Nucleic Acids Research

Branch point selection in alternative splicing of tropomyosin pre-mRNAs

David M.Helfman and William M.Ricci

Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724, USA

ABSTRACT

The rat tropomyosin 1 gene gives rise to two mRNAs encoding rat fibroblast TM-1 and skeletal muscle beta-tropomyosin via an alternative splicing mechanism. The gene is comprised of 11 exons. Exons 1 through 5 and exons 8 and 9 are common to all mRNAs expressed from this gene. Exons 6 and 11 are used in fibroblasts as well as smooth muscle whereas exons 7 and 10 are used exclusively in skeletal muscle. In the present studies we have focused on the mutually exclusive internal alternative splice choice involving exon 6 (fibroblast-type splice) and exon 7 (skeletal muscle-type splice). To study the mechanism and regulation of alternative splice site selection we have characterized the branch points used in processing of the tropomyosin pre-mRNAs *in vitro* using nuclear extracts obtained from HeLa cells. Splicing of exon 5 to exon 6 (fibroblast-type splice) involves the use of three branch points located 25, 29, and 36 nucleotides upstream of the 3' splice site of exon 6. Splicing of exon 6 (fibroblast-type splice) or exon 7 (skeletal muscle type-splice) to exon 8 involves the use of the same branch point located 24 nucleotides upstream of this shared 3' splice site. In contrast, the splicing of exon 5 to exon 7 (skeletal muscle-type splice) involves the use of three branch sites located 144, 147 and 153 nucleotides, upstream of the 3' splice site of exon 7. In addition, the pyrimidine content of the region between these unusual branch points and the 3' splice site of exon 7 was found to be greater than 80%. These studies raise the possibility that the use of branch points located a long distance from a 3' splice site may be an essential feature of some alternatively spliced exons. The possible significance of these unusual branch points as well as a role for the polypyrimidine stretch in intron 6 in splice site selection are discussed.

INTRODUCTION

The formation of a mature mRNA from a primary transcript generally requires excision of intervening sequences (introns) as well as cleavage and polyadenylation at the 3' end. Considerable progress has been made in recent years concerning the mechanisms and biochemistry of pre-mRNA processing (for reviews see references 1−3). The development of systems that accurately splice exogenous pre-mRNA *in vitro* has permitted a biochemical analysis of the splicing reaction (4−6). The overall splicing reaction between two exons can be described as occurring in two stages (6−9). In the first stage, a 5' exon is cleaved and two intermediates are formed: a 5' exon with a 3' hydroxyl end and an RNA species containing the intron sequence and 3' exon in which the 5' terminal guanosine nucleotide of the intron is covalently linked to a residue, usually an adenosine, located 18−40 nucleotides upstream of the 3' splice site via a 2'-5' phosphodiester bond. In the second stage, cleavage at the 3' splice site results in the release of a lariat intron and concomitantly the two exons are ligated together.

The identification of sequences in the pre-mRNA required for accurate and efficient splicing has provided important information about the mechanism of the splicing reaction.

The 5' splice site consensus sequence AG/GTPuAGT, has been shown to interact with the 5' end of the U1 small nuclear RNA (10−14). The 3' splice site consensus sequence $(Py)_n NPyAG/G$ (10,11) interacts with several cellular factors, whose precise nature has not been completely established. One of these, however, is a A 70−100 Kd protein, which is probably associated with U5 snRNP (15,16). Analysis of the branch points of a number of cellular and viral genes identified a loosely conserved sequence PyNPyTPuAPy, with the adenosine residue serving as the site of branch formation (17−19). U2 snRNP is thought to interact with the pre-mRNA at the branch point region, perhaps by direct base pairing with the pre-mRNA (13, 20−22). A factor termed U2AF (U2 snRNP auxiliary factor) has been reported to be necessary for the binding of U2 snRNP to the branch point and splicing complex assembly (23).

While minimal sequence requirements have been established for splicing in simple transcription units, it is unlikely that these sequences alone account for the choice between alternative splice sites in complex transcription units. How alternative splice sites are selected is not known. Sequence comparisons between splice junctions of alternative and constitutive exons have failed to identify any significant differences (24). This suggests that the regulation of splice site selection in transcripts containing alternative 5' or 3' splice sites involves other cis-acting elements. Studies of alternatively spliced genes have identified a number of features in the pre-mRNA which may be involved in alternative splice site selection. These include intron size (25), the relative strengths of 5' splice sites (26), the pyrimidine content of a 3' splice site (27), exon sequences (28−31), and multiple alternative branch points (32−34). In addition, exon and intron sequences may be involved in the formation of RNA secondary structures that could play a role in alternative RNA processing by regulating the accessibility of different exons to the splicing machinery (35−41). A number of studies have also suggested the existence of cell type-specific factors that interact with sequence elements in the pre-mRNA to promote differential splice site selection (40,42). However, the identity of these putative trans-acting factors and the sequences they recognize to regulate alternative splicing remain largely unknown.

