Transposable Element Insertions Respecify Alternative Exon Splicing in Three Drosophila Myosin Heavy Chain Mutants

Mary Beth Davis, Jon Dietz, David M. Standiford and Charles P. Emerson, Jr.

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

Manuscript received April 20, 1998
Accepted for publication August 19, 1998

ABSTRACT

Insertions of transposable elements into the myosin heavy chain (Mhc) locus disrupt the regulation of alternative pre-mRNA splicing for multi-alternative exons in the Mhc2, Mhc3, and Mhc4 mutants in Drosophila. Sequence and expression analyses show that each inserted element introduces a strong polyadenylation signal that defines novel terminal exons, which are then differentially recognized by the alternative splicing apparatus. Mhc2 and Mhc4 have insertions located within intron 7c and exon 9a, respectively, and each expresses a single truncated transcript that contains an aberrant terminal exon defined by the poly(A) signal of the inserted element and the 3′ acceptor of the upstream common exon. In Mhc3, a poly(A) signal inserted into Mhc intron 7d defines terminal exons using either the upstream 3′ acceptor of common exon 6 or the 7d acceptor, leading to the expression of 4.1- and 1.7-kb transcripts, respectively. Acceptor selection is regulated in Mhc3 transcripts, where the 3′ acceptor of common Mhc exon 6 is preferentially selected in larvae, whereas the alternative exon 7d acceptor is favored in adults. These results reflect the adult-specific use of exon 7d and suggest that the normal exon 7 alternative splicing mechanism continues to influence the selection of exon 7d in Mhc3 transcripts. Overall, transposable element-induced disruptions in alternative processing demonstrate a role for the nonconsensus 3′ acceptors in Mhc exons 7 and 9 alternative splicing regulation.

Alternative pre-mRNA processing depends on cis-acting sequence elements to regulate the inclusion of an alternative exon in specific cells or tissues (reviewed in Black 1995). Often, such elements act as splicing enhancers, which work in conjunction with accessory splicing factors to activate the typically weak (nonconsensus) 3′ splice acceptors or 5′ splice donors of the regulated exons (Chan and Black 1995; Min et al. 1997; Wei et al. 1997). Alternatively, but less commonly, exon use can be repressed through elements that inhibit the function of the splice acceptor or donors of the regulated exon (Ashiya and Grabowski 1997). This direct modulation of splice-site activity applies to the alternative splicing regulation of transcripts from many genes in which a single exon is differentially included in or excluded from the processed message. However, a large number of transcripts contain multiple alternative exons that are spliced into the final message in a mutually exclusive fashion. For this class of transcripts, the splicing mechanism must ensure the tissue-specific selection of the correct exon while simultaneously preventing the recognition of inappropriate alternatives to prevent the incorrect exon from being included in the message, as well as to restrict the splicing of multiple alternatives to each other. While the mechanisms that regulate the processing of mutually exclusive, multi-alternative transcripts are not well understood, data from studies examining the alternative splicing of the α- and β-tropomyosin transcripts show that alternative splicing regulation can involve positive or negative cis-acting elements, general and specific splice factors (Lin and Patton 1995; Roberts et al. 1996; Perez et al. 1997; Gooding et al. 1998), steric interactions at the branchpoint of the regulated exon (Mullen et al. 1991), and secondary structure (Clout et al. 1991). These data demonstrate the potential complexity of the mechanisms that regulate the alternative splicing of transcripts containing multi-alternative, mutually exclusive exons, which perhaps reflects the fact that the differentially expressed protein isoforms encoded by these RNAs often underlie the functional distinctions between different cells and tissues (Nadal-Ginard et al. 1991; Rio 1993; Schiaffino and Reggiani 1996).

One of the most complex transcripts subject to mutually exclusive splicing is expressed from the single muscle myosin heavy chain (Mhc) gene in Drosophila. The Mhc gene contains 19 exons, of which 5 are present as alternatively spliced groups, each containing 2 to 5 members (Figure 1; George et al. 1989). The alternative

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

1These authors contributed equally to this work.
2Present address: Department of Biology, Bryn Mawr College, Bryn Mawr, PA 19010.
3Present address: Novartis Agricultural Biotechnology Research, Inc., 3054 Cornwallis Rd., P.O. Box 12277, Research Triangle Park, NC 27709.
exons from each group are spliced in a mutually exclusive fashion and all exons within each group encode variations of the same functionally important domain of the MHC protein (George et al. 1989; Bernstein and Milligan 1997). Although providing an enormous coding potential (a possible 480 MHC isoforms), it has been demonstrated that the alternative processing of the Mhc primary transcript is highly regulated, with precise splice patterns generated to ensure that typically only a single MHC isoform is expressed within each muscle (Hastings and Emerson 1991). As a result, alternative mRNA processing determines a key functional difference among the myriad muscles found in Drosophila. This connection between alternative splicing regulation and the functional identity of individual muscles provides a unique opportunity to examine the mechanisms that establish and enforce precise tissue-specific patterns of splicing in transcripts with complex alternative exons.

