AD NUMBER

ADB225042

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies only; Proprietary Info.; Oct 96. Other requests shall be referred to Commander, U.S. Army Medical Research and Materiel Command, Attn: MCMR-RMI-S, Fort Detrick, Frederick, MD 21702-5012.

AUTHORITY

P. M. Rinehart, Deputy Chief of Staff for Info. Mgmt., USAMRMC, MCMR-RMI-S [70-1y], Ft. Detrick, MD.

AD_____

GRANT NUMBER DAMD17-94-J-4146

TITLE: In Vivo Analysis of Alternative Pre-mRNA Splicing

PRINCIPAL INVESTIGATOR: David M. Standiford, Ph.D.

CONTRACTING ORGANIZATION: Fox Chase Cancer Center
Philadelphia, PA 19111

REPORT DATE: October 1996

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1996	Final (15 Sep 94 - 14 Sep 96)

4. TITLE AND SUBTITLE In Vivo Analysis of Alternative Pre-mRNA Splicing

5. FUNDING NUMBERS
DAMD17-94-J-4146

6. AUTHOR(S)
David M. Standiford, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Fox Chase Cancer Center
Philadelphia, PA 19111

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick
Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution authorized to U.S. Government agencies only; Proprietary Information, Oct 96. Other requests for this document shall be referred to Comander, U.S. Army Medical Research and Materiel Command, ATTN: MCMR-RMI-S, Fort Detrick, Frederick, MD 21702-5012

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

Abstract. Alternative splicing of pre-mRNA transcripts is normally an important mechanism to regulate the expression of many different protein isoforms including transcription factors and other regulatory proteins. Mis-regulation of alternative splicing can lead to the expression of inappropriate forms of such proteins in cells and has been shown to play a role in the establishing the cancerous state, enhancing the virulence of certain cancers, or inhibiting the treatments of cancer. The mechanisms that regulate alternative splicing have been investigated in Drosophila using the alternatively spliced *myosin heavy chain* gene as a model. The unique molecular and genetic tools provided by this system, have allowed the characterization of the elements that participate in permitting alternative splicing in general and have generated a model to explain the overall regulation of alternative splicing in this gene. The findings here are likely applicable to the understanding of other complex alternatively spliced genes.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	45
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Limited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

DMS
____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

DMS
____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

PI - Signature

10/11/96
Date

D.M. Standiford

In Vivo Analysis of Alternative Pre-mRNA Splicing

Introduction

Alternative pre-messenger RNA splicing provides an important mechanism to increase the number of protein isoforms available from an existing number of genes (reviewed in Green 1991; Maniatis 1991; Nadal-Ginard et al. 1991; McKeown 1992; Black, 1995). In addition, alternative splicing can direct the expression of various regulatory proteins, thereby influencing important genetic pathways such as the development of the sexual phenotype in Drosophila (Baker 1989). Because alternative pre-mRNA splicing plays such a central role in the expression of cellular proteins, including those that effect cell cycle progression (Betticher, et al., 1995; Meredith et al, 1995), its disruption can have serious detrimental consequences for the organism. Aberrations of the alternative splicing process have been shown to be associated with many known disease states, including several types of cancer (Cooper, 1995; Oyama et al. 1989, 1990; Pfeffer et al. 1993). Included in this list is human breast cancer, which currently is one of the most common, yet least well understood, cancer (Gould 1993). For example, the estrogen receptor protein has been shown to be aberrantly spliced in a percentage of breast neoplasms, giving rise to cells that altered sensitivities or have become insensitive to hormone (McGuire et al. 1991; Wang and Miksicek 1991; Koehorst et al. 1993; Pfeffer et al. 1993B; reviewed in Petrangeli, et al., 1995). This is of particular consequence since hormone therapy is an important treatment for breast cancer. Also important in the development of many cancers (Akiyama and Yamada 1993; Glukhova and Theiry 1993; Humphries 1993), and likely to be in breast cancer as well (Zajchowski et al. 1991; Christensen 1992; Glukhova and Theiry 1993), is the expression of inappropriate isoforms of the fibronectin (FN) protein. Fibronectin is an important cell adhesion molecule and plays a role in specifying cell identity. Many different isoforms are produced in different tissues and many of these are generated through alternative splicing. The alteration of the normal splicing pattern of the FN pre-mRNA in cancer cells appears to be an important requisite for the later metastatic development of the cancer (Oyama et al. 1989, 1990). These examples indicate that the de-regulation of alternative mRNA splicing is an important feature in the development of many cancers. It appears evident, therefore, that in order to understand and provide better therapies for cancer, a fuller comprehension of the mechanism that regulate alternative splicing are required.

Despite its importance, the understanding of the mechanisms used to correctly regulate alternative splicing have been slowed by the general complexities of the alternative splicing process and the lack of appropriate *in vivo* models systems in which to study it. Different proteins isoforms

generated through alternative splicing are often expressed in different tissues making it a difficult process to investigate *in vivo*. Consequently, the study of alternative splicing has centered around the use of various cell extracts or cultured cell lines that treat alternative splice choices differentially. Because of these limitations, most analyses have generally focused on model transcripts with limited complexity and few alternative splice choices. In spite of these impediments, however, there has been progress made towards an understanding of alternative splicing regulation.

There are a number of different patterns in which alternative splicing occurs (see McKeown 1992 for review). The simplest situation is exon skipping, where a single alternative exon is flanked by two constitutive exons. In this case the alternative exon is skipped in non-permissive cells allowing the two flanking exons to be spliced to each other. A well studied example of this is found in the fibronectin (FN) gene pre-mRNA, which has a number of alternatively spliced exons (Huh and Hynes 1993). One such exon, EIIIB is included in embryonic FN mRNAs, but excluded in adult tissue transcripts such as from the liver. It was found in this case that the exclusion of EIIIB in non-permissive cells is largely due to sub-optimal 5' and 3' splice sites in the EIIIB exon (Huh and Hynes 1993). Improving either allowed for the inclusion of EIIIB in formerly non-permissive cells. The inclusion of EIIIB in normally permissive cells was found to require a balance between the strengths of the EIIIB exon's splice sites and that of the two flanking exons as well as an intronic sequence 3' to EIIIB. This 3' sequence of 122 nucleotides was found to be necessary for any EIIIB splicing in permissive cells and was hence termed the Intronic Control Region (ICR).

A second alternative splicing pattern is mutually exclusive splicing. Here, two or more alternative exons are arranged consecutively in the pre-mRNA, but only one is found spliced into the mature message with the choice being determined in a tissue or developmentally specified fashion. In this situation, there are questions concerning not only how which exon is specified for inclusion into the message, but also how consecutive alternative exons are prevented from splicing to each other. An example of this alternative splicing pattern is found in the a-tropomyosin gene. In a-tropomyosin pre-mRNA, there are several sets of alternatively spliced exons. In one set, alternative exons 2 and 3 mutually exclude each other from the transcript with exon 3 utilized in all cells except smooth muscle. It was found that the selection of exon 3 over exon 2 in most cells was due to a more favorable polypyrimidine tract at the exon 3' splice site, thus allowing it to out-compete exon 2 for incorporation into the message. This choice appears to be negatively regulated in smooth muscle, however, by the presence of additional *cis*-elements in exon 3 (Yeakley et al.

1993). In other alternatively spliced pre-mRNAs, exon selection can be determined by splice site competition for splicing factors (Libri et al. 1992), through RNA secondary structures that suppress splice choices in some tissues (Libri et al. 1991; Gontarek et al. 1993) or through the activities of certain *trans*-acting factors (Mayeda et al. 1993; Valca'rcel et al. 1993; Zahler et al. 1993; reviewed in Green 1991; Mattox et al. 1992).

The mechanisms that prevent consecutive alternative exons from splicing to each other also appear to be varied. For example, in a-tropomyosin, an investigation of the mutually exclusive nature of exons 2 and 3, (Smith and Nadal-Ginard 1989) revealed that the branch point sequence used by exon 3 was only 42 nucleotides down stream from the exon 2 splice donor site, thus sterically blocking the splicing of exon 2 and 3. This mechanism appears not to be a general one, however, since Graham,et al (Yeakley et al. 1993) found that specific exon sequences enforced the mutual exclusion of the two forms of exon 5 from a-tropomyosin.

More recent work done on various in vitro model systems has revealed a perhaps broader mechanism for the regulation of splice choice specificity through the use splice enhance elements (Casari et al. 1994; Humphrey et al. 1995). This cis acting elements have been described in both invertebrate and vertebrate systems and consist of short tracts of repeated sequence that is generally purine rich (Watakabe et al. 1993). These elements are most often found internal to the exon they regulate and can confer splice choice enhancement to the heterologous exons providing that additional *cis*-acting splice elements such as the donor are suitably modified.

While in these cases and in others, particular *cis* and *trans*-acting factors appear to function in the regulation of alternative splicing, but there is still little information on how the overall regulation of alternative splicing is achieved. In addition, given the diversity and complexity of alternative splicing it seems unlikely that current methods of analyses will provide a complete picture of how alternative splicing is regulated. Rather, what is required is a combination of the current molecular approach to the study of alternative splicing with a genetic analysis so that the factors regulating alternative splicing can not only be identified, but biologically characterized as well.