To determine if the splicing pathways of alternatively spliced exons exhibit mechanistic differences compared to simple transcription units a number of studies have characterized the intermediates involved in alternative splice site selection (32,34,43). Studies of alternative splicing involving SV40 and adenovirus RNAs have demonstrated that multiple branch points are associated with the processing of these pre-mRNAs (32,34). The expression SV40 large T and small t RNAs involves the joining of alternative 5' splice sites to a single shared 3' splice site. Splicing of large T RNA involved the use of one of six different branch sites located 18−32 nucleotides from the 3' splice site (32). On the other hand splicing of small t RNA involved the use of a single branch nucleotide located 18 nucleotides upstream of the 3' splice junction (32). The adenovirus E1A primary transcript is differentially spliced to yield five mRNAs that have been designated 13S, 12S, 11S, 10S, and 9S mRNAs (44,45). Splicing of E1A pre-mRNA to generate 13S, 12S and 9S involves the joining of alternative 5' splice sites to a shared 3' splice site. In contrast to studies of large T splicing, formation of E1A 13S, 12S and 9S mRNAs use the same branch adenosine residue located 28 nucleotides upstream of the common 3' splice site (43). By contrast, generation of the E1A 11S and 10S mRNAs, which uses the same 5' splice site as the 9S RNA but are joined to a different 3' splice site, uses one of three branch points located an unusually long distance 51, 55 or 59 nucleotides upstream of the 3' splice site (34). These studies of viral pre-mRNAs raise the possibility that utilization of multiple
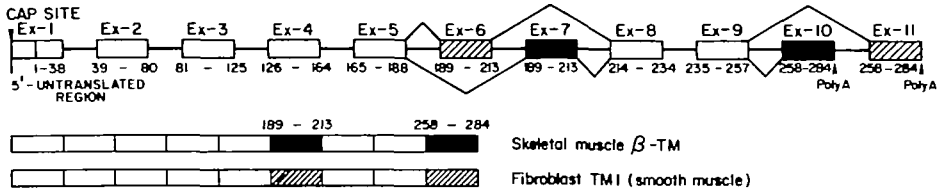
Figure 1. Schematic diagram of the rat tropomyosin 1 gene. Open boxes represent common exons, hatched and closed boxes represent tissue-specific exons as indicated. Horizontal lines represent introns; they are not drawn to scale. The amino acids encoded by each exon are indicated. The CAP site and polyadenylation sites are also indicated.

branch points may be a common feature of alternatively spliced exons. However, little information is currently available concerning branch point selection in alternative splicing of cellular genes.

We have been using the rat tropomyosin 1 gene as a model to investigate the mechanism of a type of developmental and tissue-specific alternative splicing (31,46). This gene is comprised of 11 exons (Figure 1). Exons 1 through 5 and 8 and 9 are common to all mRNAs expressed from this gene. Exons 6 and 11 are used in fibroblasts as well as smooth muscle whereas exons 7 and 10 are used exclusively in skeletal muscle. Our previous studies of tropomyosin pre-mRNA splicing *in vitro* revealed an ordered pathway of splicing in which either internal alternatively spliced exon must first be joined to the downstream common exon before they can be spliced to the upstream common exon (31). To determine if splicing of these alternative exons involves the use of multiple branch sites we analyzed the lariat intron products produced by processing of tropomyosin pre-mRNAs *in vitro*. We find that splicing of the alternative exons 6 and 7 to the downstream common exon is associated with the use the same lariat branch point. By contrast, alternative splicing of exon 5 to the downstream alternative exons 6 and 7 is associated with the use of different multiple branch points. Interestingly, the branch points used in splicing exon 5 to exon 7 (skeletal muscle-type splice) are located at an unusually long distance (144, 147, and 153 nucleotides) from the 3′ splice site of exon 7. The potential significance of these differences in branch site selection associated with utilization of these alternatively spliced exons is discussed.

MATERIALS AND METHODS
Plasmid constructions
The DNA templates for use in *in vitro* transcription with SP6 polymerase are shown in Figure 2. All are derived from the rat fibroblast TM-1 gene (46). pSP64-p2 was constructed by deletion of sequences from the unique ClaI site in exon 9 with Bal 31 to 17 nucleotides 3′ of exon 8. The DNA was then digested at a unique BalI site in exon 5. The DNA was modified with HindIII linkers and cloned into the HindIII site of SP64. Thus SP64-p2 contains an SP6 promoter with 14 nucleotides of plasmid sequence including the HindIII linker following the SP6 RNA initiation site and 66 nucleotides of exon 5 and genomic sequence extending 17 nucleotides 3′ of exon 8. Plasmid pSP64-p4 was derived from SP64-p2 by ligating the NcoI/SalI fragment of the cDNA clone encoding rat fibroblast TM-1 (47) to the NcoI/SalI sites of pSP64-p2. Plasmid pSP64-p2(7/8) was derived from pSP64-2 by removal of the intervening sequence between exons 7 and 8 using a 30
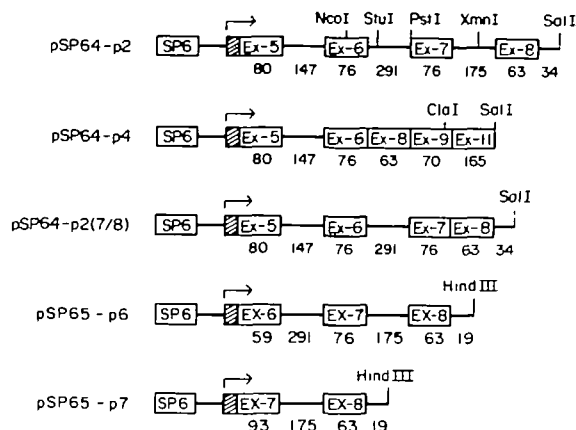
Figure 2. SP6/Tropomyosin transcriptional templates. The DNA templates for use in *in vitro* transcription with SP6 polymerase are shown. All are derived from the internal region of the TM-1 gene (exons 5 through 8). The arrows indicate the sites of transcription initiation with the SP6 sequences (cross-hatched regions) adjacent to the SP6 promoter. *In vitro* transcription terminates at indicated restriction sites. In addition, a number of internal restriction sites are also indicated that can be used for synthesis of truncated substrates. The boxes indicate positions of exon 5 (constitutive), exon 6 (fibroblast and smooth muscle-specific), exon 7 (skeletal muscle-specific) and exon 8 (constitutive). Lines represent each intron. The number of nucleotides in each exon and intron are indicated.