To examine the mechanisms that direct alternative splicing of the Mhc gene, we have analyzed the molecular nature of three mutants, Mhc2, Mhc3, and Mhc4. These mutants were isolated in a genetic screen to recover dominant flightless mutants (Mogami and Hotta 1981) and were subsequently identified as alleles of the Mhc gene (Mogami et al. 1986). In the preliminary characterization of these strains, Mogami et al. (1986) demonstrated by genomic Southern analysis that each mutant is associated with an insertion of an 8- to 10-kb DNA element, and the Northern blot analysis of pupal mRNAs revealed that each heterozygote produces a truncated Mhc mRNA in addition to the wild-type Mhc mRNAs. Subsequent work has demonstrated that each mutant results from a transposable element insertion (Dietz 1991; Cripps et al. 1994) and that these are localized to an alternative exon domain. We find here that each insert restructures the alternative exon domain with a concurrent respecification of alternative exon use during the processing of the mutant transcript. These alterations specifically change the use of only the inserted exon group and do not affect the processing of upstream alternative exons. Finally, we show that in one mutant, Mhc3, a previously unrecognized splice form results from the competition between terminal exons defined by the acceptors of either a common or represented transcripts from the wild-type form results from the competition between terminal exons. Five clones hybridized to both probes; these clones likely correspond to the approximate size of the truncated Mhc mRNA detected on Northern blots of Mhc2/SM1 adult RNAs (Mogami et al. 1986; this article). The cDNA insert was sequenced in both orientations into pEMBL18 and the insert was sequenced as described above.

**MATERIALS AND METHODS**

**Drosophila stocks and culture conditions:** Drosophila stocks were maintained at 25°C on standard cornmeal-molasses-agar medium. The Mhc mutants used in this study, Mhc2/SM1, Mhc3/SM1, and Mhc4/SM1, were originally isolated by Mogami and Hotta (1981) and are described by Mogami et al. (1986). Canton-S was used throughout these studies as the wild-type Drosophila strain. The second chromosome balancer SM1, which carries the Curly mutation, is described by Lindley and Zimm (1992) and by Ashburner (1989).

For the RNA analyses, eggs were collected for a 5-hr period from a population cage containing 4- to 5-day-old adults and allowed to develop at 25°C. Animals of different developmental stages were collected from the progeny of each Mhc mutant/SM1 heterozygous stock without the use of embryonic or larval markers. Homozygous Mhc2 and Mhc3 flies die during the larval period and the Mhc4 homozygotes die during the embryonic period (Mogami et al. 1986). For each mutant strain the 18- to 30-hr-late embryo/first-instar larval collections therefore consisted of three genotypes: Mhc mutant homozygotes, SM1 homozygotes, and Mhc mutant/SM1 heterozygotes. The 9-day pupae collected from each mutant stock, however, consisted essentially of only Mhc mutant/SM1 heterozygotes (the homozygous Curly flies, i.e., SM1 homozygotes, mostly die as larvae [Ashburner 1989]).

**RNA isolation and Northern analysis:** Total RNA from larvae, pupae, and adults was isolated using the TRIzol Reagent (GIBCO BRL, Gaithersburg, MD). Poly(A)+ mRNA was purified with oligo(dT) cellulose using the FastTrack mRNA Isolation Kit (Invitrogen Corp., San Diego, CA). mRNA was resolved on MOPS-formaldehyde-1% agarose gels (Struhl and White 1985), blotted onto Hybond-N nylon membranes (Amershams, Arlington Heights, IL), and cross-linked by ultraviolet light irradiation (Stratagene, La Jolla, CA) to the membranes. Northern blots were hybridized in 0.33 M sodium phosphate, pH 6.8, 10 mm EDTA, 10% dextran sulfate, 5% SDS, and 100 μg/ml sonicated denatured salmon sperm DNA (Struhl and White 1985) at 58°C. After hybridization the filters were washed as previously described (Struhl and White 1985) except that the temperature of the washes was 58°C.

**Radiolabeling of probes:** Restriction fragments of genomic and cDNA clones were random-primed labeled (Amershams) with [α-32P]dCTP (New England Nuclear, Boston) and oligonucleotides (see Table 1) were 3'-end labeled with terminal deoxynucleotidyl terminal transferase (Promega, Madison, WI). Hybridization signals were visualized either by standard autoradiography using Kodak XAR film (Eastman Kodak Co., Rochester, NY) or by imaging with a Molecular Dynamics (Sunnyvale, CA) PhosphorImager 445 SI and IPLab Gel Software (Signal Analytics Corp., Vienna, VA).