The potential success of this approach has been demonstrated in the study of the sex determination pathway in Drosophila (Baker 1989). In this case, genetic methods have identified important regulatory genes in sex determination, as well as demonstrating how these genes function to regulate the pathway. Interestingly, several of these regulatory genes are alternatively spliced in a male or female specific patterns to give products that function to maintain the correct sexual phenotype. Much has been learned about the control of alternative splicing through the

study of sex determination in Drosophila (Burtis and Baker 1989; Hoshijima et al. 1991; Ryner and Baker 1991; Valca'rcel et al. 1993), but because of the complexity of the pathway and the sterility of splicing mutants, the genetic approach here is limited.

There is, however, an additional example of alternative splicing that provides the opportunity to explore the mechanisms of alternative splicing *in vivo* using both molecular and genetic methods and that is found in the Drosophila muscle myosin heavy chain gene (*Mhc*). Alternative splicing of the Drosophila *Mhc* transcript provides all muscle MHC isoforms found in the fly (at least 15) and is theoretically capable of producing 480 Mhc isoforms (George et al. 1989). The 21 kilobase muscle *Mhc* gene is composed of 19 exons with most of these coding for protein regions common to all MHC isoforms (Figure 1). Five exons, however, exist as alternatively splicing groups made up of 2-5 related members. Although only one exon from each group is included in the mature *Mhc* mRNA, the combinatorial use of exons from these five sets generates MHC isoforms that have variant functional properties likely to be important in the contraction of specific muscles.

Alternative splicing of exon 11 in the thoracic muscle of the fly has been shown to be tightly regulated, with individual exons expressed only in specific muscle groups (Hastings and Emerson, 1991). A major question, therefore, is what are the elements that enforce the muscle-type specific splicing of exon 11. Based on the role of the 5'splice donor in regulating alternative splicing in other systems (Nasim et al. 1990; Talerico and Berget 1990; Kuo et al. 1991; Hodges and Bernstein 1992; Wyatt et al. 1992), it is one hypothesis that the 5 different, non-consensus exon 11 splice donors confer muscle type specificity for inclusion into the mature transcript, perhaps through interactions with muscle-specific trans-acting factors, such as a U1 snRNP. Further, it is hypothesized that intronic domains may function generally to facilitate the splicing of exon 11 in a manner similar to that seen for the ICR described in the case of exon EIIIB splicing in the fibronectin pre-mRNA. Alternatively, specificity may reside in the exons themselves. Also possible is a situation were multiple factors participate to regulate alternative splicing. Experiments below have attempt to distinguish among these possibilities.

Results

Conservation of the Mhc exon 11 domain

As a first step toward identifying *cis*-regulatory elements , the Mhc exons 10-12 of Drosophila virilis, which is 60 million years divergent from D. melanogaster, was sequenced and compared to that of D. melanogaster in an effort to identify conserved and therefore, likely to be important, sequences. This comparison showed that the overall structure of the exon 10-12 domain is identical between the two species and that protein coding regions are highly conserved (Fig. 2). In addition, this analysis also revealed a number of sequence domains outside of the protein coding regions that are also highly conserved. Of these, perhaps most interesting are the conserved 5' splice donors for each of the five exon 11s and a large block of conserved sequence between the terminal exon 11 and exon 12. *In situ* hybridizations of D. virilis thoracic sections using exon 11 specific probes from D. melanogaster demonstrated that exon usage was also the same in both melanogaster and virilis, suggesting that the mechanisms regulating splicing are also conserved. .

Splice Regulatory Elements are Local to Exon 11

A strategy to analyze cis-acting elements important in regulating exon 11 alternative splicing *in vivo* is based on the use of *Mhc* minigenes introduced transgenically into the fly. Such a minigene construct (gD1048; Fig 3A) containing the wild type exon 11 domain in conjunction with the reporter gene, ß-galactosidase (ß-gal), has been used to transform flies. This construct contains the entire exon 11 domain and is flanked upstream by the constitutive exons 1, 2 and 10 and downstream by exons 12 and 13. Exon 13 coding sequence has been fused in-frame with that of the ß-galactosidase protein and thus will provide a colorimetric method to assay the pattern of splicing and expression of the transgene. An important feature of this construct is however the fact that between exon 10 and any exon 11, there is a split codon such that in order for there to be in frame b-galactosidase message, only one exon 11 can be included in the message. Thus, the presence of ß-galactosidase activity indicates that the mechanisms preventing constitutive exons 10 and 12 from splicing to each other or exon 11's from splicing to each other are still intact.

In the flies transformed with the gD1048 construct, ß-gal is detected in all muscles, showing that splicing is occurs normally with respect to only one exon from exon 11 being included in the final transcript (Fig. 3B). Thus, the information to enforce the utilization of only a single exon 11

resides within the exon 11 domain contained in the gD1048 construct. This was further demonstrated by the gD1043 construct in which one nucleotide was added to exon 11e and a nucleotide was deleted from exon 12, such that only exon 11e containing processed messages are in-frame. Flies transformed with this construct (Figure 3C), had staining only in the IFM, showing that exon 11e use is restricted to this muscle.

To assess whether the correct exons are being utilized from the gD1048 transgene, a reverse-transcriptase- polymerase chain reaction (RT-PCR) assay has been developed (Fig. 4). RT-PCR is a powerful technique that allows for the detection of specific RNAs from very small amounts of tissues and the assay developed for use in this study has been designed to detect messages arising specifically from either the endogenous *Mhc* gene or from the *Mhc* mini-gene. Further, a method has been developed to allow the efficient dissection of specific muscles from the adult thorax, from which RNA can be prepared and utilized in the RT-PCR assay. This has allowed the indirect flight muscle (IFM) and the tergal depressor of the trochanter (TDT) both to be directly examined and splicing products and from the endogenous Mhc gene, as well as from the Mhc minigene, simultaneously detected. The use of this assay confirms that the exon 11's from the gD1048 transgene are utilized correctly in the fly and provides and mechanism to detect changes in the pattern of exon 11 splicing as a result of mutations made to the gD1048 construct.

Non-consensus 5' splice donors are important for alternative splicing

As described above, non-consensus 5' splice donors have been shown to be a part of the apparatus involved in regulating alternative splicing. Given this and their evolutionarily conserved sequence, the non-consensus exon 11 5' splice donors of Mhc exon 11 were examined for their role in Mhc exon 11 slice regulation. One possible role of the 5' splice donors is to provide an exon specific signal that would function only in the proper muscle, perhaps through the presence of muscle specific alternative splicing factors. This hypothesis is supported by the fact that the conserved donor sequence elements for each of the different exon 11's are unique. If splice choice is being regulated through the donor elements, then switching donors between exons should result in the switching of exon usage as well. This was tested in the gD1105 construct where the donor of exon 11e has been replace with that of exon 11b and vice versa (Fig. 5A). This construct was tested in transgenic flies and the first analyses revealed that the thoracic muscles of this fly where normal for the expression of ß-gal, indicating that overall splicing was not disrupted. When the RT-PCR assay was done, it was found that the expression of both exon 11e and 11b was still restricted to their normal muscles (11e -IFM; 11b- TDT). These data indicated that the donors

themselves are not involved in specifying exon choice. Further, the fact that there was no apparent disruption of splicing in general shows that there is no requirement for the particular exons to be associated with particular donors in spite of the conserved nature of the donors. This suggests that the non-consensus 5' donors act as simply as weakened splicing substrates, which is often the case in alternatively spliced exons. Regulation might occur, then, through the interaction of these weakened donors with additional exon specific *cis* elements to provide a competitive advantage to the correct exon in the appropriate muscle again through the interactions of muscle type specific trans-factors.

To test the role of competition among exons in splice regulation constructs were made that replaced the weak, wild type donor of either exon 11e or 11b with a strong consensus donor. The first of these constructs, gD1090, replaces the exon 11e donor with a consensus donor (Fig. 5B). When tested in transgenic flies, this construct was first shown to direct normal general splicing as indicated by the presence of ß-gal in the muscle of the thorax. When the IFM and TDT were assayed using RT-PCR, it was found that exon 11e was being utilized in the TDT and to the apparent complete exclusion of exon 11b. The replacement of the exon 11e donor with a strong consensus donor, in combination with the consensus 3' splice acceptor, should result in the creation of a essentially constitutive exon. This appears to be the case in terms of the exon 11e usage in all muscle. Unexpected, however, was the observation that the transformation of exon 11e to a constitutive exon did not activate the splicing of this exon to downstream alternatives. This indicates that the addition of a consensus donor to 11e, while eliminating the specificity of the splicing reaction, does not over-ride the mechanisms that ensure only single alternative exons are included in the processed message.

These conclusions were further tested with the gD1177 construct, in which the donor of exon 11b was replaced with a consensus donor (Fig. 5C). The positive results of ß-galactosidase staining in all muscles indicates that splicing still occurs normally with respect to only a single exon 11 being utilized in the transcript. Analyses of the TDT and IFM with the RT-PCR assay, revealed that exon 11b is used in the IFM indicating that splice specificity is eliminated in the gD1177 construct. As with the gD1090 construct, the inclusion of a consensus donor in exon 11b did not disrupt the mechanisms to prevent alternatives from splicing to each other.

Interestingly, exon 11e in gD1090 and exon 11b in gD1177 were active in larvae, where neither exon is normally utilized in any muscle (George el al, 1989) and mutual exclusive splicing was maintained, suggesting that the overall mechanisms regulating exon usage and mutual exclusivity are the same in larvae and adults.

