that alternative to a constitutively utilized exon that completely represses the use of the normal alternative.

The ability of a consensus donor to confer splice-choice dominance to exons other than that of exon 11e was tested with the gD1177 transgene (Figure 6C), in which the donor of exon 11b is replaced with a consensus donor. Transgenic flies containing the gD1177 construct were found to stain for β-gal in all thoracic muscles, again establishing that the inclusion of a consensus donor into exon 11b does not induce the addition of multiple exon 11s into the processed transcript. When RT-PCR was used to assay the pattern of exon usage in the IFM and TDT of gD1177-transformed flies, it was found that exon 11b was now expressed in both the IFM and TDT. However, in the IFM, the minigene exon 11e is still detectable by RT-PCR, indicating that the substitution of a consensus donor in exon 11b does not completely repress the use of exon 11e from the gD1177 construct in this muscle. This is in contrast to the behavior of the modified minigene exon 11e in the gD1090 construct. The inclusion of the consensus donor into exon 11b however does result in this exon being used exclusively in transcripts arising from the gD1177 transgene in larvae (not shown). Together, the results obtained with the gD1090 and gD1177 minigenes show that the replacement of nonconsensus exon 11 donors with a consensus donor converts a regulated exon to a splice-dominant exon that represses the normal exon choice but does not disable the mechanism that maintains the mutual exclusivity of alternative exon splicing.

Role of conserved intron and exon sequences: A domain of striking sequence conservation (ICR) occurs in the intron between exon 11d and exon 12 (Figure 5), where it might function to regulate splicing through

exon 11b was converted to consensus in the gD1177 minigene construct. RT-PCR analysis of exon use revealed that, in the IFM, products arising from the minigene now contain exon 11b. This appears to occur in conjunction with the expression of some exon 11e, as this exon is also detected in the minigene products. Only exon 11b was detected in minigene products expressed in the TDT.

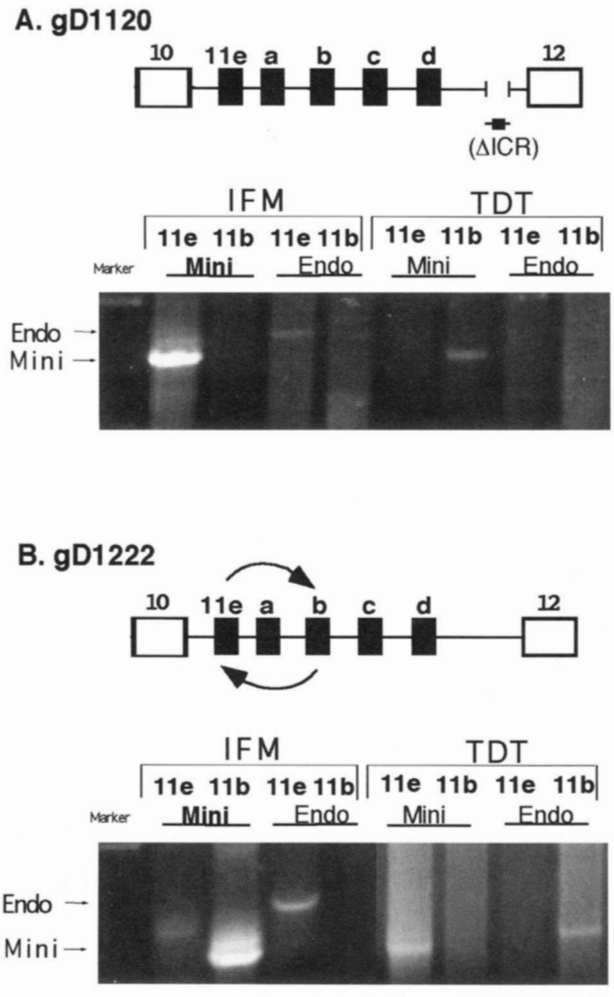

FIGURE 7.—Constructs designed to test the function of the ICR and exonic sequences in alternative splicing regulation. (A) The deletion of the ICR was analyzed in the gD1120 construct by RT-PCR and found to have no effect on the expression pattern of minigene exons 11e and 11b in either the IFM or TDT. (B) Minigene exons 11e and 11b were positionally exchanged in the gD1222 construct and this was tested by RT-PCR analysis for its effects on alternative exon usage. The results of this analysis revealed that the exchanged exons were also exchanged in their utilization with 11b now expressed in the IFM while exon 11e from the minigene is expressed in the TDT.

mechanisms such as providing alternative or regulated branch point sequences or by modifying the access of splicing factors to the 3′ splice acceptor of exon 12. To examine these possibilities, the gD1120 transgene was constructed (Figure 7A), in which the ICR domain was deleted. Flies transformed with gD1120 express β-gal throughout the thorax, indicating that this deletion does not induce exon skipping or the aberrant splicing of multiple alternative exons. Further, the removal of the ICR does not alter the splice usage of exons 11e or 11b nor is the pattern of exon use changed in larvae, when exon usage was assayed with RT-PCR (not shown), indicating that the ICR is not essential to regulate the specificity of alternative exon usage.

Although the alternatives of exon 11 each code for the same protein domain of MHC, there is significant divergence both at the protein and nucleotide level within these exons. While these differences are likely to be important for the function of the myosin protein itself, we considered the possibility that these exon sequences serve to participate in splicing regulation, perhaps as splicing enhancers, to directly tie alternative splicing regulation to the alternative exons themselves. The gD1222 transgene, in which exons 11e and 11b are exactly switched in their positions, was developed to test this hypothesis (Figure 7B). If the exon sequences alone direct their own usage, then exon position with respect to other exons in the domain should not influence exon choice. Flies transformed with gD1222 are positive for β-gal staining throughout the thorax, indicating that splicing of multiple alternatives or exon skipping does not occur for this construct. However, in transcripts arising from gD1222, 11e is now expressed in the TDT and exon 11b is expressed in the IFM when spliced products were assayed for exon usage by RT-PCR, while the larval pattern of exon use was undisturbed. Thus, substitution of exon 11b in the position of exon 11e results in exon 11b utilization in the IFM as if it were exon 11e. Thus, these data demonstrate that the *cis*-information specifying exon choice is not contained within the exons themselves and must either reside within the intron or is contained within the order of the exons themselves.