**Construction and screening of the Mhc2/SM1 cDNA library:** A cDNA library from Mhc2/SM1 first-instar larvae was constructed in lambda-gt10 and was screened according to the procedures in George et al. (1989). Probes corresponding to either the 5′ (genomic DNA spanning exons 2–7c) or the 3′ end (a cDNA clone containing exons 14, 15b, and 16) of the Mhc gene were used to screen duplicate filters from the unamplified cDNA library. Nine positive clones were recovered. Five clones hybridized to both probes; these clones likely represented transcripts from the wild-type Mhc allele present on the SM1 balancer chromosome. The remaining four clones hybridized only to the 5′ probe. Three clones had inserts that were <1 kb and were not characterized. One clone from this group, designated cDNA01, had an insert of ~3.7 kb, which corresponds to the approximate size of the truncated Mhc mRNA detected on Northern blots of Mhc2/SM1 adult RNAs (Mogami et al. 1986; this article). The cDNA01 insert was sequenced in both orientations into pEMBL18 and the insert was sequenced as described above.

**PCR amplification and RT protocols:** Genomic DNA for PCR analysis was isolated as described by McGinnis et al. (1983). Oligonucleotide primer pairs were designed for each mutant such that an interval could be amplified from the long terminal repeat (LTR) of the inserted element to flanking
genomic DNA from the Mhc allele. PCR reactions were performed using Perkin Elmer-Cetus (Norwalk, CT) reagents in a Perkin Elmer-Cetus PCR thermocycler. The PCR products were subcloned into either the pUC18 plasmid or Bluescript (Stratagene) plasmid by standard cloning methods (Sambrook et al. 1989) and were sequenced using the Sequenase kit from United States Biologicals. The sequence data were analyzed with the Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI.

Total RNA for reverse-transcriptase dependent PCR (RT-PCR) was isolated from first instar larvae or late pupae using the Trizol reagent (GIBCO BRL), and the RT reactions were performed as described in Standard (1997). A total of 10 μl of the RT reaction was used in each PCR. To recover the truncated Mhc cDNA fragment the RT reaction was primed with an oligo(dT) adaptor oligonucleotide; the primers for the subsequent PCR were a sense oligonucleotide in Mhc exon 4 and the oligo(dT) adaptor oligonucleotide. The PCR product was purified from an agarose gel using the Gene Clean kit (BIO-101 Inc., Vista, CA), then sequenced. cDNA fragments from the 1.7-kb Mhc mRNA were isolated from a RT reaction of total Mhc/SM1 larval RNA that was primed with an antisense oligonucleotide to the splicing LTR that had an Xho I restriction site incorporated onto the 5' end of the oligonucleotide. A primer from exon 4 and the LTR-XhoI primer were used to amplify the Mhc3 mutant RNA; the resulting PCR product was digested with StuI (site in exon 4) and XhoI (from the end of the LTR-XhoI primer) and subcloned into a Bluescript II plasmid vector (Stratagene) that had been digested with EcoRV and XhoI. The clones were analyzed by restriction enzyme digestion and DNA sequencing.

### RESULTS

**Determination of the molecular nature of the Mhc2, Mhc3, and Mhc4 mutant alleles:** Mhc2, Mhc3, and Mhc4 were isolated in a genetic screen to recover dominant flightless mutants (Mogami and Hotta 1981), and subsequently identified as mutations of the muscle myosin heavy chain (Mhc) gene (Mogami et al. 1986). Their initial characterization by genomic Southern analysis demonstrated that the mutants Mhc2, Mhc3, and Mhc4 were each associated with an insertion of an 8-10-kb DNA element within the Mhc locus. Later data (Dietz 1991; Cripps et al. 1994) suggested that the Mhc2 and Mhc3 mutations were due to insertions of the 8.8-kb copia-like transposable element springer, and that the Mhc4 mutation was due to the insertion of a B104 element. Using primers directed against sequences in the LTRs of the respective transposons and primers generated from the flanking DNA of the Mhc gene (Table 1; George et al. 1989), we amplified the segments of genomic DNA that corresponded to the junction fragments at the 5' and 3' ends of the inserted elements from each mutant in the Mhc gene. These restriction fragments were subcloned and sequenced to confirm the identities of the inserted elements in the different Mhc mutants as springer elements (Karlik et al. 1984) in Mhc2 and 3, and B104 (Scherrer et al. 1982) in Mhc4. The springer element in Mhc2 is inserted 5 bp 5' to exon 7d and in Mhc3 the element is inserted 117 bp 3' to exon 7d (Figure 1B). The springer insertion in both Mhc mutants is flanked by a 4-bp duplication of the target site; however, the target sites are different (Figure 1C).