Exon position and splice specificity

As demonstrated by the gD1105 clone, the specificity of muscle specific exon choice does not reside within the 5' splice donors of the alternative exons themselves, indicating that other cis-elements are involved in this process. One such element might be the exon itself. Other example of exons containing tissue specific splice enhancers exist and in the case of the Mhc exon 11's, each has a sequence of approximately 30 nucleotides in the middle of the exon that all differ significantly from each other, which could function to direct muscle specific splicing. To test the role of the exons themselves in splice regulation, the gD1122 clone was constructed in which the positions of the two exons, 11e and 11b, were switched (Fig. 6). If the exons themselves contain the information tot direct their muscle specific inclusion, then swapping position should not affect exon utilization. When the dD1222 construct was tested transgenically, the flies were found to still express ß-galactosidase in all muscles indicating that general splicing was not disrupted. When the expression pattern of the exon 11e and exon 11b was assayed through RT-PCR, it was found that in the IFM, there was a complete switch from exon 11e to exon 11b usage. In the TDT, there was also a switch from exon 11b to exon 11e. These results indicate that the exons themselves do not direct there own splice utilization and that intronic sequences must also play a role. Further, this indicates that the position of the exons within the exon 11 domain must be important for maintaining splice choice specificity. Interestingly, these results also show that isoform expression can be altered by the switching of exons within an alternative group.

The role of intronic sequences

The role of intronic sequences was first examined by removing the intronic conserved region (ICR) in the gD1120 construct (Fig. 7). Several cases exist where a intronic domain functions to activate or enhance the splicing of alternative exons in a cell or tissue specific fashion and a similar role might be envisioned for the ICR. However, when examined in transgenic flies, the gD1120 construct was spliced correctly to produce ß-galactosidase in all muscles. Analysis of exon usage from this transgene with RT-PCR showed in addition that both exon 11e and 11b were utilized correctly in the IFM or TDT. Therefore, the definition of splice choice in exon 11 does not appear to be dependent of a single domain of intronic sequence. The role of the very well conserved ICR may lie within maintain the efficiency of alternative splicing or may be related to other events not related to RNA processing.

The influence of exon position on splice choice can be mediated through a scanning mechanism,

where the splicing apparatus surveys through the alternative domain, and chooses an exon for inclusion based on its position relative to the other exons. So for the IFM, exon 11e might be selected based on its being the first exon in the linear order. This might function through the activity of splicing factors such as ASF/SF2 (Krainer et al. 1990), which has been shown to shift the splice choice toward the upstream exon with increasing concentration. Thus, the IFM might have the highest concentration of ASF/SF2 so that the most 5' splice choice is always made in this muscle. One conserved feature of the exon 11 domain is the order and position of the alternatives within it, indicating that this is important in splice regulation. To test this, the gD1168 clone was constructed in which exon 11e was removed with enough flanking sequence to move exon 11a to a primary position essentially equivalent to that of exon 11e with respect to exon 10 (Fig. 8). If a scanning mechanism functions in defining the splice choice, then exon 11a, as the first exon, will get selected for inclusion in the IFM. Analysis of flies transformed with the gD1168 transgene, however, first revealed that there was no ß-galactosidase staining in the IFM, while other muscles stained normally. This indicated that normal splicing was not occurring in the IFM. RT-PCR analysis showed that exon 11b was still correctly utilized in the TDT and that the pattern of exon usage in the larvae was normal. These results indicate that exon scanning does not occur and that specific exons are recognized in a context that does not depend on their position relative to other exons, but rather is dependent on intronic sequence elements.

The removal of exon 11e from the gD1168 construct resulted in the loss of any exon being used in the IFM, while rest of the exons in the domain remained normal for splicing. Further, the activation of exon 11b when it is in exon 11e's position suggests that the sequences immediately flanking the exons themselves participate in splice regulation. This suggests that the intronic sequence immediately flanking exon 11e are required and perhaps sufficient for correct splicing. To test this, a segment of exon 11e approximately equal to that removed in the gD1168 construct was placed into background from which all other exons and most of the ICR had been deleted. This construct, gD1060, was found to direct the expression of ß-galactosidase exclusively in the IFM, indicating that the 360 nt of intronic sequence included in this minimal construct contains the sequences required for IFM-specific splicing of exon 11e (Figure 9). A comparison of these domain from *D. melanogaster*, *D. virilis*, and *D. hydei* (Miedema et al. 1994) showed the presence of two well conserved elements residing in the intron downstream of exon 11e (Figure 10). The first of these, Element I (EI) is repeated twice in *melanogaster* and *virilis*, but is only single copy in *hydei*. Element II (EII) is found separated from EI by ~20 nt in *melanogaster* and *virilis*, but is contiguous with EI in *hydei*..

Preliminarily data suggests that small conserved elements might function in the splice regulation of other alternatives is found in the gD1204 construct, which is similar to gD1060, except that exon 11b is inserted into the deleted background and the conserved elements downstream of exon 11b are absent (not shown). Here, no ß-galactosidase staining was detected in flies transformed with this construct, nor could the protein be detected by immunoblot assay. RT-PCR analysis revealed that spliced products included a skipped product, which consisted of exon 10 spliced directly to exon 12. Normal splice products were also detectable (i.e. exon 10-11b-12), indicating that exon 11b can still be included in the message, but the fact that ß-galactosidase protein could not be detected indicates that normal splice products are very rare and that, likely, most of the spliced products are following the skipping pathway. Thus, the result that exon 11b is aberrantly spliced in the absence of conserved intronic sequences suggests that these elements might be important generally for directing the muscle specific splicing of alternative exon 11s.

One final construct which has been examined has addressed the question of mutual exclusivity in exon choice. This construct, gD1223, contains both exon 11e and 11b in and otherwise exon-deleted background (not shown). This construct is also spliced poorly in transgenic flies and the absence of detectable ß-galactosidase and the presence of a skipped product likely indicates that most of the message is aberrantly skip-spliced. However, normally spliced exon 11e and 11b are detected in the both TDT and IFM, indicating that in this background, there is no competition between the two exons. There is also no evidence of exon to exon splicing between 11e and 11b. This indicates that the *cis*-elements that mediate this process are contained within the limited sequence domain defined by this construct.

Summary

The alternatively spliced Mhc gene exon 11 contains five member alternative exons, of which only one can be included in the final message. The mechanisms that determine the exact choice of exon to be included in that message function in a very precise muscle specific fashion, thus assuring that only the correct MHC isoform is expressed in that muscle. Prior to this study, little work had been done on complex alternative exons such as Mhc exon 11, and therefore, little was known concerning the mechanisms that regulate alternative splicing in exons with multiple members. In the work presented above, several key components important for splice regulation have been identified *in vivo*, and these are likely to have meaning and relevance towards the understanding of other complex alternatively spliced messages.

The exon definition model (Berget, 1995) holds that in a general splicing, individual exons are defined through the interactions of the splicing machinery with both the 3' splice acceptor and the 5' splice donor. This recognition is facilitated by the strength of the splice elements (i.e. similarity to consensus) and the context of the exon itself (i.e. its relationship to other exons; the size of flanking introns, etc...). Once exons themselves are defined through this process, then splicing between exons can occur. In the case of Mhc exons 10, 11 and 12, both exons 10 and 12 have consensus splice sites and should both be identified readily as exons, whereas the non-consensus 5' splice donor of each exon 11 are expected to make them generally difficult to be recognized by the splicing machinery. Therefore, regulatory elements in addition to those that conduct constitutive splicing are expected to participate in alternative splicing regulation.

The work presented above demonstrates that there are several such components that function in cis to properly regulate the alternative splicing of the Mhc exon 11 alternatives. First, although highly conserved and different from each other, the 5' splice donors do not provide muscle specificity to exon choice. However the donors were found to be essential for allowing alternatives exon choice to occur in any muscle. This is shown by the replacement of native 5' splice donors with consensus donor sequences, which results in the dominant use of that alternative exon that contains the consensus donor. Second, this work shows that a small domain of intronic sequence located within the gD1060 can properly direct the muscle specific use of a single alternative, namely exon 11e in the IFM. This is the first example of cis-regulatory sequences that can function to direct alternative splicing in a particular muscle type and suggests a model for how the overall regulation of splicing regulation of exon 11 occurs. In this model Figure 12), intronic sequence elements, perhaps the EI and EII for exon 11e, serve to attract the

splice activating factors (SAF), which are themselves important for promoting the recognition by the regular splicing apparatus of the exon 11e non-consensus donor site. The IFM SAF might interact with the regular splicing components, or might require additional intermediate factors, but importantly, they would be expressed specifically in the IFM to ensure the exclusive use of this exon in the IFM. Other muscles might then express different SAFs to direct the splicing of other alternative exons. Interestingly, each alternative exon 11 has associated with in a one or two short well conserved intronic sequence elements, suggesting that this model might be applicable to the exon 11 alternative splicing in general.