IFM splice specificity is local to exon 11e: The linear order of the *Mhc* exon 11 alternatives is conserved and might function to regulate splicing if the mechanism depends on the muscle-specific balance of accessory splicing factors, such as SR proteins and hnRNPs, which have been shown to influence splice choice (MAYEDA and KRAINER 1992; CACERES *et al.* 1994; reviewed in CHABOT 1996). The gD1168 transgene, in which exon 11e is deleted, was developed to test the possibility that the order of exons in the exon 11 domain participates in the definition of exon usage (Figure 8). In this construct, exon 11a is moved into the position proximal to exon 10. If exon order dictates exon usage, then exon 11a should be activated in the IFM. Flies transformed with the gD1168 transgene, however, were found to lack β-gal staining in the IFM while other muscles of the thorax stain normally, indicating that no exon 11 is used correctly in the IFM. Further, in gD1168, RT-PCR analysis showed that exon 11b was still expressed normally in the TDT, as were other alternative exons in larval muscles. These data show that shifting of exon 11a into the position closest to exon 10 does not alter the pattern of exon use and that the IFM is unable to select another alternative in the absence of exon 11e. This suggests that the sequences specifically required to permit the choice of exon 11e in the IFM are absent in the gD1168 construct or they are incompatible with other exon 11s.

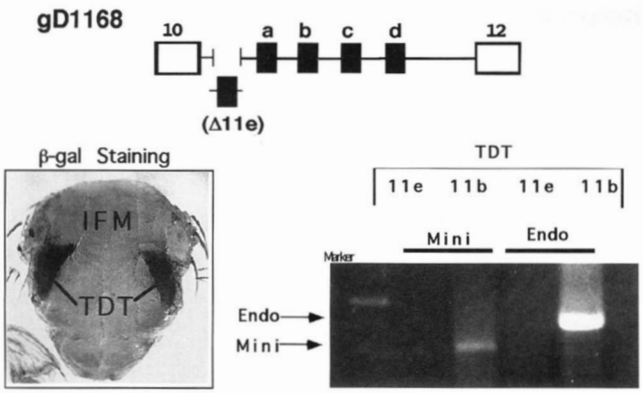

FIGURE 8.—The gD1168 minigene construct was developed to test the function of exon position in alternative splicing regulation. (A) Exon 11e was deleted from the *Mhc* exon 11 minigene, which moved exon 11a to the proximal position with respect to exon 10. The effect of this mutation was tested by β-gal staining of thoracic sections which showed that there is no staining in the IFM, while functional β-gal was present in the TDT. (B) RT-PCR analysis of the TDT showed that exon 11b was still processed normally into messages arising from the gD1222 minigene. Analysis of the total RNA collected from gD1222 larvae showed that exon usage was not disrupted (not shown).

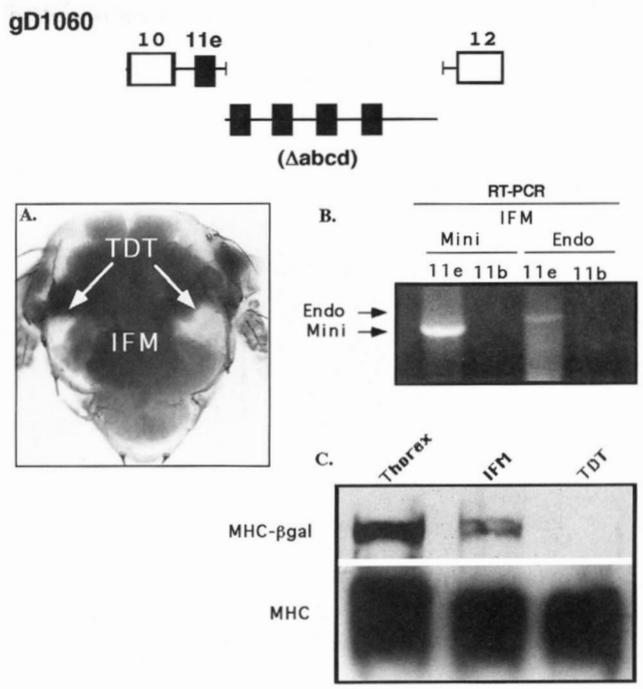

FIGURE 9.—The gD1060 minigene construct isolates sequences required for the IFM-specific inclusion of exon 11e into the processed message. A minimal construct containing only exon 11e in addition the entire intron 10 and 75 bp of intron 11e was analyzed transgenically. (A) In thoracic sections of flies that contain the gD1060 construct, the IFM was seen to stain positively for β-gal but not the TDT. (B) RT-PCR analysis of the IFM showed that exon 11e was normally processed into the transcripts arising from the gD1060 minigene. (C) Immunoblot analysis of total thoracic proteins, IFM proteins or TDT proteins shows that the MHC-β-gal fusion protein expressed from the gD1060 minigene occurs only in the IFM, and no product is detectable in the TDT. Myosin heavy chain (MHC) protein was immunodetected in each sample as a loading control.

The finding that neither the position or sequence of the exon itself specify usage indicates that the sequences flanking the exons contain regulatory sequences that are important for splicing regulation. To directly test this possibility, the gD1060 transgene was developed (Figure 9), in which all alternative exon 11s except exon 11e are deleted from the minigene. The entire 180-bp intron between exon 10 and 11e and 180 bp 3′ to exon 11e remain intact and these 11e sequences are ligated to a position ~190 bp 5′ of exon 12 removing 60 bp of the ICR domain. Flies transformed with gD1060 express β-gal in the IFM, but not in TDT, indicating that this construct contains *cis*-information sufficient to specifically direct splicing of exon 11e in the IFM. This was further tested by RT-PCR analysis of larval muscles, which showed that no minigene exon was expressed here (not shown), and by immunoblotting, which demonstrated that the MHC-β-gal product is present only in the IFM. These data indicate that gD1060 contains the elements required for the IFM-specific use of exon 11e. Further, given the interchangeability of splice donors (gD1105; Figure 6), the neutrality of the exons themselves in alternative splice choice specification (gD1222; Figure 7) and the failure of sequences other than those removed in gD1060 to direct splicing in the IFM (gD1168; Figure 8), these results indicate that the specificity elements are entirely contained within the 360 nt of sequence flanking exon 11e.