**Exon 7 alternative splicing disrupted in Mhc2 transcripts:** Northern blot analysis of RNAs from late Mhc2/SM1 pupae using probes directed against Mhc exons 2, 3b, and 4, identifies a single truncated Mhc mRNA that is 3.7 kb in size, in addition to the three larger, normally processed transcripts of the wild-type Mhc allele (George et al. 1989), as has been previously shown by Mogami et al. (1986). In light of the complexity of the alternative RNA processing of the Mhc primary transcripts we were led to examine the structures of the mutant RNA transcripts to determine if these transposons and the activity of their associated polyadenylation signals alter Mhc alternative splicing.

### TABLE 1

<table>
<thead>
<tr>
<th>Oligonucleotides used in this study</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>oD166: LTR-XhoI antisense</td>
<td>CGGCCCTGAGTTTGACCTTGTGCTGGTGCGGACGATC</td>
</tr>
<tr>
<td>oD167: -exon 8 antisense</td>
<td>CATGACAGGGCGGTTGACCTGGTACAC</td>
</tr>
<tr>
<td>oD168: -exon 4 sense</td>
<td>GATCATGCTGGCTGGTCTCCGCGTG</td>
</tr>
<tr>
<td>oD169: -exon 4 antisense</td>
<td>GAGGCCCTGGTAAGCCAAAGACG</td>
</tr>
<tr>
<td>oD170: -exon 9a sense</td>
<td>ATTCCTAACCTCAATGTA</td>
</tr>
<tr>
<td>oD170: -exon 9a sense</td>
<td>TGCAGCTGGAATTCCCTTACCG</td>
</tr>
<tr>
<td>oD182: -exon 7d antisense</td>
<td>GGATACCCCGGTGCTCAACGATGG</td>
</tr>
<tr>
<td>oD183: -exon 7c sense</td>
<td>GGATACCCCGGTGCTCAACGATGG</td>
</tr>
<tr>
<td>oD167: -exon 4 antisense</td>
<td>GGATACCCCGGTGCTCAACGATGG</td>
</tr>
<tr>
<td>oD169: -exon 4 antisense</td>
<td>GGATACCCCGGTGCTCAACGATGG</td>
</tr>
<tr>
<td>oD170: -exon 9a sense</td>
<td>GGATACCCCGGTGCTCAACGATGG</td>
</tr>
<tr>
<td>oD182: -exon 7d antisense</td>
<td>GGATACCCCGGTGCTCAACGATGG</td>
</tr>
<tr>
<td>oD183: -exon 7c sense</td>
<td>GGATACCCCGGTGCTCAACGATGG</td>
</tr>
</tbody>
</table>
the ATG in exon 2 and is identical to the wild-type $M_{hc}$ mRNA through the end of exon 6. The transcript contains exon 3b as the exon 3 alternative. At exon 6, the sequence is the same as wild-type genomic sequence through the exon 7 domain and includes all introns and exons until 5 bases upstream of exon 7d (Figure 2). Here, the sequence matches that of the LTR from the springer element. The cDNA includes 198 bp of the springer element, excluding the poly(A) tail, and contains a consensus polyadenylation signal 21 bases from the cleavage/polyadenylation site. Thus, the processing of this truncated transcript is normal for exons 2-6, whereas all exon 7's are unprocessed and truncated by premature polyadenylation of the $M_{hc}$ mRNA by the polyadenylation signal provided in the springer LTR. Only this single mutant transcript was identified by RT-PCR analysis of pupal and larval mRNA collected from $M_{hc2}$ mutants using primers directed against $M_{hc}$ exon 4 and the springer LTR (not shown), indicating that exon 7 alternative splicing is completely blocked.

**Alternative processing of the $M_{hc3}$ mutant transcripts:** When RNAs collected from late $M_{hc3}$/SM 1 pupae were probed in a Northern blot with a cDNA probe consisting of $M_{hc}$ exons 2, 3b, and 4, RNAs of 4.1 and 1.7 kb were detected in addition to the normal $M_{hc}$ processed transcript (Figure 3). While the 1.7-kb transcript was not detected in the $M_{hc3}$ RNA by Mogami et al. (1986), it is perhaps due to the different stage at which the RNA was collected from mutant pupae in this study.