Mutual exclusivity and splice regulation

One critical aspect to splice regulation in exon 11 is its mutual exclusivity. This process assures that only a single exon from this group is included in the message, which prevents the introduction of reading-frame shifts that would render the protein useless. In the gD1090 and gD1177 clones, a consensus donor was placed into alternative exons, essentially converting them to constitutive exons apparently free to splice to any other exon, including another alternative. In these mutants, however, mutual exclusivity still functions and all splicing occurs between the modified alternative and the flanking constitutive exons 10 and 12. The elements that enforce this mutual exclusivity are perhaps defined in the gD1223 clone in which both exon 11e and 11b are together in a largely deleted background. While most of the spliced product from this clone appears to be skip-spliced, there is still detectable message that is normally spliced exon 10 to 11e and 11b. There is, however, no detectable exon 11e to 11b splicing indicating that the mechanism of mutual exclusive splicing has not been lost in this construct. Much future work will be aimed at determining the rules that govern this process.

Conclusions and Future Work

Several conclusions concerning the nature of cis-regulation of alternative splicing can be drawn from the work presented above. The first of these is that the weak 5' donors are essential for allowing alternative splicing to occur. A second feature is that while the exons themselves do not appear top participate in directing their own splicing, they do appear to rely on sequences that immediately flank the exons. In the case of exon 11e, this intronic information appears to contribute directly to the specific utilization of this alternative in the IFM. A third point is that the mechanisms that enforce mutual exclusivity are strong enough to over come strong exons created within the exon 11 domain and appear to operate independently of splice choice specificity.

Future work will be directed at a better understanding of the role the intronic sequences play in splice regulation. In particular, the role of EI and EII will be assessed through deletion and mutation analysis. Further, the role of the conserved domains associated with other alernative exon 11s will be tested using mutation, deletion and position swapping experiments.

Having identified a limited domain of intronic sequence that is able to direct the IFM specific use of exon 11e in the gD1060 construct, we are now making progress using *Drosophila* genetic screens to identify the trans-acting factors that are likely to interact with these intronic sequences. This screen employed a powerful new method for making mosaic flies (Xu and Rubin 1993), and was based on the fact that the IFM is dependent on the presence of its own specific isoform of myosin for flight. This, of course, depends on alternative splicing, and mutations that disrupt this process in the IFM should lead to flightless flies. Flies with defects in the IFM are completely viable, so mutants can be isolated and analyzed for splicing defects. The construction of this screen allows for a single chromosomal arm (one sixth of the genome) to be examined at a time and when one such arm was examined in a previous screen, some 30 recessive flightless mutants were recovered. While these are being examined for splicing mutants, a screen of a second chromosomal arm is being developed. This screen will have much more stringent criteria for the types of mutants to be examined, which will lead to a more rapid identification of flies mutant for alternative splicing.

The work performed so far has progressed to a point where a nearly complete picture of how alternative splice choices are made in different muscles during the processing of the Mhc pre-mRNA transcript can be presented. The process leading to a correctly spliced message is, like the message itself, a complex one, and consist of separate mechanisms that provide an context in which alternative splicing can occur, permit specific exons to be selected for inclusion into the processed message in a muscle specific fashion and a mechanism to prevent the inappropriate splicing of consecutive alternative exons from splicing to each other. An exact myosin heavy chain protein isoform is required in the IFM or the fly will likely be unable to fly. It is therefore critical to the animal that the correct splice choice be made in the IFM and it has evolved the mechanisms to do this. As discussed above, the deregulation of the differentiated cell pattern of splicing for many vertebrate transcripts appears to be associated with the establishment or progression of cancer. Thus, maintaining tight control of alternative exon usage is a generally shared need and as is anticipated from the further study of the process in Drosophila, will be based on shared mechanisms as well.

Bibliography

Adams MD, Rudner DZ and Rio, DC (1996) Biochemistry and regulation of pre-mRNA splicing. Curr. Op. Cell Biol. 8: 331-339.

Akiyama SK, Yamada KM (1993) Introduction: adhesion molecules in cancer. Part I. Sem Cancer Biol 4:215-218

Baker BS (1989) Sex in flies: the splice of life. Nature 340:521-523

Betticher DC, Thatcher N, Altermann HJ, Hoban P, Ryder WD, and Heighway J. (1995) Alternative splicing produces a novel cyclin D1 transcript. Oncogene 11:1005-1011.

Cooper DL (1995) Retention of CD44 introns in bladder cancer: understanding the alternative splicing of pre-mRNA opens new insights into the pathogenesis of human cancers. J. Pathology 177:1-3

Black DL (1995) Finding splice sites in a wilderness of RNA. RNA 1:763-771

Berget SM. (1995) Exon recognition in vertebrate splicing. J. Biol. Chem. 270:2411-2414.

Burtis KC, Baker BS (1989) Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. Cell 56:997-1010

Christensen L (1992) The distribution of fibronectin, laminin and tetranectin in human breast cancer with special attention to the extracellular matrix. APMIS 26(Suppl.):1-39

Deshpande, G. Samuels, ME, Schedl, PD (1996) Sex-lethal interacts with splicing factors in vitro and in vivo. Mol. Cell. Biol. 16:5036-5047.

Erlich HA (ed) (1989) PCR Technology: Principles and Application for DNA Amplification. Stockton Press, New York

George EL, Ober MB, Emerson Jr CP (1989) Functional domains of the Drosophila melanogaster muscle mysosin heavy-chain gene are encoded by alternatively spliced exons. Mol Cell Biol 9:2957-2974

Glukhova MA, Theiry J (1993) Fibronectin and integrins in development. Sem Cancer Biol 4:241-249

Gontarek RR, McNally MT, Beemon K (1993) Mutation of an RSV intronic element abolishes both U11/U12 snRNP binding and negative regualtion of splicing. Genes Dev 7:1926-1936

Gould MN (1993) Cellular and molecular aspects of the multistage progression of mammary carcinogenesis in humans and rats. Sem Cancer Biol 4:161-169

Green MR (1991) Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu Rev Cell Biol 7:559-599

Hastings GA, Emerson Jr CP (1991) Myosin functional domains encoded by alternative exons are expressed in specific thoracic muscles of Drosophila. J Cell Biol 114:263-276

Hodges D, Bernstein SI (1992) Suboptimal 5' and 3' splice sites regulate alternative splicing of Drosophila melanogaster myosin heavy chain transcripts in vitro. Mech Devel 37:127-140

Hoshijima K, Inoue K, Higuchi I, Sakamoto H, Shimura Y (1991) Control of doublesex alternative splicing by transformer and tranformer-2 in Drosophila. Science 252:833-836

Huh GS, Hynes RO (1993) Elements regulating alternative spliced exon of the rat fibronectin gene. Mol Cell Biol 13:5301-5314

Humphrey, M.B., Bryan, J., Cooper, T.A. and Berget, S.M. (1995) A 32-nucleotide exon-splicing enhancer regulated usage of competing 5' splice sites in a differential internal exon. Mol. Cell. Bio. 15:3979-3988

Humphries MJ (1993) Fibronectin and cancer: rationals for the use of antiadhesives in cancer treatment. Sem Cancer Biol 4:293-299

Krainer, A.R., Conway, G.C., Kozak. D. (1990) The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. Cell 62: 35-42

Koehorst SG, Jacobs HM, Thijssen JH, Blankenstein MA (1993) Wild type and alternatively spliced estrogen receptor messenger RNA in human meningioma tissue and MCF7 breast cancer cells. J Steroid Biochem & MolBiol 45:227-233

Kuo H, Nasim FH, Grabowski PJ (1991) Control of alternative splicing by the differential binding of U1 small nuclear ribonuclear particle. Science 251:1045-1050

Libri D, Piseri A, Fiszman MY (1991) Tissue-specific splicing in vivo of the ß-tropomyosin gene: dependence on an RNA secondary structure. Science 252:1842-1845

Libri D, Balvay L, Fiszman MY (1992) In vivo splicing of the ß-tropomyosin pre-mRNA: a role for the branch point and donor site competition. Mol Cell Biol 12:3204-3215

Maniatis T (1991) Mechanisms of alternative splicing. Science 251:33-34

Mattox W, Ryner L, Baker BS (1992) Autoregulation and multifunctionality among trans-acting factors that regulate alternative pre-mRNA processing. J Biol Chem 267:19023-19026

Mayeda A, Helfman DM, Krainer AR (1993) Modulation of exon skipping and inclusion by heterogenous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. Mol Cell Biol 13:2993-3001

McGuire WL, Chamness GC, Fuqua SA (1991) Estrogen receptor variants in clinical breast cancer. Molecular Endocrinology 5:1571-1577

McKeown M (1992) Alternative mRNA splicing. Annu Rev Cell Biol 8:133-155

Meridith Jr., J., Takada, Y., Fornaro, M., Languino, L.R., and Schwartz, M.A. (1995) Inhibition of cell cycle progression by the alternatively spliced integrin b1C. Science 269:1570-1572

Miedema, K., H. Harhangi, S. Mentzel, M. Wilbrink, A. Akhmanova, M. Hooiveld, P. Bindels, and W. Hennig. (1994). Interspecific sequence comparison of the muscle-myosin heavy-chain genes from Drosophila hydei and Drosophila melanogaster. *Journal of Molecular Evolution.* 39:357-368.