Splicing efficiency and exon usage: As seen above, conversion of the exon 11 5′ splice donors to consensus (gD1090, gD1177; Figure 6) results in the dominant activation of the exon that contains consensus donor. While the consensus donor exon is the predominant exon included in the processed transcript, it is possible that the splice reaction rate is affected by this donor change. Such changes might result from a competition between the normally defined exon and the consensus donor exon for splicing factors, or the consensus donor exon might need to overcome some inhibitory effect that normally represses the use of that exon in inappropriate muscles. Alternatively, the dominant use of the consensus donor exon might result from its ability to assemble more rapidly the splicing apparatus, thus processing of this exon might be more efficient than the processing of the wild-type exons. Further, although spliced products arising from the gD1090 and gD1177 transgenes were not seen to contain multiple alternative exons by RT-PCR, these reaction products would be out-of-frame and perhaps subject to turnover rates high enough to prevent their detection. To determine the effects of the different mutations on splicing efficiency and or RNA turnover, we took advantage of the fact that the all the *Mhc* minigenes used in this study contain

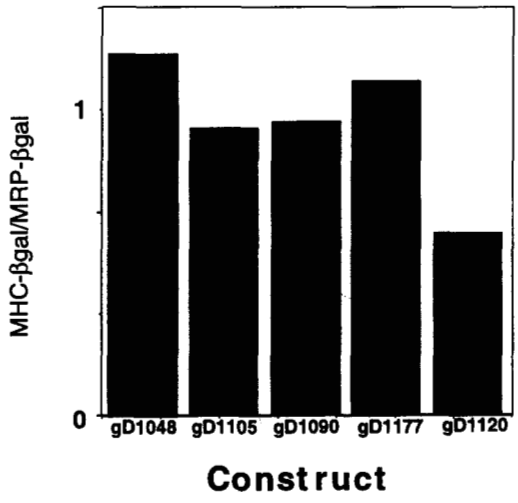

FIGURE 10.—Analysis of splicing efficiency for the different splicing mutants examined in this study. Splicing efficiency was determined by comparing the levels of the protein products arising from the *Mhc* promoter (MHC-β-gal) to the proteins products arising from the *Mrp* promoter (MRP-β-gal), both of which are contained in all minigene constructs. MHC-β-gal product is dependent on alternative splicing for its expression, while MRP-β-gal is expressed directly. Changes in the ratio between these proteins that result from modifications to the minigenes are likely to indicate changes in alternative splicing. This analysis shows that the swapping of exon 11e and 11b donors on the gD1105 construct does not alter the ratio of MHC-β-gal to MRP-β-gal nor does the replacement of their natural donors with consensus donor sequences in the gD1090 or gD1177 constructs. In the gD1120 construct, where the ICR is deleted, there is an ~40% reduction in the ratio of MHC-β-gal to MRP-β-gal compared to gD1048, indicating a loss of splicing efficiency in this construct.

both the *Mhc* promoter, which is active in all muscles, and the *Mrp* promoter, which is active in many muscles but not in the IFM or TDT. Transcripts arising from the minigene produce both a β-gal fusion protein that is 175 kD (MHC-β-gal) for *Mhc* generated transcripts and a 125 kD (MRP-β-gal) from the *Mrp* promoter. Because the fusion MRP-β-gal protein is not dependent on splicing for its expression, changes in the ratio between MHC-β-gal/MRP-β-gal that occur following mutations to the minigene exon 11 domain are interpreted to indicate changes in splicing efficiency or transcript stability. Further, the level of the 125-kD MRP-β-gal product serves a control for positional effects such that the level of MHC-β-gal can be compared among the different constructs.

When this analysis was applied to the gD1090 or gD1177 constructs, which have a consensus donor at either 11e or 11b, respectively, it was found that the ratio of the MHC-β-gal products to the MRP-β-gal products was 1:1 for gD1090 and 1.2:1 for gD1177, while the ratio from the wild-type construct gD1048 is 1.3:1 (Figure 10). Thus, while all processed transcripts from gD1090 or gD1177 include the exon that has the consensus donor, this appears to result in neither a marked reduction or enhancement in the level of MHC-β-gal. Thus, unless translation rates are severely limiting, this result indicates that processing rates are not affected by the inclusion of a consensus donor into an alternative exon 11. Further, these results indicate that aberrant splicing of multiple alternatives and rapid turnover of transcripts are not induced by the inclusion of a consensus donor, at least not to the level where measurably lower levels of functional transcripts are available for translation. These results have important implications for the mechanism of splice choice selection, as will be discussed below. Also examined was the gD1105 construct, in which the splice donors of exons 11e and 11b are switched. In this clone, the ratio of MHC-β-gal product to MRP-β-gal product is 1:1, indicating that exchanging the 5′ splice donors does not lead to a marked enhancement or reduction of splicing efficiency, which supports the neutrality of these elements in alternative exon specification.

Interestingly, the gD1120 construct, which lacks the ICR but retains proper alternative exon specification, has a ratio of 0.75:1.0 between the MHC and MRP forms of β-gal, which is <60% of the wild-type (gD1048) ratio. Although the gD1120 deletion may alter directly the activity of the two promoters, thereby changing the ration of the two products in our assay, this result might also indicate that the ICR functions as a general element important for splicing efficiency of all alternative exons. The ICR is in the position generally associated with the location of branchpoint sequences (MOUNT *et al.* 1992) and, although no branchpoint consensus sequences occur in the ICR, its activity might be related to the function of branchpoints elsewhere in this domain. Additionally, the short purine-rich domains in the ICR resemble several known splice enhancers (SUN *et al.* 1993; LYNCH and MANIATIS 1996), thus the ICR might serve as a general exon 11 splicing enhancer.