To determine the exon composition of the 4.1- and 1.7-kb mRNAs, Northern blots containing $M_{hc3}$/SM 1 pupal RNA were assayed with probes directed against the 7a intron or exon 7d. As seen in Figure 3, the intronic probe hybridizes to only the 4.1-kb RNA, whereas the exon 7d probe hybridizes to both RNAs. The presence of intronic sequence in the 4.1-kb RNA suggests this transcript contains a truncated message that is unprocessed across the exon 7 domain, while the size and lack of intronic sequence in the 1.7-kb transcript predicts that this RNA incorporates a processed exon 7. To examine this possibility further, RT-PCR using primers directed against the LTR of the springer insert and exon 4 was used to amplify and clone the 1.7-kb truncated product. Southern analysis of the RT-PCR product with probes against each of the exon 7 alternatives showed that only exon 7d hybridized (not...
shown), and restriction analysis of 30 individual clones indicated that all 1.7-kb transcripts contain only exon 7d. Further, several clones were sequenced showing normal processing for exon 4, 5, and 6. Exon 7d was also incorporated into the sequence with its normal acceptor spliced to exon 6, but the native 7d donor was not recognized, and instead the message was cleaved and polyadenylated within the springer element LTR at the site of the inserted poly(A) signal (Figure 3B). These results show that exon 7d is recognized by the splicing apparatus as part of a novel terminal exon defined by the exon 7d acceptor and the springer element poly(A) signal that can be differentially selected in addition to the exon 6-springer poly(A)-defined terminal exon.

Normally, the Mhc exon 7d is spliced into adult Mhc transcripts, but not those in larval muscles (George et al. 1989). To determine if the selection of the 7d-poly(A)-defined terminal exon was also subject to this regulation, a Northern blot using mRNA prepared from Mhc3/SM1 larvae was probed with exon 7d, which, as seen in Figure 3C, hybridizes to both the 4.1- and 1.7-kb mutant transcripts, but not to the wild-type Mhc transcripts. Thus, the 7d-poly(A) terminal exon is recognized in larvae, but, when compared to a Northern blot prepared from pupal Mhc3 RNA, the ratio of the 1.7- to 4.1-kb transcripts shifts from 1:3 in larvae to 3:1 in pupae. The increased use of the 7d-poly(A)-defined terminal exon in pupae is consistent with the normal adult-specific use of exon 7d, and these results suggest a model in which the 7d-poly(A) exon is defined at a low, constitutive level in muscles that do not ordinarily use exon 7d (i.e., larval and some adult muscles), but its recognition is enhanced by the alternative splicing pathway in muscles that normally select exon 7d. A possible outcome of this model is that in muscles where exon 7d is normally selected, only the exon 7d-poly(A) terminal exon will be recognized in mutant transcripts. To test this possibility, RNA isolated from the indirect flight muscle (IFM), where only exon 7d is used, or from whole pupae, was reverse transcribed using a springer LTR-specific primer. These fractions were then examined using PCR and exon 4- and 7d-specific primers to detect the 7d-poly(A)-defined terminal exon transcript, or exon 7c and 7d primers to detect the common exon 6-poly(A)-defined terminal exon product. As is seen in Figure 3D, fragments corresponding to the 1.7- and 4.1-kb transcripts were amplified in fractions from both the indirect flight muscle and whole pupae, showing that, while its recognition and use are enhanced in pupae, the exon 7d-poly(A) terminal exon is not exclusively defined in an adult muscle that normally selects alternative exon 7d.

Exon 9a alternative splicing is repressed in the Mhc4 mutant transcript: The B104 element in Mhc4 inserted into exons 9a (Figure 4) results in the aberrant production of a single, truncated 2.4-kb RNA (Mogami et al. 1986). The detection of this RNA with Northern analysis using a probe directed against Mhc intron 8 shows that this RNA is unprocessed between exons 8 and 9 (Figure 4A). Further, RT-PCR of pupal RNA with oligo(dT) and a primer to exon 6 generates a product of the