nucleotide long deoxyoligonucleotide containing 15 nucleotides of exon 7 and 15 nucleotides of exon 8 (48). Plasmid pSP64-p2p6 was constructed by joining the regions of the cDNA encoding fibroblast TM-1 corresponding to exons 9 through 11 to pSP64-p2. Plasmid pSP65-p6 was derived from pSP64-p2 by modifying the NcoI site in exon 6 of pSP64-p2 with EcoRI linkers, and the resulting EcoRI-HindIII fragment, containing sequences from the middle of exon 6 through sequences downstream of exon 8, was ligated into EcoRI/HindIII double-cut pSP65 plasmid DNA. Plasmid pSP65-p7 was derived from pSP65-p6 by modifying the PstI site upstream of exon 7 with EcoRI linkers and removing the upstream EcoRI/PstI fragment.

Synthesis of RNA and in vitro splicing

The ^{32}P-labelled SP6/tropomyosin transcripts were synthesized *in vitro* and primed with CAP analog as described (49). The ^{32}P-labelled pre-mRNAs contained a small amount of prematurely terminated transcripts, but their presence did not appear to affect the *in vitro* splicing reactions and therefore the RNA substrates were not further purified.

HeLa cell nuclear extracts were prepared as described (6,50). *In vitro* splicing reactions were carried out at 30° for the indicated times (0−240 min). Standard assay conditions used for these studies consisted of 15 μl of nuclear extract in a final reaction volume of 25 μl containing 1 mM MgCl$_2$, 500 μM ATP, 20 mM creatine phosphate, 2.7% (w/v) polyvinyl alcohol, 15−30 ng pre-mRNA, 12.8 mM Hepes (pH 8), 14% (v/v) glycerol, 60 mM KCL, 0.12 mM EDTA and 0.7 mM DTT.

Analysis of processing products

After incubation for the indicated times the reactions were stopped by addition of a solution containing SDS-proteinase K and the RNA recovered as described (6). The products of the reaction were analyzed on denaturing urea-polyacrylamide gels. Intermediates were distinguished from final products by their relative appearance and disappearance in time

course experiments. Lariat RNAs were identified by their aberrant mobility on different percentage polyacrylamide gels (7,8). In addition, RNAs containing lariats were verified with treatment by a lariat debranching activity (51). Processing products were also analyzed by primer-extension analysis. Primer extension analysis was carried out essentially as described (52,53). Oligodeoxynucleotides complementary to sequences encoding amino acids 198−204 within exon 6 (5'-GAGAGCCTGGTCCATGGTTC-3'), amino acids 198−204 within exon 7 (5'- CAAGTTGTTGGTAACAATTT-3'), amino acids 223−229 within exon 8 (5'-TCCAGAAGTTTGATCTCTTC-3'), the 3' end of intron 6 (5'-CTGCAGGGGCGTGTG-3') and at position −94 through −110 nucleotides upstream of the 3' splice site of exon 7 (5'-GAGGGGTGGCAGAGTGG-3') of the rat fibroblast TM-1 gene were used as primers. Primers were end labelled with (gamma-^{32}P)ATP using T4 polynucleotide kinase. The primers were annealed to the RNA by incubation for 3 minutes at 55°C, then 30 minutes at 37°C in 30 microliters of primer extension buffer (50 mM Tris-Cl pH 8.3, 140 mM KCl, 10 mM MgCl$_2$). For primer extension, reactions were transferred to 42°C and supplemented with DTT (final concentration of 5 mM) and deoxynucleotide triphosphates (final concentration of 0.5 mM) and 10 U of avian myeloblastosis virus reverse transcriptase. After 30 min, the reactions were terminated by addition of EDTA to a final concentration of 25 mM, and the radiolabelled pre-mRNA was degraded by treatment with RNAse A prior to gel electrophoresis. In order to sequence the primer extension products the standard splicing reactions were scaled up 10−30 fold, the primer-extension products were resolved on sequencing gels and the bands cut out of gels, recovered and subjected to sequence analysis (54). In some experiments DNA sequence was determined by the dideoxy termination procedure (55).

RESULTS

Tropomyosin gene constructs used for generation of pre-mRNAs for in vitro splicing
We prepared a number of DNA templates designed to analyze the internal alternatively spliced region of the rat tropomyosin 1 gene (Figure 2). The constructions used in the present studies contain 1, 2 or 3 introns of the region between exons 5 through exon 8, inserted downstream of the SP6 promoter.