Nadal-Ginard B, Smith CWJ, Patton JG, Breitbart RE (1991) Alternative splicing is an effecient mechanism for the generation of protein diversity: contractile protein genes as a model system. Adv Enz Regul 31:261-286

Nasim FH, Spears PA, Hoffman HM, Kuo H, Grabowski PJ (1990) A sequential splicing mechanism promotes selection of an optional exon by repositioning a downstream 5' splice site in preprotachykinin pre-mRNA. Genes Dev 4:1172-1184

Oyama F, Hirohashi S, Shimosato Y, Titani K, Sekiguchi K (1989) Deregulation of alternative splicing of fibronectin pre-mRNA in malignant human liver tumors. J Biol Chem 264:10331-10334

Oyama F, Hirohashi S, Shimosato Y, Titani K, Sekiguchi K (1990) Oncodevelopmental regulation of the alternative splicing of fibronectin pre-messenger RNA in human lung tissue. Cancer Res 50:1075-1078

Petrangeli E, Lubrano C, Ortolani F, Ravenna L, Vacca A, Sciacchitano S, Frati L and Gulino A (1995) Estrogen receptors: new perspectives in breast cancer management.
J. Ster. Bioch. Mol. Biol. 49:327-331.

Pfeffer U, Fecarotta D, Castagnetta L, Vidali G (1993) Estrogen receptor variant RNA lacking exon 4 in estrogen-responsive human breast cancer lines. Cancer Res 53:741-743

Ryan KJ and Cooper TA (1996) Muscle specific splicing enhancers regualte inclusion of the cardiac toponin T alternative exon in embryonic skeletal muscle. Mol. Cell. Biol. 16:4014-4023.

Ryner LC, Baker BS (1991) Regulation of doublesex pre-mRNA processing occurs by 3'-splice site activation. Genes Dev 5:2071-2085

Smith CWJ, Nadal-Ginard B (1989) Mutually exclusive splicing of a-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. Cell 56:749-758

Talerico M, Berget SM (1990) Effect of 5'splice site mutations on splicing of the preceding intron. Mol Cell Biol 10:6299-6305

Valcàrcel J, Singh R, Zamore PD, Green MR (1993) The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. Science 362:171-175

Wang Y, Miksicek RJ (1991) Identification of a dominant negative form of the human estrogen receptor. Molecular Endocrinology 5:1707-1715

Watakabe, A., Tanaka, K, and Shimura Y. (1993) The role of exon sequences in splice site selection. Genes Dev. 7:407-418

Yeakley JM, Hedrjran F, Morfin JP, Merillat N, Rosenfeld MG, Emeson RB (1993) Control of calcitonin/calcitonin gene related peptide pre-mRNA processing by constitutive intron and exon elements. Mol Cell Biol 13:5999-6011

Zajchowski DA, Band V, Trask DK, Kling D, Connolly JL, Sager R (1991) Suppression of tumor-forming ability and related traits in MCF-7 human breast cancer cells by fusion with immortal mammary epithelial cells. Proc Nat Acad Science USA 87:2314-2318

Appendix I

Myosin Rod Protein

In the course of the work described within the body of this document, a new contractile protein was discovered in Drosophila. This protein, the Myosin Rod Protein (MRP) is a novel contractile protein expressed from a gene that is internal to the myosin heavy chain gene (see Figure 1). The gD1048 clone actually contains the promoter for Mrp and, although beginning as troublesome background expression from the Mhc minigene mutant constructs, this fact has allowed for the initial characterization of the MRP protein. The MRP is unique in that it is contains the rod domain of myosin, but not the motor head domain. So while the MRP can form filaments similar to muscle thick filaments, it is contractually inert. This feature leads directly to some unprecedented structural and perhaps physiological changes in the muscles that express MRP. Interestingly, despite its unusual and unexpected properties, the MRP is expressed in a variety of muscles, including the heart. A better understanding of this protein might then provide further information about the structure and function of this organ, as well as muscles in general. Further, the MRP is expressed in a non-muscle tissue, this being the testis, where it appears to be involved in important sperm cell mutation events that potentially involve cell shape changes and cytoplasmic movement. Thus, this interesting protein appears to have a multiple roles in the fly, both in the function of muscle and in differentiation of germ cells in Drosophila.

Bibliography

Standiford, D.M., Davis, M.E., Sun, W., and Emerson, C.P. Jr. (1997) Splice-junction elements and intronic sequences regulate alternative splicing of the Drosophila myosin heavy chain gene transcript. **Genetics**: In Press.

Standiford, D.M., Davis, M.B., Miedema, K., Franzini-Armstrong, C. and Emerson, C.P. (1997) Myosin Rod Protein: a novel thick filament component of Drosophila muscle. **J. Mol. Biol**. 265:40-55

Standiford, DM, Davis, MB, Sun, W. and Emerson CP. Analysis of intronic sequences in the regulation of myosin heavy chain alternative splicing in *Drosophila*.. (In Progress)

Davis, MB, Standiford, DM, Emerson CP. Analysis of *trans*-acting factors regulating alternative splicing of the myosin heavy chain gene in *Drosophila*. (In Progress)

Tung, W, Standiford, D, Davis, MB and Emerson CP. Analysis of myosin heavy chain gene structure in arthropods. (In Progress)

FIGURE LEGENDS

FIGURE 1. Diagram of the 36B myosin heavy chain (*Mhc*) gene in *Drosophila melanogaster* showing the distribution of the 5 alternatively spliced exon groups (exons 3, 7, 9, 11 & 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.

FIGURE 2. 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 homology, in the same order as the *melanogaster* alternatives. Exon size is conserved as is the approximate distance between each alternative. *(B)* Exon 11 alternative exon splice donors are conserved and all are non-consensus (stars indicate non-consensus nucleotides). *(C)* A large intronic conserved region (ICR) between exon 11d and 12 is present and represents the largest site of intronic conservation. Several domains of purine rich sequence are underlined. *(D)* The regulation of the alternative exons in *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*.

FIGURE 3. *(A)*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. *(B)* 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 ß-gal. *(C)* The gD1043 construct contains an insertion in exon 11e and a deletion in exon 12 and expresses ß-gal in the IFM but not the TDT.

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. Larger products corresponding to unprocessed message or DNA priming are occasionally seen (11b minigene lane). 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.

FIGURE 5. Constructs designed to test the function of the exon 11 non-consensus 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 non-consensus 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 utilized 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 non-consensus 5' splice donor of 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 6. Constructs designed to test the function exonic sequences in alternative splicing regulation. Minigene exons 11e and 11b were positionally exchanged in the gD1222 construct and this was tested by RT-PCR analysis for its affects 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.

FIGURE 7. Constructs designed to test the function of the ICR. 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.

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 ß-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-ß-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 immuno-detected in each sample as a loading control.

FIGURE 10. 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 *melanogaster* and *virilis*, but is contiguous with E-I in *D. hydei*.

FIGURE 11. Model of IFM specific splicing of exon 11e. *(A)* Intronic specificity elements (possibly E-I and E-II) are specifically recognized by splice activation factors (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).

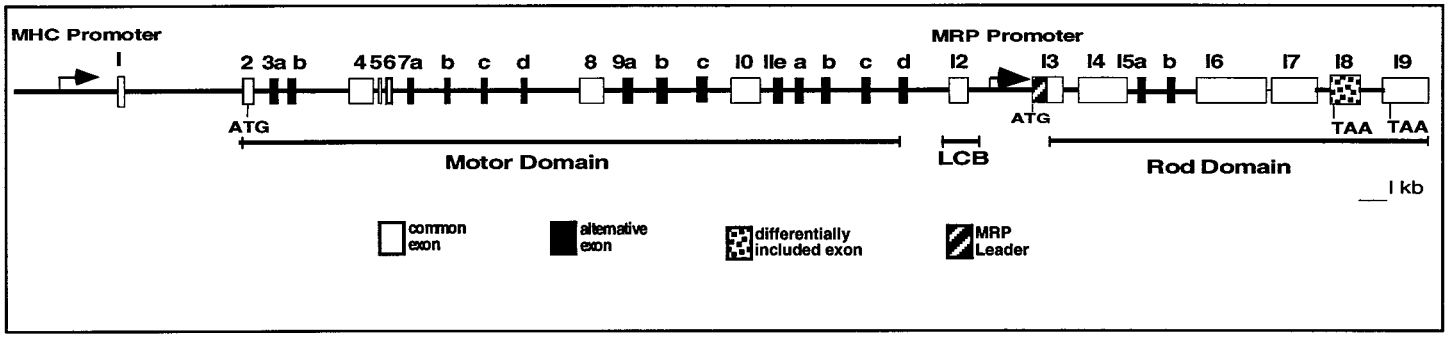
MHC Promoter
MRP Promoter
1
2 3a b
4 5 6 7a b c d
8 9a b c
10 11e a b c d
12
13
14 15a b
16
17
18
19
ATG
ATG
TAA
TAA
Motor Domain
LCB
Rod Domain
1 kb
common exon
alternative exon
differentially included exon
MRP Leader