DISCUSSION

The experiments presented above address the mechanisms that regulate alternative splicing of the Drosophila *Mhc* primary transcript in individual muscles. Using transgenic analysis and directed mutations, we demonstrate that large, complex alternatively spliced exon groups can be successfully studied *in vivo*. These data, summarized in Figure 11, show that *Mhc* exon 11 *cis*-regulation is local to this exon, and that *Mhc* alternative splicing is directed by multiple sequence elements that include nonconsensus 5′ splice donors and intronic elements that are local to the alternative exon.

***Cis*-regulation is local to exon 11:** Our findings that the isolated exon 11 domain in the context of a transgenic minigene is properly splice regulated demonstrates that a system to coordinate splicing across alternative exon groups is not obligatory for exon 11 muscle-type specific splicing. Such a system has been implied by the observation that there is a restriction in

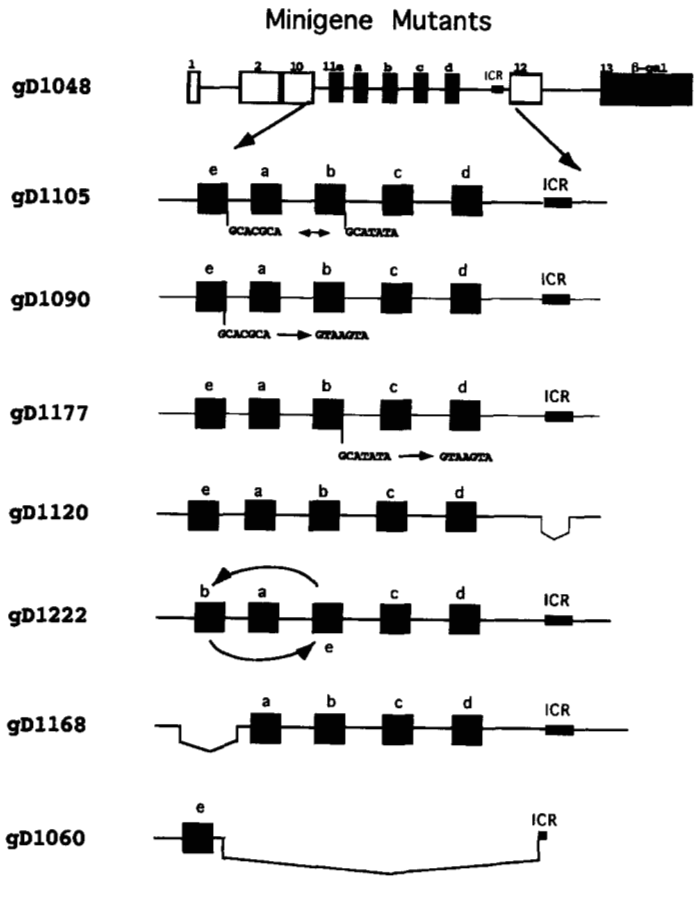

FIGURE 11.—Summary diagram of *Mhc* minigenes used in this study listing the mutation and the effect of the mutation on alternative splicing regulation.

the usage of exons from different groups and some exon combinations are not produced, indicating that the inclusion or exclusion of a particular exon from one alternative group influences the inclusion or exclusion of alternatives from other exon groups. The ability of exon 11 to regulate its own alternative splicing appears to be a common feature of *Mhc* alternative exon splicing in general because neither the disruption of *Mhc* exon 9 splicing in the IFM in the *Mhc11* mutation (KRONERT *et al.* 1991) or the truncation of the *Mhc* mRNA at exon 7 caused by an inserted transposable element in the *Mhc2* mutation (MOGAMI *et al.* 1986; M. B. DAVIS, J. DEITZ and C. P. EMERSON, unpublished data) disturb the processing of other alternative exon groups.

Nonconsensus splice donors are essential: Splice donor swap experiments (gD1105; Figure 6) show that *Mhc* splice donors are interchangeable, at least between exons 11e and 11b and, thus, we conclude that the nonconsensus donors do not provide information sufficient to direct muscle-specific splice choice. However, when these elements are replaced with consensus donors, splicing regulation is lost such that the exon with the consensus donor is spliced in the place of the other exons. The equivalent nature of the native donors and the ability of a consensus donor to repress their ability to participate in the splicing reaction suggest that the exon 11 splice donors function to provide a splice site that is not effectively recognized by the constitutive splicing machinery, but can be selectively activated through mechanisms separate from normal splicing. Nonconsensus donors are often associated with alternatively spliced exons and appear to be generally important for the selective enhancement of alternative exons (HODGES and BERNSTEIN 1994; BLACK 1995; ZHANG *et al.* 1996), but their role in an alternative exon groups as complex as *Mhc* exon 11 has not been previously determined. All *Mhc* alternative exons have donors that are nonconsensus and their function may be similar to that seen for the exon 11 donors. It is interesting to note, however, that the degree of nonconformity to the consensus differs markedly among donors both within and between exon groups. For instance, exons 11e, 11a, 11b and 11c all have a C at the +2 position (AAG|GC . . . . .) in the donor, which is a very rare nucleotide at this position (occurring in <1% of donor sites examined by MOUNT *et al.* 1992), while exon 11d is consensus at this site (AAG|GT . . . . .). Further, exon 3b is nonconsensus only at the −3 and −1 positions (TAC|GTAAGTA), which is similar to the degree

of nonconformity to consensus seen in several of the common exon donor sequences. Such observations suggest that the role of splice donors may differ among alternatives.

The loss of splice-choice specificity in exon 11 by the replacement of nonconsensus donors with consensus donors (gD1090; gD1177: Figure 6) is in contrast to the results seen for the differentially included *Mhc* exon 18, in which both the splice acceptor and splice donor were converted to consensus before the splicing mechanism became deregulated (HODGES and BERNSTEIN 1992). Thus, for alternative exons 11e and 11b, only a change to consensus in the splice donor is required to convert these alternatives to constitutively spliced exons. This observation suggests that the exon 11 3′ splice acceptors, which all contain poor polypyrimidine tracts, might serve to balance the nonconsensus donors in splice strength, but do not provide information directing the specificity of the splicing reaction.