Splicing of exon five to exon six (fibroblast-type splice)
To analyze the branch point used during splicing of exon 5 to exon 6 (fibroblast-type splice), we used a pre-mRNA derived from pSP64-p4 linearized with ClaI (Figure 2). We previously demonstrated that exon 5 is spliced accurately and efficiently to exon 6 using this precursor (31). The products of the *in vitro* splicing reactions with pre-mRNA transcripts from pSP64-p4 linearized with ClaI are shown in Figure 3. In this experiment RNA transcripts were incubated in *in vitro* splicing reactions for 30 min, 45 min, 60 min, and 120 min. Following the splicing reactions the RNA was divided and a portion was fractionated on a 4% polyacrylamide-urea gel (Figure 3a). To identify the branched nucleotides, a portion of the RNA was analyzed by primer extension. Reverse transcription has been shown to stop at branch nucleotides (19). Half of the RNA (debranched Figure 3b) was incubated in a HeLa S-100 fraction that contains a lariat debranching activity (51) prior to primer extension analysis. The other half was analyzed directly (branched, Figure 3b). Primer extension was performed using a 20 nucleotide long ^{32}P-labelled deoxyoligonucleotide complementary to sequences within exon 6. cDNA products were analyzed on an 8% polyacrylamide-8M urea gel along with dideoxy sequence generated from the appropriate single stranded M13 template (Figure 3b). As indicated in Figure 3b, a 129 nucleotide
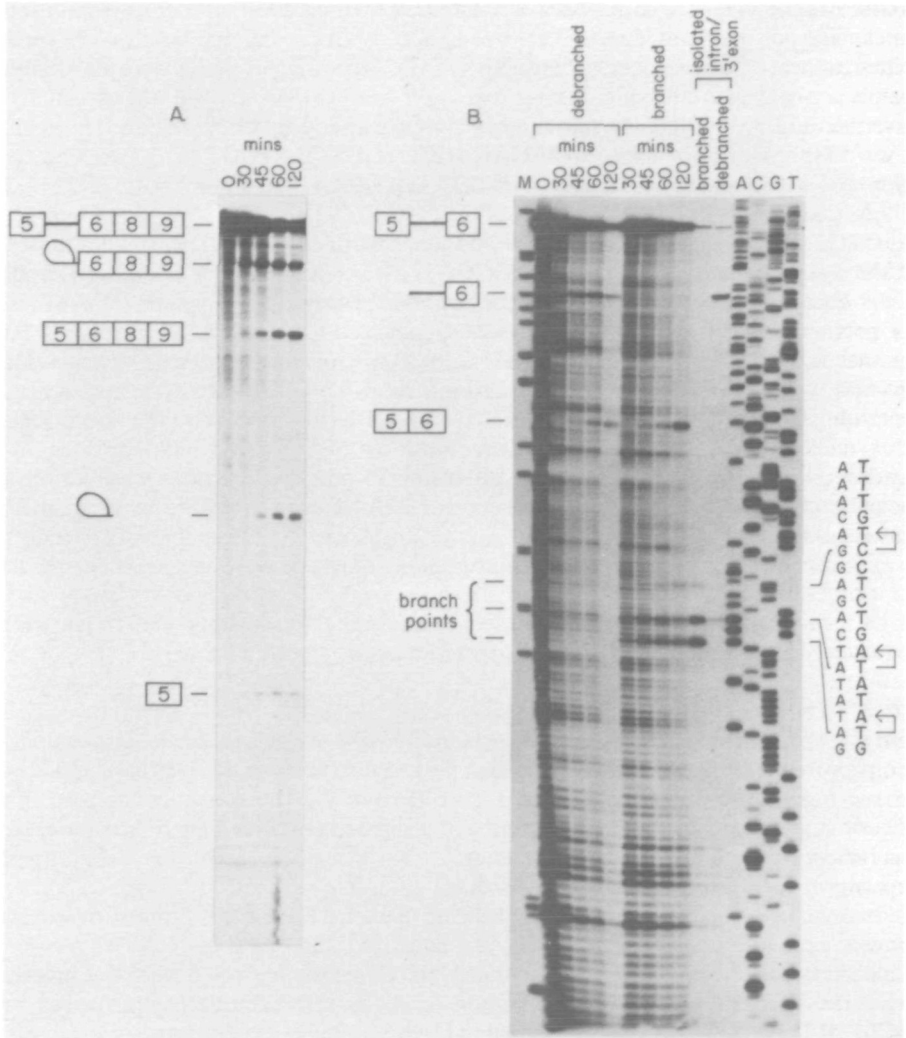
Figure 3. *In vitro* splicing of exon five to exon six (fibroblast-type splice). (Panel A) Time course (min) of formation of spliced products from transcripts terminating at the ClaI site of plasmid pSP64-p4. The ^{32}P-RNA products were separated in 4% denaturing polyacrylamide gels. Schematic representation of products are indicated; the boxes represent exon sequences and the lines intron sequences. (Panel B) The *in vitro* processing products shown in panel A were analyzed by primer-extension using a ^{32}P-labelled oligodeoxynucleotide primer specific for sequences in exon 8. Part of the processed RNA was treated with debranching extract prior to reverse transcription (lanes debranched). (panel B) The lariat-exon intermediate indicated in panel A was isolated and subjected to primer extension analysis. Position of primer extension products corresponding to splice products and intermediates are indicated. The primer extension products were separated on a 8% denaturing polyacrylamide gel and electrophoresed in parallel with dideoxy sequencing reactions performed using single stranded M13 templates primed with the same ^{32}P-labelled oligonucleotide primer to the right. The sequences corresponding to positions 23 through 40 nucleotides upstream from the 3′ splice site and corresponding branch nucleotides are indicated.