Figure 1

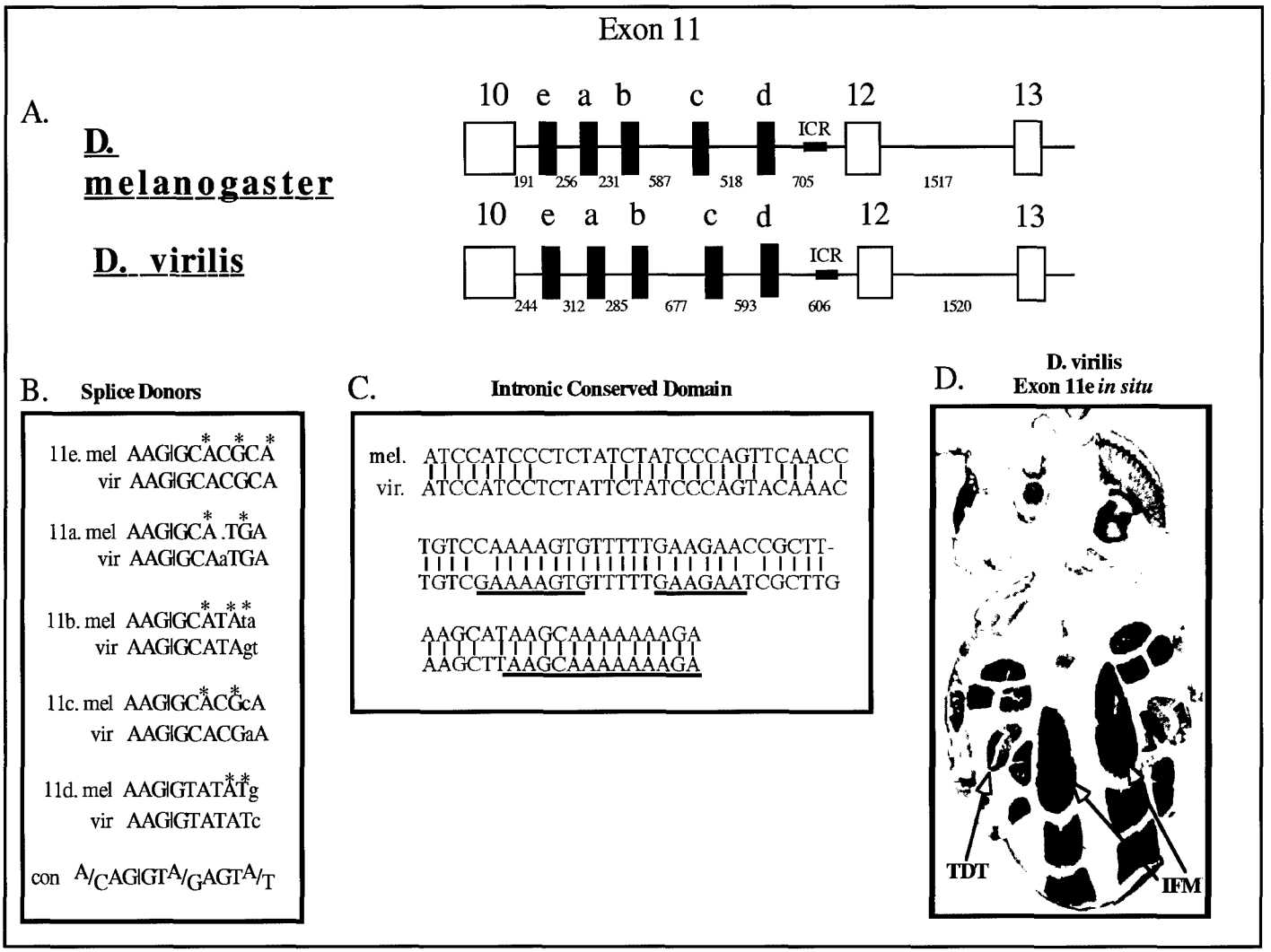
Exon 11
A.
D. melanogaster
D. virilis
10 e a b c d 12 13
ICR
191 256 231 587 518 705 1517
10 e a b c d 12 13
ICR
244 312 285 677 593 606 1520
B. Splice Donors
11e. mel AAGIGCACGCA
vir AAGIGCACGCA
11a. mel AAGIGCA.TGA
vir AAGIGCAaTGA
11b. mel AAGIGCATAta
vir AAGIGCATAgt
11c. mel AAGIGCACGcA
vir AAGIGCACGaA
11d. mel AAGIGTATATg
vir AAGIGTATATc
con A/CAGIGTA/GAGTA/T
C. Intronic Conserved Domain
mel. ATCCATCCCTCTATCTATCCCAGTTCAACC
vir. ATCCATCCTCTATTCTATCCCAGTACAAAC
TGTCCAAAAGTGTTTTTGAAGAACCGCTT-
TGTCGAAAAGTGTTTTTGAAGAATCGCTTG
AAGCATAAGCAAAAAAAGA
AAGCTTAAGCAAAAAAAGA
D. virilis
Exon 11e in situ
TDT
IFM