An interesting result of the splice donor conversion to consensus is the observation that the now constitutively spliced consensus-donor exon continues to be normally spliced downstream to exon 12. The exon definition model of splicing regulation (ROBBERSON *et al.* 1990; BERGET 1995) predicts that the consensus donor in exon 11e (gD1090) would permit the normal splicing of this exon in the IFM, since no other exon is normally selected here. In the TDT, however, this model predicts the normal definition of exon 11b, but exon 11e should also be defined by virtue of its consensus donor. Thus, in the TDT, both 11e and 11b should be defined and spliced to each other. Our observation that the presence of a consensus donor in exon 11e completely represses the use of the normal exon and does not result in the aberrant splicing of multiple alternatives indicates that mechanisms in addition to exon definition are required to direct *Mhc* alternative exon selection. Such additional mechanisms are likely to involve those that enforce mutual exclusivity. In the case of the either gD1090 or gD1177, while both the normal exon and the consensus exon might be defined, the mutual exclusivity mechanism prevents the splicing of the alternatives to each other. Exon use might then be determined by competitive effects that act through the splice donors. How mutual exclusivity is maintained in the presence of multiply defined alternatives is not clear, but could involve incompatible branchpoints or splice donors and/or acceptors or other novel elements.

Analysis of the ICR, exon sequence and exon position: Although the conservation and position of the ICR in the exon 11 domain suggested a role in alternative splicing regulation, our analysis revealed that the deletion of this domain did not affect muscle-specific alternative exon use and, therefore, the ICR appears not to play a role in the regulation of alternative splice choice specification. However, the conservation of this domain across *D. melanogaster, D. virilis* and *D. hydei* suggests some function for the ICR and the observation that the expression of the MHC-β-gal reporter protein is diminished in the gD1120 construct, where the ICR is removed, suggests that it may enhance the rate of transcript processing. In support of this are the several domains of purine-rich sequence in the ICR whose composition is reminiscent of several known splicing enhancers. Given the nonconsensus donors and the poor polypyrimidine tracts found in all exon 11 alternatives, an element that could function to generally enhance the splicing of several alternatives would be a novel and intriguing possibility.

While exonic enhancers are found in a number of alternatively spliced transcripts, including the *doublesex* transcript in Drosophila (LYNCH and MANIATIS 1995; HERTEL *et al.* 1996), such a mechanism appears not to play a role in determining splice specificity in *Mhc* exon 11. This conclusion was drawn from the gD1222 construct in which exons 11e and 11b were positionally swapped, resulting in their expression also being swapped such that 11b was expressed on the IFM and 11e was expressed in the TDT. Thus, exons 11e and 11b are dependent on information found elsewhere in the exon 11 domain for their correct splice specification. This result is interesting because it suggests that the evolution of the protein-coding capacity of the exon 11 alternatives is not tied to their muscle-type specific splicing regulation. Whether this indicates that the mechanisms to direct alternative splicing existed prior to the elaboration of the complex *Mhc* alternative exon groups or if both protein coding capacity and the splicing regulatory mechanism co-evolved will be an intriguing question for future analysis. However, while these results demonstrate that the specificity of alternative exon use does not depend on sequence information contained in the alternative exons themselves, these data do not exclude the existence of general splicing enhancers within the alternative exons.

A mechanism to explain the interchangeability of the alternatives seen in the gD1222 construct is one that allows the alternative splicing apparatus to scan the entire exon 11 domain and select the appropriate exon by its position relative to other alternatives. For instance, in the IFM, such a scanning mechanism might always choose the exon in the position most proximal to exon 10, while in the TDT, the third exon is scanned and used. Such a mechanism has been demonstrated for other gene transcripts both *in vitro* and *in vivo*, where splice site selection can be influenced through the antagonistic activities of the splicing factors: ASF/SF2 and the hnRNP A1 proteins (MAYEDA and KRAINER 1992; MAYEDA *et al.* 1993; KRAUS and LIS 1994). Although conservation of the order and spacing of the exon 11 alternatives supports this model, the failure of the IFM to select a single exon in the gD1168 transgene shows that splice choice in this muscle is not directed simply

to the exon most proximal to exon 10. Further, since the pattern of usage for the other alternative exons was the same as in the wild-type, our results indicate that splice choice specification is not defined by the linear order of the exons and is, then, not likely to depend on mechanisms that work through the balance of splice factors. Further, these data are consistent with the results of the gD1105 donor swap experiments, which demonstrate that the alternative splice choice is not determined through a competition hierarchy among splice elements. Rather, these data indicate that alternative splicing requires elements that are directly involved in the specification and activation of the correct alternative exon and that these are likely localized to the intron.

Exon 11e is specified by local intronic sequences: In the gD1060 construct, we show that the alternative exon 11e in association with a limited domain of flanking intronic sequence and, in the absence of any other alternative exon, is still correctly and exclusively specified for inclusion in the IFM. These data show that the sequence required to direct this muscle-type specific choice is entirely contained within this construct. Further, we have shown that the splice donor and exon 11e itself are not important for splice specificity, demonstrating that the gD1060 IFM-specific alternative splicing regulatory elements are contained within the 360 nt of intronic sequence included in this construct. Examination of this domain from *D. melanogaster, D. virilis* and *D. hydei* (MEIDEMA *et al.* 1994) reveals several well-conserved sequences located in the intron between exons 11e and 11a (Figure 12). These first of these conserved elements (E-I) consist of two repeats in *D. melanogaster* and *D. virilis,* but only one perfectly conserved sequence in *D. hydei.* Interestingly, the second element, E-II, is separated by 24 nt in *D. melanogaster* and *D. virilis,* but is contiguous with E-I in *D. hydei.* Small, conserved intronic elements are associated with alternatively spliced exons in other Drosophila gene transcripts (THACKERAY and GANETZKY 1995) and have been shown to play a role in splicing regulation in other alternatively spliced transcripts as well. For instance, the repeated hexanucleotide TGCATG regulates the alternative splicing of the fibronectin gene transcript and can be found in several other alternatively spliced transcripts (HUH and HYNES 1994). A similar role for the small conserved elements identified in the gD1060 construct is, therefore, suggested.