products appeared after 45 min incubation and increased in intensity with longer incubation times. This product corresponds to the size of the primer extension product of exon 5 joined to exon 6. Three shorter primer extension products accumulated in a time dependent manner, and disappeared after debranching. Debranching of the reaction products also gave rise to a new cDNA product of 196 nucleotides, resulting from a run-off extension at the 5' splice cleavage site of intron 5. A similar analysis with the isolated lariat-exon intermediate confirmed that three branched nucleotides were associated with use of this 3' splice site (Figure 3b). Comparison of the dideoxynucleotide sequence with the primer extension products indicates that the three branched nucleotides correspond to positions 25, 29, and 36 nucleotides upstream from the 3' splice site of exon 6. The relative intensity of the primer extension products indicates that the branch sites at positions -25 and -29 are preferred over the branch site at -36. In addition, the weak branch site at -36 is unusual in that it does not utilize an adenosine residue but corresponds to a uridine residue.

Splicing of exon six to exon eight (fibroblast-type splice)

To analyze the branch point used during splicing of exon 6 to exon 8 we used precursor RNA derived from pSP65-p6 linearized with HindIII (Figure 2). RNA transcripts were incubated in *in vitro* splicing reactions for 30 min, 45 min, 60 min, and 120 min, and the products were analyzed as described before. The RNA products were loaded on a 4% polyacrylamide-urea gels (Figure 4a). We observed four processing products corresponding to i) free exon 6, ii) exon 6 spliced to exon 8, iii) free lariat containing intron 6/exon 7/intron 7, and iv) the lariat containing intron 6/exon 7/intron 7/exon 8 sequences. Since HeLa cells express the same isoform as rat fibroblasts, this is the normal splice pattern expected if this pre-mRNA was being spliced in a cell type-specific manner. In addition, there was no indication that either the 5' or 3' splice sites of exon 7 were utilized with this substrate. To determine the location of the branch point, we carried out primer extension analysis using a 20 nucleotide long ^{32}P-labelled oligodeoxynucleotide primer complementary to sequences in the middle of exon 8 (Figure 4b). Figure 4b shows the time course of the reaction products analyzed by primer-extension analysis. One product of 106 nucleotides appeared after 45 min incubation and increased in intensity with longer incubation times. This product corresponded to the size of the primer extension product expected if exon 6 was correctly joined to exon 8. We confirmed this observation at the nucleotide sequence level, by direct sequencing of this primer extension product (data not shown). We also observed a shorter primer extension product of approximately 70 nucleotides long that accumulated in a time dependent manner. This product corresponded to a branch point sequence located 24 nucleotides upstream of the 3' splice site of exon 8. This product was found to be sensitive to treatment with debranching activity (see below, Figure 6).

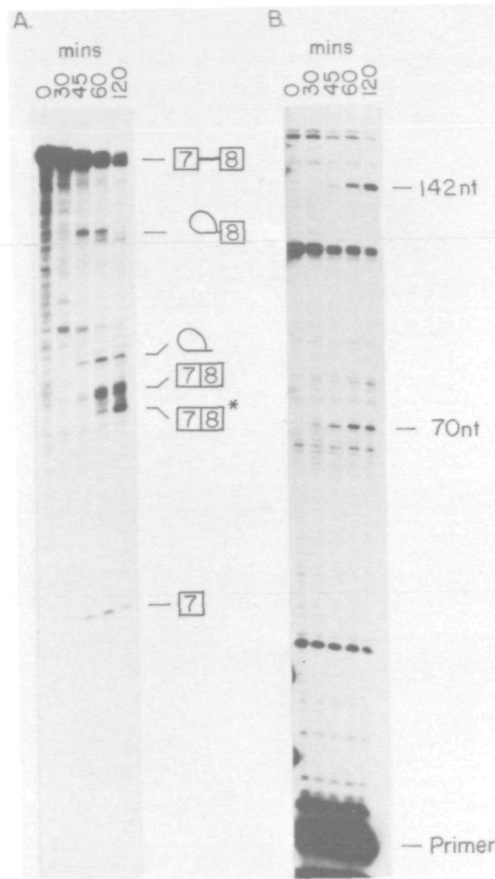*Splicing of exon seven to exon eight (skeletal muscle-type splice)*