Figure 2

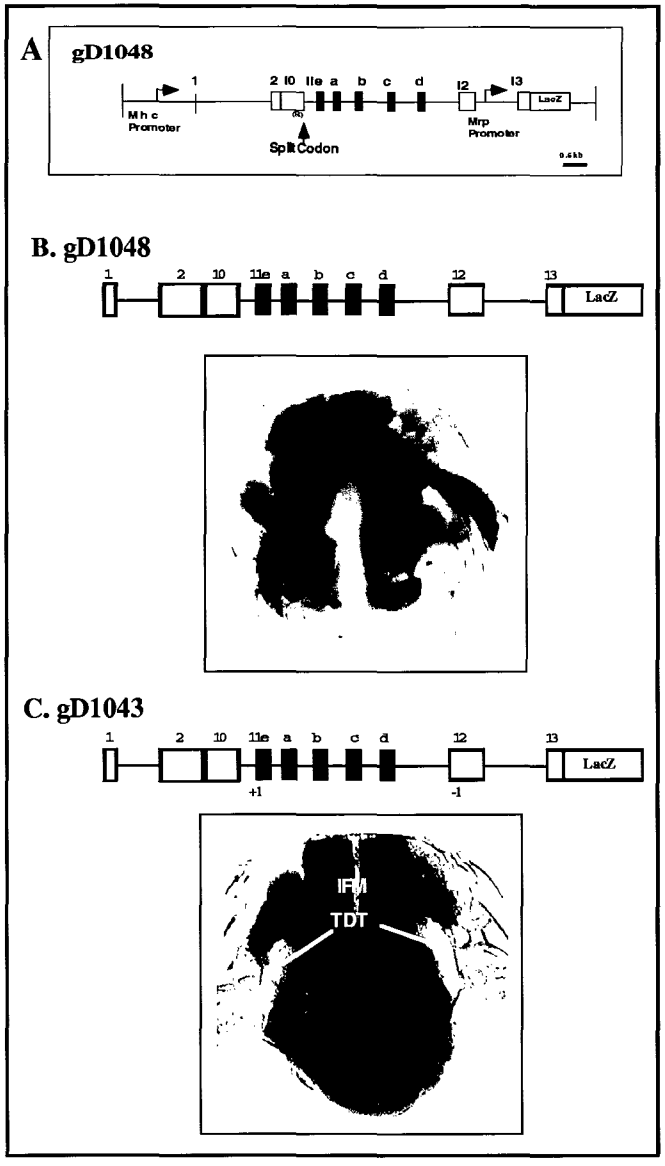

Figure 3

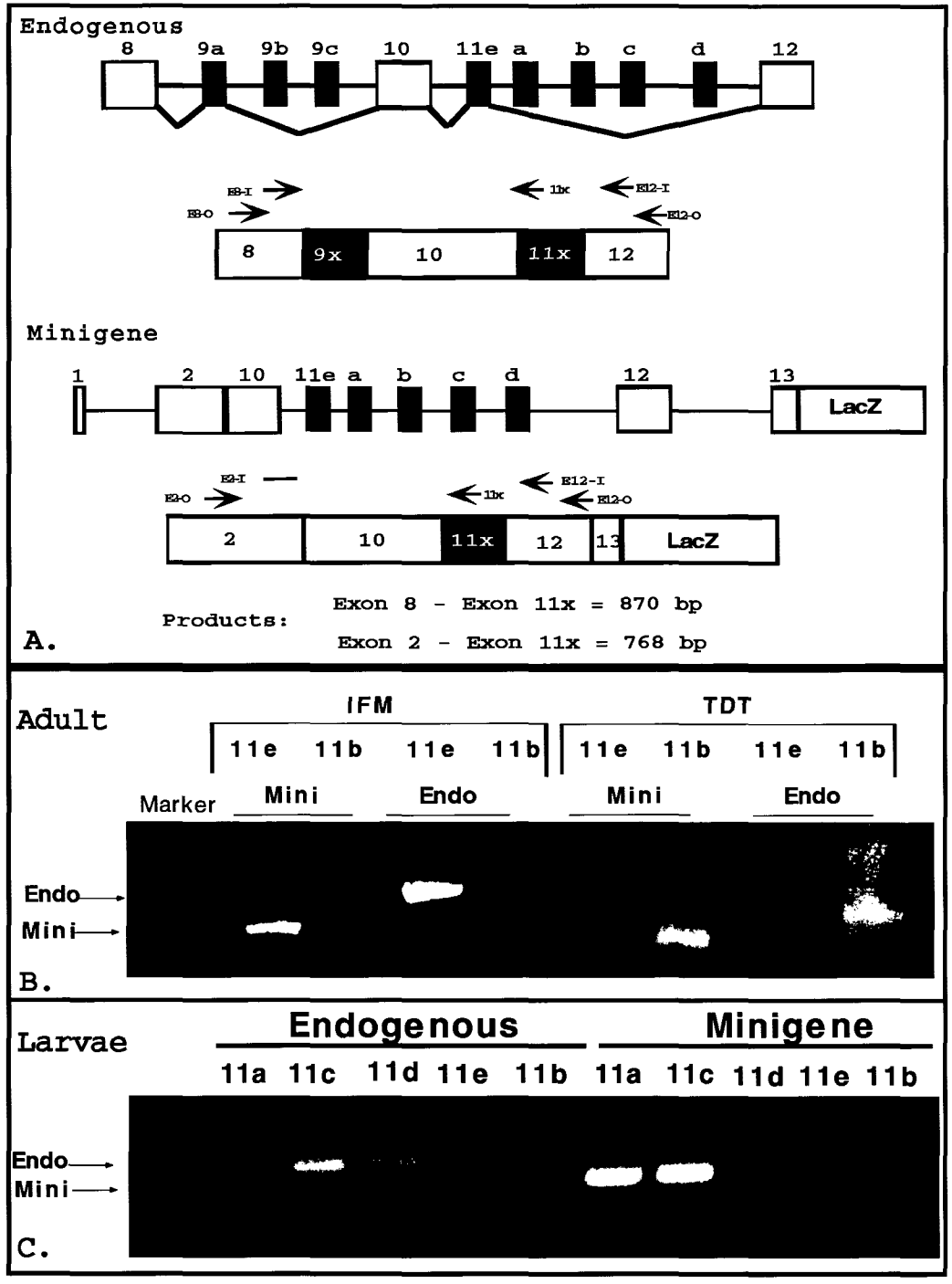

Figure 4

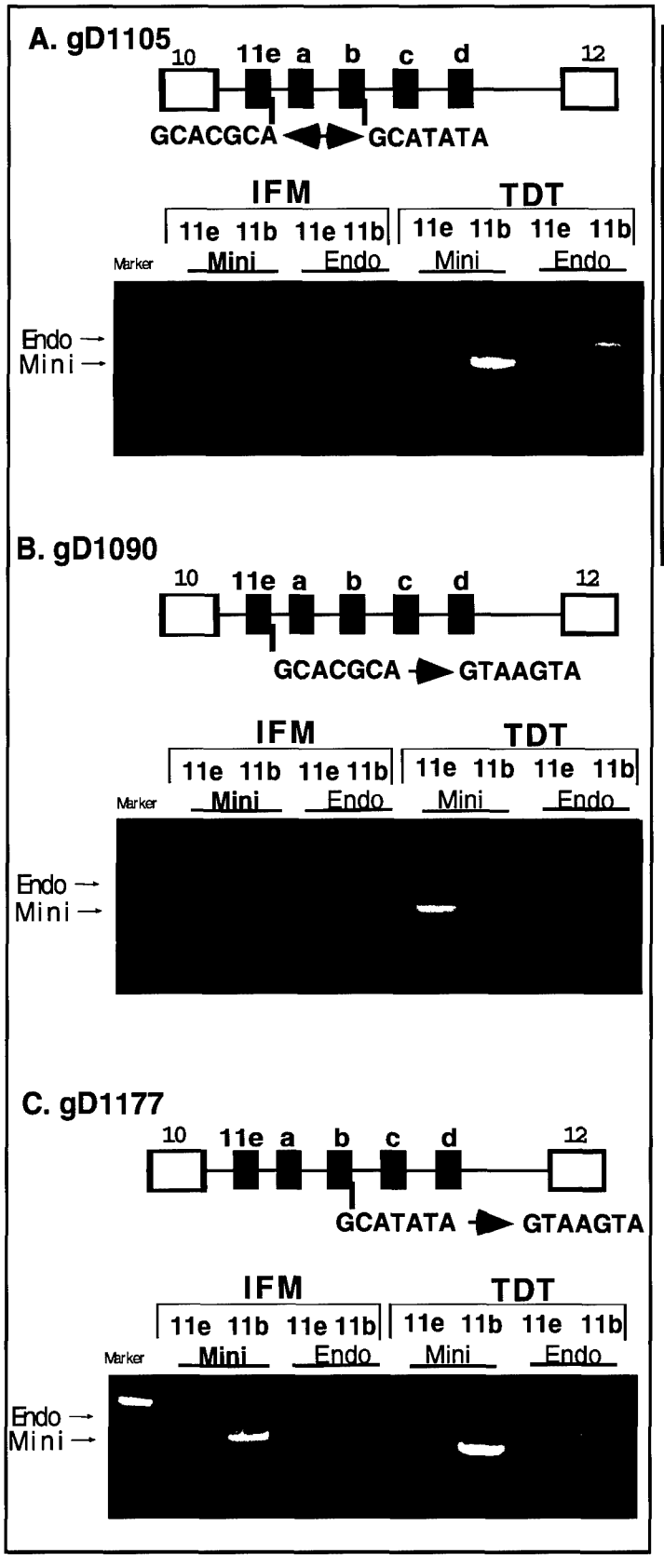

Figure 5

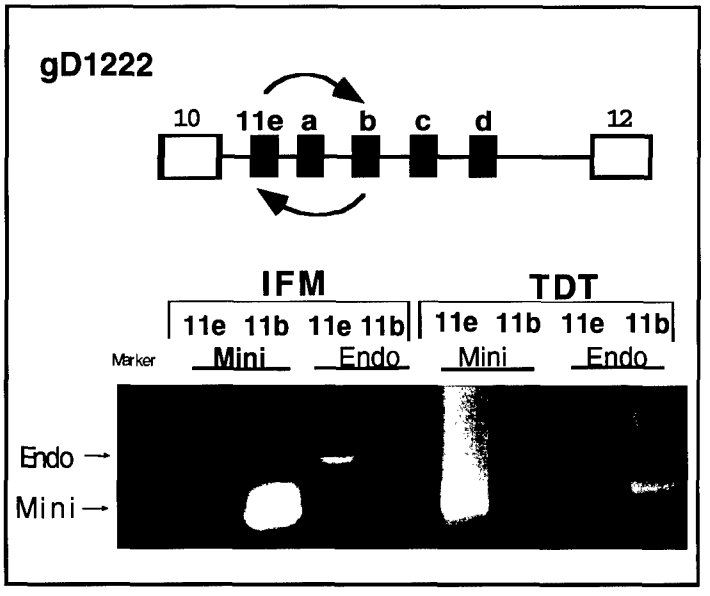

Figure 6

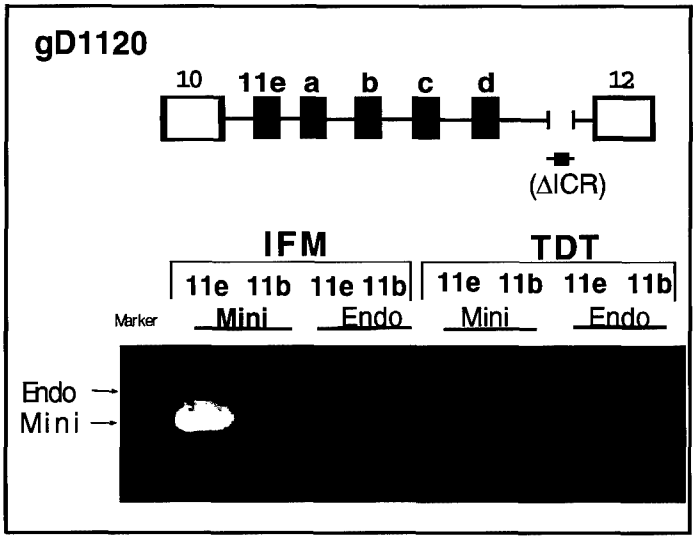

Figure 7

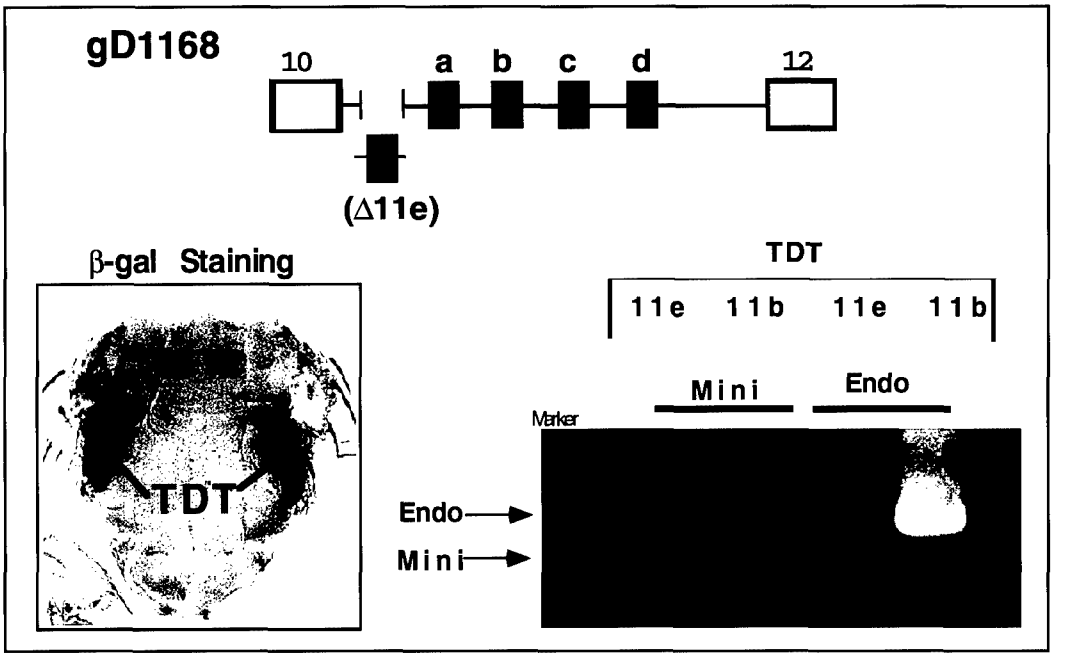

Figure 8

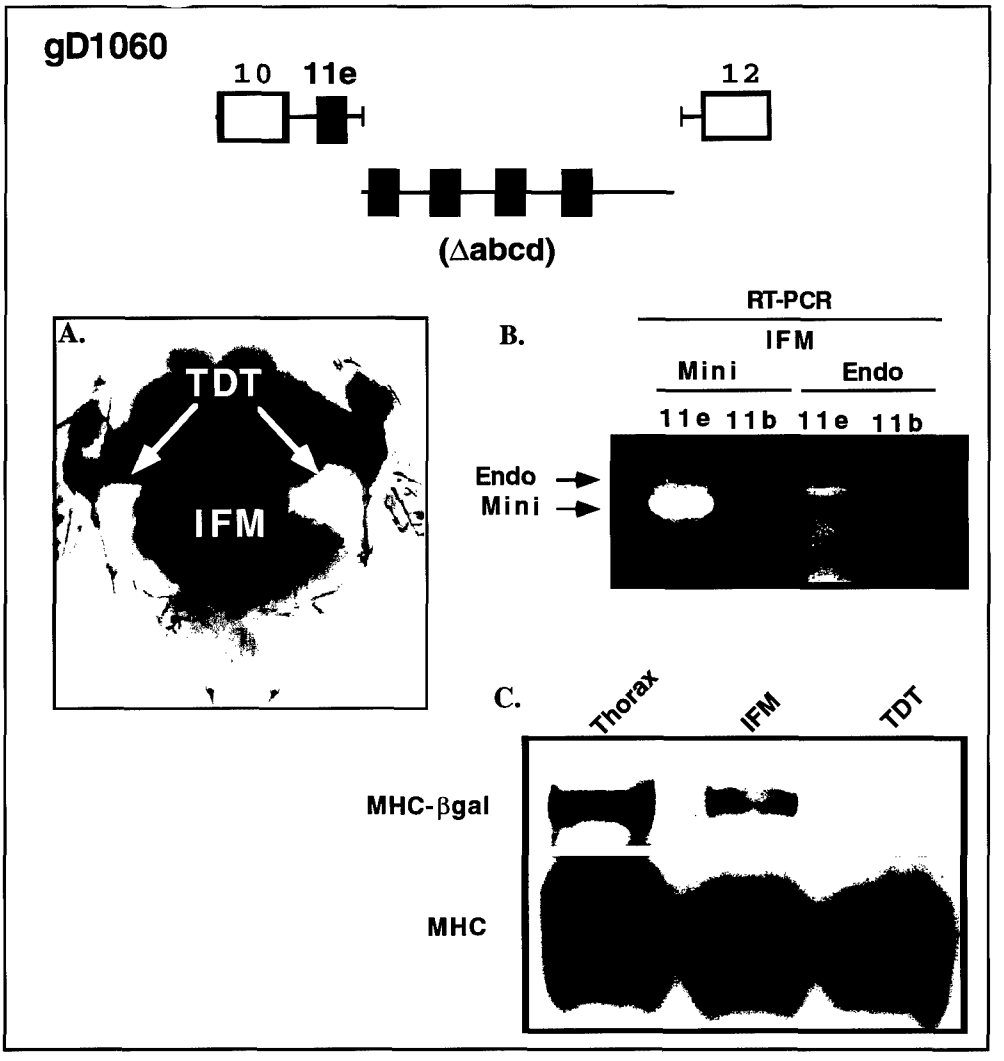