Alternative splicing regulation of *Mhc* exon 11: Consistent with the data presented here, the intronic elements identified in the gD1060 construct are modeled to function in alternative splicing regulation either through the repression of exon 11e in all muscles except the IFM or by activating this exon exclusively in the IFM. Specifically, negative regulation would involve intronic elements that interact with *trans*-acting factors to repress the use of exon 11e. Such repressive factors

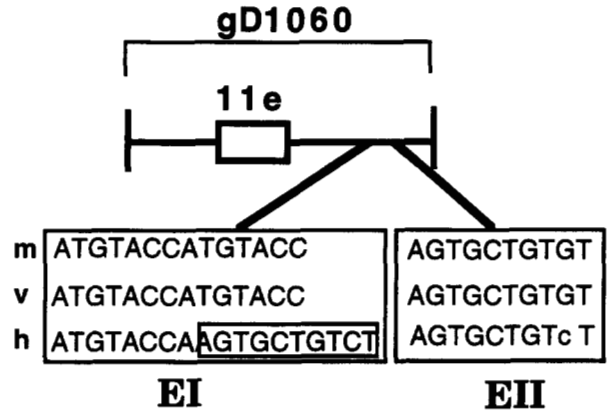

FIGURE 12.—Small conserved sequence elements retained in the gD1060 construct are potential sites of muscle-specific *cis*-regulation of alternative splicing. The 5′ element E-I is repeated twice in *D. melanogaster* (m) and *D. virilis* (v) where it is both a direct and an inverted repeat, but is only a single element in *D. hydei* (h). Interestingly, the second element E-II is 24 nt downstream of E-I in *D. melanogaster* and *D. virilis,* but is contiguous with E-I in *D. hydei.*

would be present in all muscles except for the IFM, which would instead express the factors required to suppress the inclusion of the other alternatives in the exon 11 group. Similarly, each muscle would then express the set of factors required to repress all exons except the included one.

A positively acting model of alternative splicing regulation is mechanistically simpler, involving only interactions between the intronic sequences in gD1060 and IFM-specific splicing factors that together direct the inclusion of exon 11e in the IFM. A model for the positive regulation of exon 11e in the IFM is diagramed in Figure 13 and depicts interactions between intronic elements that are local to the exon, which serve as alternative splicing specificity elements (hypothetically E-I and E-II) to provide information to direct the IFM-specific use of this exon, and the nonconsensus 5′ splice donor, which provides a context that is permissive for alternative splice selection. This interaction is mediated through IFM-specific *trans*-acting splice-activation factors (SAF) that associate specifically with the intronic specificity elements. This interaction then serves to attract the splicing apparatus such that the exon can be defined and incorporated into the processed message. This process might occur by the direct interactions of the SAFs and standard splicing factors, such as the U1 snRNP (REED 1996), or might require intermediate interactions with accessory splicing factors, such as the SR proteins (MANLEY and TACKE 1996), similar to the assembly of the *dsx*RE-associated complex in the positive activation of the female-specific exon of the *doublesex* transcript (HEINRICHS and BAKER 1995; LYNCH and MANIATIS 1995; LYNCH and MANIATIS 1996; HEINRICHS and BAKER 1997). In the case of *Mhc* exon 11 splicing, the use of intermediate factors might be indicated by the ability of an alternative exon that contains

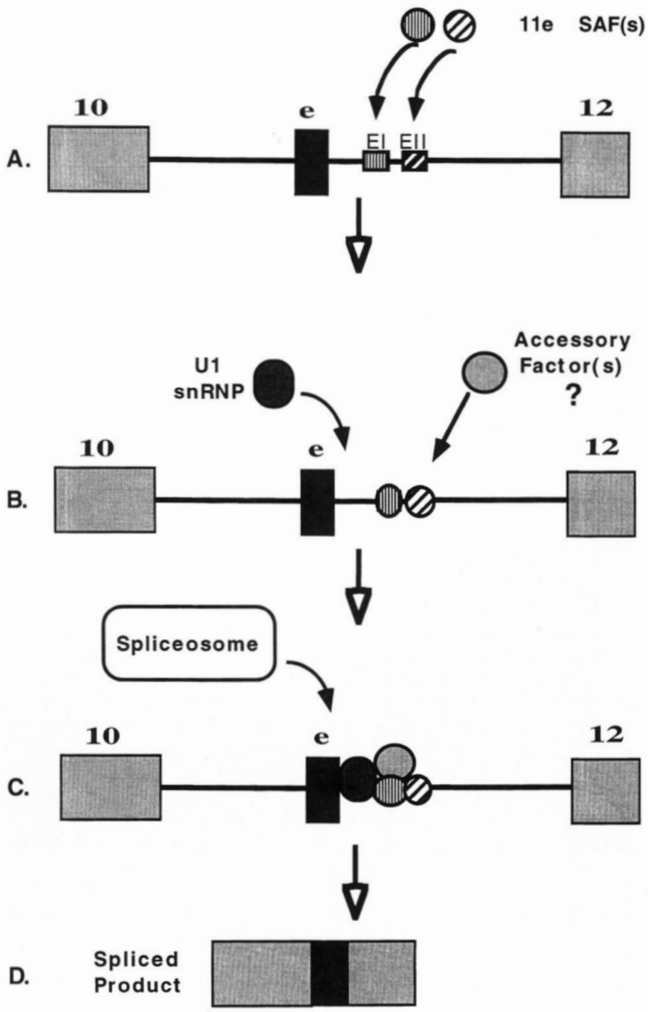

FIGURE 13.—Model of IFM-specific splicing of exon 11e. (A) Intronic specificity elements (possibly E-I and E-II) are specifically recognized by SAF. These factors are, in the case of exon 11e splicing, expressed only in the IFM. (B) This interaction permits the recruitment of the early components of the regular splicing machinery, such the U1snRNP, either directly or through interactions with additional accessory factors. (C) This complex then induces the assemble of the spliceosome, which results in the inclusion of exon 11e specifically in the IFM (D).

a consensus donor to completely repress the normal alternative, since a multistep assembly pathway might depress the normal assembly kinetics compared to that of an alternative exon that contains a consensus donor. The alternative splicing-specificity elements and the SAFs are required to enhance the ability of the splicing apparatus to recognize the alternative exon nonconsensus splice donor, thus promoting the inclusion of only that exon in the processed message. Similar splice-specificity elements might be associated with the all of the exon 11 alternatives and these could serve to interact with different SAFs to regulate the inclusion of the correct exon in the appropriate muscle.