To analyze the branch point used during splicing of exon 7 to exon 8 we have used precursor RNA derived from pSP65-p7 linearized with HindIII (Figure 2). RNA transcripts were incubated in *in vitro* splicing reactions and a portion of the RNA was fractionated on a 4% polyacrylamide-urea gel (Figure 5a). We observed four processing products corresponding to i) free exon 7, ii) exon 7 spliced to exon 8, iii) free lariat, and iv) the lariat-exon 8 intermediate. Primer extension analysis was carried out using a 20 nucleotide long ^{32}P-labelled oligodeoxynucleotide primer complementary to sequences in the middle of exon 8 as a primer (Figure 5b). Figure 5b shows the time course of the reaction products analyzed by primer extension analysis. One product of 142 nucleotides appeared after 45
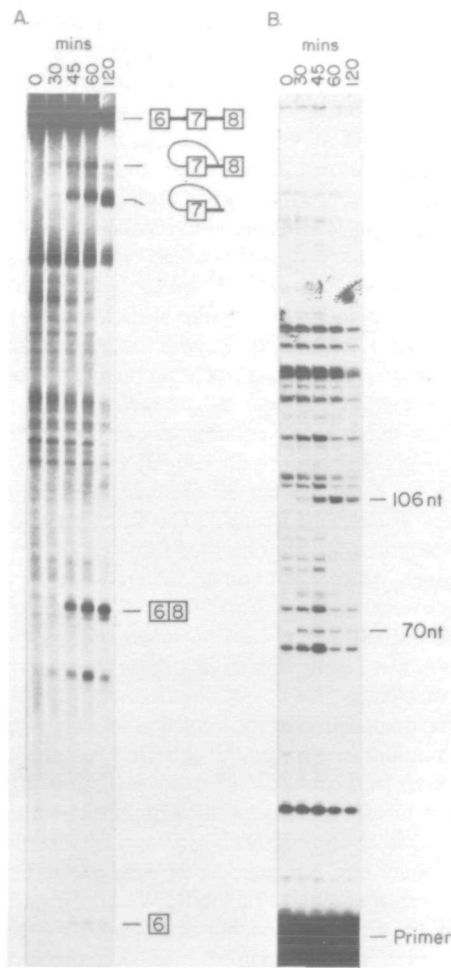
Figure 4. *In vitro* splicing of exon six to exon eight (fibroblast-type splice). (Panel A) Time course (min) of formation of spliced products from transcripts terminating at the SalI site of plasmid pSP65-p6. The ^{32}P-RNA products were separated in 4% denaturing polyacrylamide/ gels. Schematic representation of products are indicated; the boxes represent exon sequences and the lines intron sequences. (Panel B) The *in vitro* processing products shown in panel A were analyzed by primer-extension using a ^{32}P-labelled oligodeoxynucleotide primer specific for sequences in exon 8. Positions of primer extension products corresponding to splice products and intermediates are indicated.

minutes incubation and increased in intensity with longer incubation times. The product corresponded to the size of the primer extension product expected if exon 7 was spliced to exon 8. This was confirmed by direct sequencing of this primer extension product (data not shown). A shorter primer extension product of approximately 70 nucleotides long accumulated in a time dependent manner. This product corresponded to a branch point sequence located 24 nucleotides upstream of the 3' splice site of exon 8, the same branch
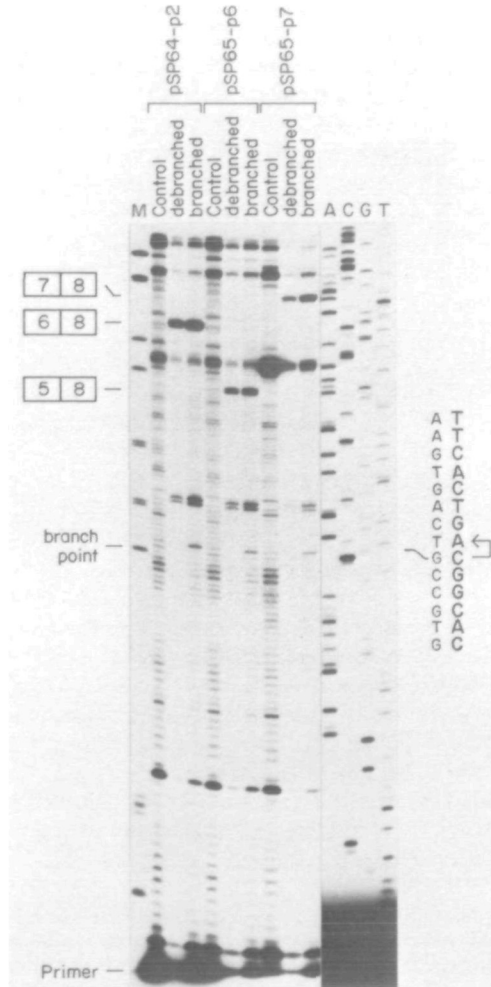
Figure 5. *In vitro* splicing of exon seven to exon eight (skeletal muscle-type splice). (Panel A) Time course (min) of formation of spliced products from transcripts terminating at the HindIII site of plasmid pSP65-p7. The ^{32}P-RNA products were separated in 4% denaturing polyacrylamide gels. Schematic representation of products are indicated; the boxes represent exon sequences and the lines intron sequences. Due to the presence of a 3' exonuclease activity the splice product appears as two bands and is indicated with an asterisk. (Panel B) The *in vitro* processing products shown in panel A were analyzed by primer-extension using a ^{32}P-labelled oligodeoxynucleotide primer specific for sequences in exon 8. Position of primer extension products corresponding to splice products and intermediates are indicated.

point used for the splicing of exon 6 to exon 8 with the pSP65-p6 SalI substrate (Figure 4). This was confirmed when this product was found to be sensitive to treatment with debranching activity (see below, Figure 6).