Figure 2.—(A) Northern blot analysis of mRNA collected from Canton-S (wild type) or Mhc2/SM1 pupae probed with cDNA homologous to Mhc exon 2, 3b, and 4, shows the presence of an aberrant 3.7-kb band only in the Mhc2 RNA. Other bands indicated are the normally processed 8.6-, 8.0-, and 7.2-kb mRNA arising from the wild-type Mhc allele(s). (B) Partial cDNA sequence of truncated mRNA cloned from Mhc2/SM1 flies. All isolated clones contained the normally spliced exons 1-6 and were then unspliced throughout the exon 7 domain to the site of the springer insertion, which is 738 nt downstream of exon 7c (underlined). The element (bold) is cleaved and polyadenylated at a site that is 19 nt downstream from a consensus polyadenylation signal (underlined).
Figure 3.—(A) Northern blot analysis of mRNA collected from Mhc3/SM1 pupae. When probed with labeled DNAs containing Mhc exons 2, 3b, and 4, transcripts of 4.1 and 1.7 kb are seen in Mhc3, but not in wild-type (Canton-S) mRNA. Both the 4.1- and 1.7-kb transcripts are detected by probes containing exon 7d, but only the 4.1-kb RNA hybridizes to a probe that contains Mhc intron 7a. (B) Partial sequence of RT-PCR product generated from the 1.7-kb Mhc3 truncated mRNA in which Mhc exons 5 and 6 are correctly spliced. Exon 6 is spliced to exon 7d via the normal exon 6 splice donor and 3' acceptor of exon 7d, but the 7d donor (boxed) is not recognized and the message contains unprocessed 7d intronic sequence (small case) and is truncated within the LTR of the springer element (bold). 21 nt downstream of a consensus polyadenylation signal (AAUAAA). (C) Northern analysis of Mhc3/SM1 mRNA collected from pupae and larvae showing that the 1.7-kb mRNA is generated at a higher level than the 4.1-kb transcript in pupae, where the ratio is 3:1, while in larvae the ratio is 1:3, showing the 4.1-kb transcript is favored. (D) RT-PCR assay to examine the alternative splicing of the Mhc3 transcript in an individual adult muscle. RNAs collected from the IFM or from whole pupae were reverse-transcribed using a primer specific for the LTR of springer and primers for Mhc exon 4 and 7d amplify a product of 325 bp in RNAs from the IFM and whole pupae, indicating the presence of the 1.7-kb mRNA in both fractions. A 900-bp product is generated from exons 7c- and 7d-specific primers in the IFM and whole pupae RNAs, indicating the presence of the 4.1-kb mRNA in both fractions. No products were detected in reactions performed without RT (not shown).

appropriate size for the Mhc4 mutant message that contains a normally processed exon 7 alternative, but is unprocessed in intron 8 (not shown). Sequencing of this product shows that the mutant transcript includes 35 nucleotides of exon 9a followed by 402 bp from the B104 LTR, which contains a consensus polyadenylation signal 23 bases upstream of the poly(A) tail. Together, these results show that, similar to the Mhc2 and the 4.1-kb Mhc3 truncated mRNAs, the mutant Mhc4 mRNA is correctly spliced upstream of the domain containing the transposon, while RNA processing in the alternative exon 9 domain is blocked and the mRNA is polyadenylated at a site provided by the B104 LTR.

On the basis of the analyses performed here, we conclude that the local insertions of transposable elements into the alternatively processed exon 7 and 9 domains result in either a complete or a partial blockage of alternative splicing, and this is associated with the recognition and use of polyadenylation signals associated with the inserted elements. The results of this analysis are depicted in Figure 5 and are discussed below.

DISCUSSION

The molecular effects of the Mhc2, 3, and 4 mutations on alternative splicing have been examined. Each mu-
Alternative Splicing in Drosophila

**Figure 4.** (A) Northern analysis of mRNA collected from Mhc4/SM1 pupae. When probed with labeled DNA containing Mhc exons 2, 3b, and 4, a 2.4-kb transcript is detected in mutant RNAs that is not present in the wild-type (Canton-S) mRNA. This same band is detected with a DNA containing Mhc intron 8, indicating that this band is not processed normally. (B) Partial sequence of RT-PCR product generated from the Mhc4 mutant transcript using a primer directed against intron 8 and oligo(dT). Shown is a portion of exon 9a (underlined), which is interrupted by the insertion of the B104 element (bold). The element is cleaved and polyadenylated 21 nt downstream of a consensus poly(A) signal (AAUAAA).

**Polyadenylation signal insertions respecify alternative splicing:** In each Mhc mutant examined here, the insertion of a transposable element introduces a poly(A) signal into an alternatively spliced exon domain. In addition to directing the cleavage and polyadenylation reactions during mRNA 3' end processing, polyadenylation