Although our data do not distinguish between the negative or positive models of alternative splicing regulation, the predicted effects of removal of the intronic sequences, if negative regulation occurs, is promiscuous expression of an alternative exon, while, if positively regulated, their removal would result in the loss of expression. Thus, the mechanism can be readily identified with the further analysis of the alternative splicing regulatory elements.

Conclusion: In this paper, we have used the Drosophila *Mhc* exon 11 group as a novel system for the *in vivo* analysis of alternative splicing in muscle. This system has allowed us to examine a large and complex exon group with multiple alternatives and to address the mechanisms through which individual muscles specify and direct the inclusion of particular alternatives into the processed message. We show here that the alternative splicing regulatory system relies on a combination of *cis*-elements that include nonconsensus splice donors and intronic elements to direct the use of alternatives in individual muscles. The continued study of the *Mhc* exon 11 will allow the further description of these sequence elements and, in combination with the genetic analysis available in Drosophila, the identification of *trans*-acting factors that are important for alternative splicing regulation. Ultimately, the use of this system will allow a better understanding of the mechanisms that regulate alternative splicing in muscle and alternative pre-mRNA processing in general.

We thank BRIAN BRUNK and ELIZABETH WILDER for their constructive comments on this manuscript and we thank members of the EMERSON laboratory for their interest and discussion throughout the course of this project. This work was initiated at the Fox Chase Cancer Center. This work was supported by a U.S. Army Breast Cancer Initiative Postdoctoral Fellowship (D.M.S.), a Muscular Dystrophy Association Postdoctoral Fellowship (M.B.D.), an American Cancer Society grant NP-841 (C.P.E.) and by a National Institutes of Health R01-AR42363 grant (C.P.E.).

LITERATURE CITED

ABMAYR, S. M., M. S. ERICKSON and B. A. BOUR, 1995 Embryonic development of the larval body wall musculature of *Drosophila melanogaster*. Trends Genet. **11:** 153–159.

ASHBURNER, M., 1989 *Drosophila: A Laboratory Handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

BANDMAN, E., 1992 Contractile protein isoforms in muscle development. Dev. Biol. **154:** 273–283.

BERGET, S. M., 1995 Exon recognition in vertebrate splicing. J. Biol. Chem. **270:** 2411–2414.

BERNSTEIN, S. I., K. MOGAMI, J. J. DONADY and C. P. EMERSON, JR., 1983 Drosophila muscle myosin heavy chain encoded by a single gene in a cluster of muscle mutations. Nature **302:** 393–397.

BERNSTEIN, S. I., P. T. O'DONNELL and R. M. CRIPPS, 1993 Molecular genetic analysis of muscle development, structure, and function in Drosophila. Int. Rev. Cytol. **143:** 63–152.

BEVERLEY, S. M., and A. C. WILSON, 1984 Molecular evolution in Drosophila and the higher Diptera II. A time scale for fly evolution. J. Mol. Evol. **21:** 1–13.

BLACK, D. L., 1995 Finding splice sites within a wilderness of RNA. RNA **1:** 763–771.

BRAY, S. J., and J. HIRSH, 1986 The *Drosophila virilis* dopa decarboxylase gene is developmentally regulated when integrated into *Drosophila melanogaster*. EMBO J. **5:** 2305–2311.

CACERES, J. F., S. STAMM, D. M. HELFMAN and A. R. KRAINER, 1994 Regulation of alternative splicing *in vivo* by overexpression of antagonistic splicing factors. Science **265:** 1706–1709.

CHABOT, B., 1996 Directing alternative splicing: cast and scenerios. Trends Genet. **12:** 472–478.

CLACKSON, T., G. M. REINACH, G. M. CHUMBLEY and A. R. MACLEOD, 1989 Sticky feet directed mutagenesis and its application to swapping antibody domains. Nucleic Acids Res. **17:** 10163–10170.

CLOUET D'ORVAL, B., Y. D'AUBENTON CARAFA, P. SIRAND-PUGNET, M. GALLEGO, E. BRODY *et al.*, 1991 RNA secondary structure repression of a muscle-specific exon in HeLa cell nuclear extracts. Science **252:** 1823–1828.

CSANK, C., F.M. TAYLOR and D.W. MARTINDALE, 1990 Nuclear pre-mRNA introns: analysis and comparison of introns sequences from *Tetrahymena thermophila* and other eukaryotes. Nucleic Acids Res. **18:** 5133–5141

DEAR, S., and R. STADEN, 1991 A sequence assembly and editing program for the efficient management of large projects. Nucleic Acids Res. **19:** 3907–3911.

EMERSON, C. P., JR., 1993 Skeletal myogenesis: genetics and embryology to the fore. Curr. Opin. Genet. Dev. **3:** 265–274.

EPSTEIN, H. F., and S. I. BERNSTEIN, 1992 Genetic approaches to understanding muscle development. Dev. Biol. **154:** 231–244.

FYRBERG, E., and C. BEALL, 1990 Genetic approaches to myofibril form and function in Drosophila. Trends Genet. **6:** 126–131.

GEORGE, E. L., M. B. OBER and C. P. EMERSON, JR., 1989 Functional domains of the *Drosophila melanogaster* muscle myosin heavy-chain gene are encoded by alternatively spliced exons. Mol. Cell. Biol. **9:** 2957–2974.

GRAHAM, I. R., M. HAMSHERE and I. C. EPERON, 1992 Alternative splicing of a human alpha-tropomyosin muscle-specific exon: identification of determining sequences. Mol. Cell. Biol. **12:** 3872–3882.