Figure 9

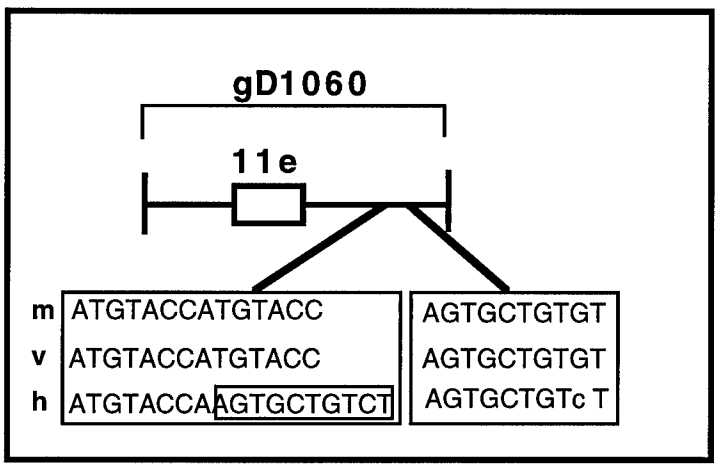
gD1060
11e
m ATGTACCATGTACC AGTGCTGTGT
v ATGTACCATGTACC AGTGCTGTGT
h ATGTACCAAGTGCTGTCT AGTGCTGTc T

Figure 10

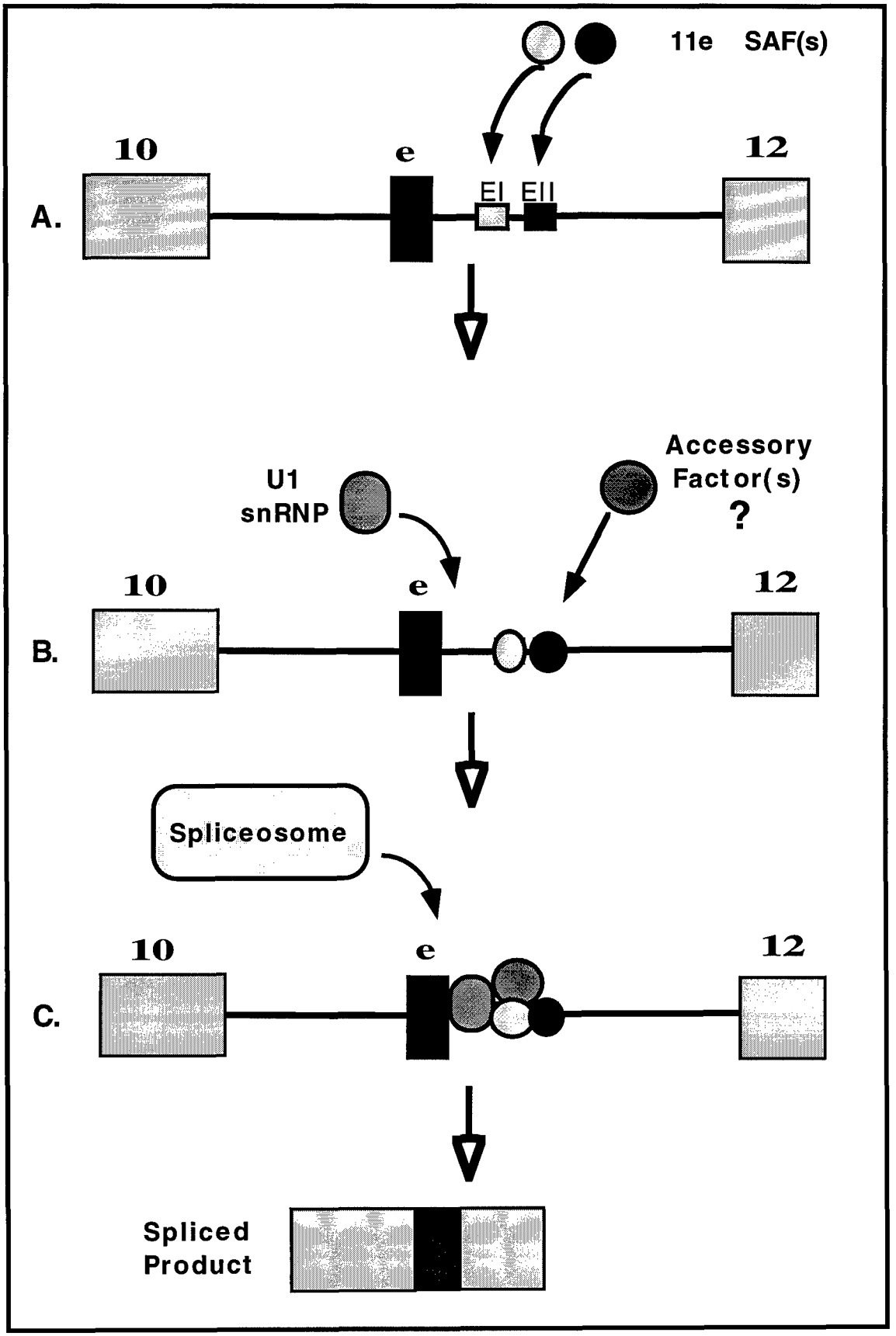
11e SAF(s)
10
e
12
EI EII
A.
U1
snRNP
Accessory
Factor(s)
?
10
e
12
B.
Spliceosome
10
e
12
C.
Spliced
Product

Figure 11