**Figure 5.** Model showing the effects of inserted poly(A) on splicing in the Mhc2, 3, and 4 mutants. In the Mhc2 mutant, the inserted element contributes a poly(A) signal downstream of exon 7c. This site functions to define the 3' end of a 2431-nt terminal exon in a splicing reaction, while the 5' junction is defined in this mutant exclusively by the consensus (CAG) 3' acceptor of the common exon 6, which likely serves as a better splicing substrate than the nonconsensus splice acceptors (TAG) of the exon 7 alternatives. The Mhc3 insert places a poly(A) site downstream of exon 7d and is compatible with the formation of two terminal exons at this position. A 2658-nt terminal exon is defined by the poly(A) signal and the consensus acceptor of the common exon 6. The selection of the nonconsensus acceptor of alternative exon 7d, which is alternatively used to define a 420-nt terminal exon, is potentially enhanced by its proximity to the poly(A) signal. The further use to this site seen in the increased level of expression of the 1.7-kb Mhc3 transcript in adult muscle, suggests that the activity of the 7d acceptor is sensitive to the mechanism that normally directs the adult-specific use of alternative exon 7d. The Mhc4 insertion places a poly(A) signal within alternative exon 9a, which represses exon 9 alternative splicing and instead directs the expression of a 437-nt terminal exon that is defined by the poly(A) signal and the consensus donor of the common exon 8. Exon 7 splicing is not disrupted. The inability of the 9a acceptor to define a terminal exon in conjunction with the poly(A) signal despite their proximity, suggests an intrinsic incompatibility between these elements.
signals also participate in RNA splicing, where they define the terminal exon in conjunction with an upstream 3’ acceptor element in a mechanism that is functionally and perhaps mechanistically similar to the role of the 5’ splice donor during exon definition (Niwa and Berget 1991; Wasser man and Steitz 1993; Cooke and Alwine 1996; Colgan and Manley 1997; Wahle and Kuhn 1997). Evidence for this connection is shown by the ability of the U1A component of the spliceosome to interact with both the poly(A) signal and 5’ splice donors (Lut tz et al. 1996) and in transcripts such as those from the mu heavy chain and CT/CGRP gene, where tissue-specific alternative end processing is regulated through competitive interactions between polyadenylation signals and 5’ splice donors (Seipel and Peterson 1995; Takagaki et al. 1996). Thus, poly(A) signals play an important role in both the constitutive and regulated definition of terminal exons. Given this, it is not surprising that the introduction of a poly(A) signal into the alternative exons of Mhc results in the premature truncation and polyadenylation of Mhc transcripts. However, while element-induced truncations underlie other mutations in Drosophila (Mount et al. 1988; Searles et al. 1990; Pret and Searles 1991; Ishimaru and Saigo 1993; Fridell and Searles 1994), the insertion of poly(A) signals into Mhc exons 7 and 9 affects alternative splicing and, therefore, provides insight into the normal mechanisms that regulate the alternative processing of these exons.

The truncated mutant transcripts examined in Mhc2, 3, and 4 contain terminal exons defined by the poly(A) signal and by the aberrant use of an upstream 3’ acceptor. In Mhc2 and Mhc4, the terminal exon is defined through the use of the upstream common exon acceptor, which also occurs in the 4.1-kb transcript in Mhc3. Thus, although the acceptors of alternative exons are available for use in generating the terminal exons, these are not used in favor of the upstream common exon. A simple cis-competition model explains this observation, where the nonconsensus acceptors of the exon 7 or 9 alternatives function as poor splicing substrates and, therefore, do not compete successfully with the consensus acceptor of the upstream common exon 6 or 8 when in combination with the inserted poly(A) signal. In support of this model, Niwa et al. (1992) showed that a weakened 3’ acceptor upstream of a poly(A) signal severely represses polyadenylation in a chimeric transcript, and Nesic and M aquat (1994) found that the efficiency of polyadenylation is tied to the efficiency of terminal intron removal, both demonstrating that 3’ end processing is affected by the composition of upstream splice sites. These data suggest that the consensus acceptor of the upstream common Mhc exon 6 or 8 can promote more efficient 3’ end processing than the nonconsensus acceptors of the alternative exons and will, therefore, be the favored site for terminal exon definition. While interactions between poly(A) signals and 3’ acceptors have been examined in the case of the Drosophila dsx transcript (H edley and M aniat is 1991), the observation here of competition among different 3’ acceptors for terminal exon definition has not been reported previously.

In the Mhc3 mutation, the 3’ acceptor of alternative exon 7d is also used to form a processed alternative terminal exon that generates a 1.7-kb transcript. Given that the 7d acceptor is nonconsensus and of similar composition to the other exon 7 acceptors, its ability to compete with the exon 6 acceptor for 3’ end formation suggests that a mechanism exists to enhance its activity relative to that of other exon 7 acceptors. One possibility is that this enhancement is provided by the proximity of the 7d acceptor to the insertion, which might improve the ability of the splicing apparatus to define this terminal exon. Although little data are available on the effects of distance between competing 3’ acceptors and poly(A) signals on 3’ end formation, the fact that no other truncated forms are observed would indicate that other alternative exon 7’s are too distant from the inserted poly(A) signal to compete effectively with the exon 6 acceptor. A similar model also holds for Mhc2 3’ end formation, in which the distance of all alternative exon acceptors to the poly(A) signal makes them unable to compete. The nonconsensus exon 9a acceptor site is relatively close (437 bp) to the insertion and might, according to this model, also be expected to be enhanced for defining the terminal exon in conjunction with the poly(A) signal of the B104 LTR. However, the data presented here show that the acceptor of exon 8 is exclusively used in Mhc4, indicating that the proximity of the 9a acceptor to the inserted poly(A) signal does not promote its use in terminal exon formation. Interestingly, in the Mhc11 mutant (K ronen et al. 1991), a single nucleotide change in the donor of exon 9a activates a downstream cryptic donor splice site in the 9c intron. This, however, is not used in conjunction with the 9a acceptor and rather defines an exon in combination with the acceptor of exon 8, indicating that there is a strict requirement for the correct 9a donor for the 9a acceptor to be functional. Thus, we suggest that, in the Mhc4 insertion, a heterologous “donor” site is introduced that is incompatible with the 9a acceptor, forcing the use of the upstream exon 8 acceptor.