HARRIDGE, S. D. R., R. BOTTINELLI, M. CANEPARI, M. A. PELLEGRINO, C. REGGIANI *et al.*, 1996 Whole-muscle and single-fibre contractile properties and myosin heavy chain isoforms in humans. Pflueg. Arch. Eur. J. Physiol. **432:** 913–920.

HASTINGS, G. A., and C. P. EMERSON, JR., 1991 Myosin functional domains encoded by alternative exons are expressed in specific thoracic muscles of Drosophila. J. Cell Biol. **114:** 263–276.

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.

HEINRICHS, V., and B. S. BAKER, 1997 *In vivo* analysis of the functional domains of the Drosophila splicing regulator RBP1. Proc. Natl. Acad. Sci. USA **94:** 115–120.

HERTEL, K. J., K. W. LYNCH, E. C. HSIAO, E. H. LIU and T. MANIATIS, 1996 Structural and functional conservation of the Drosophila *doublesex* splicing enhancer repeat elements. RNA **2:** 969–981.

HODGES, D., and S. I. BERNSTEIN, 1992 Suboptimal 5′ and 3′ splice sites regulate alternative splicing of *Drosophila melanogaster* myosin heavy chain transcripts *in vitro*. Mech. Dev. **37:** 127–140.

HODGES, D., and S. I. BERNSTEIN, 1994 Genetic and biochemical analysis of alternative RNA splicing. Adv. Genet. **31:** 207–281.

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.

KRAUS, M. E., and J. T. LIS, 1994 The concentration of B52, an essential splicing factor and regulator of splice site choice *in vitro*, is critical for Drosophila development. Mol. Cell. Biol. **14:** 5360–5370.

KRONERT, W. A., K. A. EDWARDS, E. S. ROCHE, L. WELLS and S. I. BERNSTEIN, 1991 Muscle-specific accumulation of Drosophila myosin heavy chains: a splicing mutation in an alternative exon results in an isoform substitution. EMBO J. **10:** 2479–2488.

LIBRI, D., A. PISERI and M. Y. FISZMAN, 1991 Tissue-specific splicing *in vivo* of the beta-tropomyosin gene: dependence on RNA secondary structure. Science **252:** 1842–1845.

LOWEY, S., G. S. WALLER and K. M. TRYBUS, 1993 Function of skeletal muscle myosin heavy and light chain isoforms by an *in vitro* motility assay. J. Biol. Chem. **268:** 20414–20418.

LYNCH, K. W., and T. MANIATIS, 1995 Synergistic interactions between two distinct elements of a regulated splicing enhancer. Genes Dev. **9:** 284–293.

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.

MANLEY, J. L., and R. TACKE, 1996 SR proteins and splicing control. Genes Dev. **10:** 1569–1579.

MAYEDA, A., and A. R. KRAINER, 1992 Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell **68:** 365–375.

MAYEDA, A., D. M. HELFMAN and A. R. KRAINER, 1993 Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. Mol. Cell. Biol. **13:** 2993–3001.

MCCULLOUGH, A. J., and M. A. SCHULER, 1993 AU-rich intronic elements affect pre-mRNA 5′ splice site selection in *Drosophila melanogaster*. Mol. Cell. Biol. **13:** 7869–7697.

MIEDEMA, K., H. HARHANGI, S. MENTZEL, M. WILBRINK, A. AKHMANOVA *et al.*, 1994 Interspecific sequence comparison of the muscle-myosin heavy-chain genes from *Drosophila hydei* and *Drosophila melanogaster*. J. Mol. Evol. **39:** 357–368.

MIEDEMA, K., M. HANSKE, A. AKHMANOVA, P. BINDELS and W. HENNIG, 1995 Minor-myosin, a novel myosin isoform synthesized preferentially in *Drosophila testis* is encoded by the muscle myosin heavy chain gene. Mech. Dev. **51:** 67–81.

MOGAMI, K., P. T. O'DONNELL, S. I. BERNSTEIN, T. R. WRIGHT and C. P. EMERSON, JR., 1986 Mutations of the Drosophila myosin heavy-chain gene: effects on transcription, myosin accumulation, and muscle function. Proc. Natl. Acad. Sci. USA **83:** 1393–1397.

MOUNT, S. M., C. BURKS, G. HERTZ, G. D. STORMO, O. WHITE *et al.*, 1992 Splicing signals in Drosophila: intron size, information content, and consensus sequences. Nucleic Acids Res. **20:** 4255–4262.

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.

REED, R., 1996 Initial splice-site recognition and pairing during pre-mRNA splicing. Curr. Opin. Genet. Dev. **6:** 215–220.

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.

ROBERTS, G. C., C. GOODING and C. W. J. SMITH, 1996 Smooth muscle alternative splicing induced in fibroblasts by heterologous expression of a regulatory gene. EMBO J. **15:** 6301–6310.

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.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SCHIAFFINO, S., and C. REGGIANI, 1996 Molecular diversity of myofibrillar proteins: gene regulation and functional significance. Physiol. Rev. **76:** 371–423.

STANDIFORD, D. M., M. B. DAVIS, K. MIEDEMA, C. FRANZINI-ARMSTRONG and C. P. J. EMERSON, 1997 Myosin rod protein: a novel thick filament component of Drosophila muscle. J. Mol. Biol. **265:** 40–55.

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.

THACKERAY, J. R., and B. GANETZKY, 1995 Conserved alternative splicing patterns and splicing signals in the Drosophila sodium channel gene *para*. Genetics **141:** 203–214.

THUMMEL, C. S., A. M. BOULET and H. D. LIPSCHITZ, 1988 Vectors for Drosophila P-element-mediated transformation and tissue culture transfection. Gene **74:** 445–456.

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.

ZARET, K. S., J. LIU and C. M. DIPERSIO, 1990 Site-directed mutagenesis reveals a liver transcription factor essential for the albumin transcriptional enhancer. Proc. Natl. Acad. Sci. USA **87:** 5469–5473.

ZHANG, L., M. ASHIYA, T. G. SHERMAN and P. J. GRABOWSKI, 1996 Essential nucleotides direct neuron-specific splicing of g2 pre-mRNA. RNA **2:** 682–698.

Communicating editor: V. G. FINNERTY