The splicing of the internal alternative exons (exons 6 and 7) to exon 8 involves joining the alternative 5' splice sites of each of these exons to a common 3' splice site. In addition, we previously demonstrated using pre-mRNAs derived from pSP64-p2 linearized with SalI, that exon 5 was efficiently and accurately spliced to exon 8 (31). It was of interest to determine if splicing of exon 5, as well as exons 6 and 7 to exon 8 involves the use of

5641

Figure 6. A single branch point is utilized during the splicing of exon five, six or seven to exon eight. Precursors derived from pSP64-p2 linearized with SalI, pSP65-p6 linearized with SalI, and pSP65-p7 linearized with HindIII were spliced for 2 h *in vitro*. The *in vitro* processing products were analyzed by primer-extension using a ^{32}P-labelled oligonucleotide primer specific for sequences in the middle of exon 8. Half of the processed RNA was treated with debranching extract prior to primer-extension analysis. Position of the primer extension products corresponding to splice products and intermediates are indicated. The primer extension products were separated on a 8% denaturing polyacrylamide gel and electrophoresed in parallel with dideoxy sequencing reactions performed using single stranded M13 templates primed with the same ^{32}P-labelled oligonucleotide primer to the right. The sequences corresponding to positions −18 through −31 nucleotides upstream from the 3′ splice site and the corresponding branch nucleotide are indicated.

the same or different branch sites. Precursors derived from pSP64-p2, pSP65-p6, and pSP65-p7 were spliced *in vitro* and the products analyzed by primer extension analysis using a 20 nucleotide long ^{32}P-labelled oligodeoxynucleotide corresponding to sequences

in the middle of exon 8 (Figure 6). cDNA products of products of branched and debranched RNA were analyzed on 8% polyacrylamide-8M urea gels along with dideoxy sequence generated on a single stranded M13 template. Comparison of the dideoxynucleotide sequence with the primer extension products indicates a single branched nucleotide is utilized that corresponds to a position 24 nucleotides upstream from the 3′ splice site of exon 8. The bands 8 and 9 nucleotides above the branch site in Figure 6 were not observed in isolated intermediates and were not believed to be represent authentic branch sites (data not shown). Thus, splicing of exon 5, 6 or 7 to exon 8 involves the use of the same branch point.

Splicing of exon five to exon seven (skeletal muscle-type splice)