Role of splice junction elements in exon 7d alternative splicing: Interestingly, we observed that the use of exon 7d to define the terminal exon in Mhc3 transcripts increases in adults compared to larvae, which is consistent with the fact that exon 7d is the major exon 7 alternative in adult transcripts, but is essentially absent in larval transcripts. This suggests that the increase in the level of exon 7d selection in adults is influenced by the alternative splicing mechanism that normally directs the use of exon 7d, and, further, that the cis-regulatory elements required for exon 7d use are intact and functional in this mutant. As in other regulated exons, the
nonconsensus 3' acceptor and 5' donor are potential sites for this regulation. However, recognition of the 5' splice donor would be expected to suppress the activity of the polyadenylation signal and instead promote the read-through and the subsequent removal of the element via the splicing of exon 7d to the downstream exon 8, which is similar to the events that occur in the regulation of the mu gene transcript (Takahashi et al. 1996). Evidence in the Mhc3 mutant of read-through and processing-out of the transposon insert was observed by Mogami et al. (1986), who noted that the mutant heterozygotes express a higher level of MHC in adult muscle than is found in Mhc haplotypes. Thus, the enhancement of the 1.7-kb Mhc3 transcript abundance in adults is not likely to occur through a mechanism that promotes 5' splice donor activity.

Regulated exons often contain 3' acceptors that have poor intrinsic splicing activity in nonpermissive tissues or in the absence of enhancing factors or cis-elements (Hodges and Bernstein 1994). This feature is demonstrated in the 3' acceptor of exon 7d by its weak ability to compete with the 3' acceptor of the upstream common exon 6 for terminal exon definition in larval muscle. Thus, it is perhaps the 3' acceptor of exon 7d that is the site of enhancement in adult muscle and, therefore, a potentially important element for Mhc exon 7 alternative splicing regulation. In support of this is the fact that all exon 7 alternatives contain evolutionarily conserved nonconsensus 3' acceptors (Miedema et al. 1994), while the exon 6 acceptor is consensus. It is notable, however, that the 3' acceptors of common exon 6 and all exon 7 alternatives lack significant polypyrimidine tracks and are essentially equivalent in the number of pyrimidine nucleotides that they contain. The most striking difference is that all alternative exon 7's contain either a nonconsensus T or A at the -3 position, while exon 6 contains a consensus C nucleotide at this position. Acceptor strength for these Mhc exons might then be determined in a large part through the composition of the nucleotide at this single position, a situation that has not been noted in other regulated exons (Lin and Patton 1995).

**Conclusion:** The alternative processing of the Mhc gene transcript requires precise, muscle-specific selection and use of alternatives from within five exon groups. Disruptions in this process caused by transposon-inserted polynucleotides have been examined here and reveal the role of 3' splice acceptors in the normal alternative splicing regulation of Mhc alternative exons 7 and 9, in contrast to the splicing regulation seen in other Mhc alternative exons (St and for d et al. 1997). These differences suggest the possibility of distinct alternative splicing mechanisms for each Mhc alternative exon group, indicating that a number of pathways are important to the splicing regulation of the complex Mhc gene transcript.

We thank James Alwine (University of Pennsylvania) and Brian Brunk (University of Pennsylvania) for their constructive comments on this manuscript. We also thank the following colleagues for technical assistance, and all members of the Emerson laboratory for their interest and discussion throughout the course of this project. This work was initiated at the University of Virginia and was supported by a US Army Breast Cancer Initiative Postdoctoral Fellowship (to D.M.S.), a Muscular Dystrophy Association Postdoctoral Fellowship (to M.B.D.), an American Cancer Society grant NP-841 (to C.P.E.), and a National Institutes of Health grant R01-AR42363 (to C.P.E.).

**LITERATURE CITED**


Communicating editor: L. L. Searles