To analyze the branch point used during splicing of exon 5 to exon 7 we used a pre-mRNA derived from pSP64-p2(7/8) linearized with SalI (Figure 2). We previously demonstrated that exon 5 is spliced accurately and efficiently to exon 7 in HeLa cell nuclear extracts using this precursor (31). The products of the *in vitro* splicing reactions with pre-mRNA transcripts from pSP64-p2(7/8) linearized with SalI are shown in Figure 7. Following the splicing reactions the products were divided and a portion loaded on a 4% polyacrylamide-urea gel (Figure 7a). To map the branch point utilized during the splicing of exon 5 to exon 7 primer extension experiments were carried using three oligonucleotides complementary to sequences in the middle of exon 7 (Ex.7), the 3′ end of intron 6 (int 6−3′), and 94−110 nucleotides from the 3′ end of intron 6 (intron 6x). Surprisingly, analysis of the primer extension products using ^{32}P-labelled oligonucleotide primers complementary to sequences in the middle of exon 7 (Ex.7 primer) or the 3′ end of intron 6 (int 6−3′ primer) revealed a putative branch nucleotide located approximately 190 and 145 nucleotides upstream of these respective primers (Figure 7b). This would correspond to a lariat branch point located 145 nucleotides upstream of the 3′ splice site of exon 7. To investigate this branch point in greater detail primer extension analysis was performed using a ^{32}P-labelled oligonucleotide complementary to sequences at the 3′ end of intron 6 (Figure 7c). Figure 7c shows the primer extension analysis using the total reaction products (Figure 7c, lanes 1 and 2) as well as the isolated lariat/exon intermediate (Figure 7c, lanes 3 and 4) and the isolated lariat (Figure 7c, lanes 5 and 6). As indicated, three primer extension products of 144, 147, and 153 nucleotides were detected using all three RNA templates. When the primer extension was carried out on debranched RNA the 144, 147 and 153 nucleotide bands disappeared, and a 514 nucleotide band appeared corresponding to extension to the 5′ end of the linearized (debranched) intron. A similar analysis was performed using a ^{32}P-labelled oligonucleotide complementary to sequences 94−110 nucleotides upstream of the 3′ splice site of exon 7 (Figure 7d, intron 6x primer). Figure 7d shows the time course of the reaction products analyzed by primer extension using this primer (lanes 1 through 6). Three products appeared as early as 30 min and accumulated in a time dependent manner. These three products corresponded to branch point sequences located 144, 147 and 153 nucleotides upstream of the 3′ splice site of exon 8. Based on the relative intensity of the primer extension stop products the branch point corresponding to position −147 was found to be the preferred branch nucleotide. The use of this branch site location was further confirmed when these products were found to be sensitive to treatment with debranching activity. Figure 7d also shows the primer extension analysis of the isolated lariat-exon 7/8 intermediate using the same oligonucleotide (lanes 7 and 8). At this level of exposure the expected band of 421 nucleotides corresponding to the 5′ end of the linearized intron following treatment with debranching activity is not visible in this autoradiograph (lane 7). Identical results were also obtained using the isolated lariat
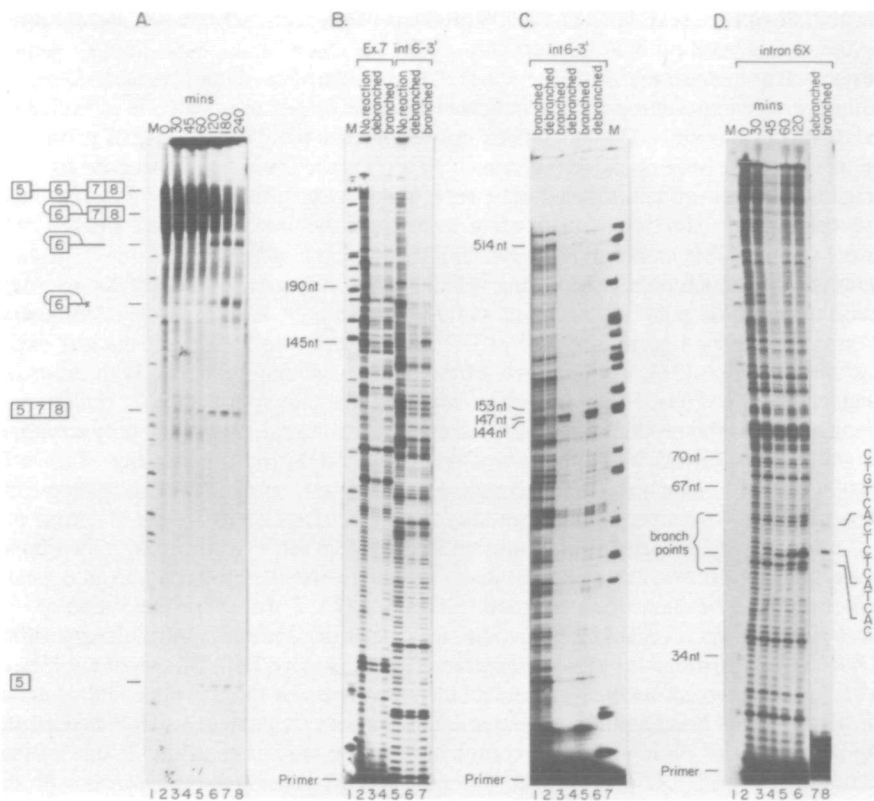
Figure 7. *In vitro* splicing of exon five to exon 7 (skeletal muscle-type splice). (Panel A) time course (min) of formation of spliced products from transcripts terminating at the SalI site of plasmid pSP64-p2(7/8). The ^{32}P-RNA products were separated on 4% denaturing polyacrylamide gels. Schematic representation of products and intermediates are indicated at the left of the panel; from top to bottom, i) pre-mRNA, ii) intron/exon intermediate, iii) free lariat, iv) free lariat but 3' end of intron 6 partly degraded by 3' exonucleolytic activity, v) splice product and vi) free exon 5. Lane M is markers using pBR322 digested with MspI. (Panel B) The *in vitro* processing products shown in the left panel were analyzed by primer-extension using a ^{32}P-labelled oligodeoxynucleotide primer complementary to sequences in the middle of exon 7 and a ^{32}P-labelled oligonucleotide primer complementary to the 3' end on intron 6. Half of the processed RNA was treated with debranching extract prior to the primer-extension reaction. Position of the extension product corresponding to spliced RNA is indicated. Lane M is markers using pBR322 digested with MspI. (Panel C) The *in vitro* processing products shown in panel A were analyzed by primer-extension using a primer complementary to the 3' end of intron 6. The positions of the branch nucleotides are indicated. The total reaction products (lanes 1 and 2) as well as the lariat-exon 7/8 intermediate (lanes 3 and 4) and free lariat (lanes 5 and 6) shown in panel A were isolated and subjected to primer-extension analysis. Half of the processed RNA was treated with debranching extract prior to the primer-extension reaction. Position of the extension products are indicated at the left of the panel. Lane M is markers using pBR322 digested with MspI. (Panel D) The *in vitro* processing products shown in panel A were analyzed by primer-extension using a primer specific for sequences in intron 6 corresponding to nucleotides 94−110 nucleotides upstream from the 3' splice site of exon 7 (intron 6x). The positions of the branch nucleotides are indicated. The lariat-exon7/8 intermediate shown in panel A was isolated and subjected to primer-extension analysis using the same ^{32}P-labelled oligodeoxynucleotide. Half of the processed RNA was treated with debranching extract prior to the primer-extension reaction. Position of the extension products and corresponding DNA sequence is shown to the left and right of panel, respectively. Lane M is markers using pBR322 digested with MspI.

(data not shown). These results confirm that splicing of exon 5 to exon 7 is associated with the use of branch points located 144−153 nucleotides upstream of the 3' splice site of exon 7.

DISCUSSION

The present study show some features of branch point selection involved in the utilization of alternative splice sites in tropomyosin pre-mRNAs including: i) the use of multiple branch points associated with utilization of each alternative 3' splice site and, ii) the use of branch points located a long distance (144−153 nucleotides) from the 3' splice site of exon 7 (Figure 8). Splicing of exon 5 to exon 6 (fibroblast-type splice) involved the use of two adenosine residues and a uridine residue at positions 25, 29, and 36 nucleotides upstream of the 3' splice site of exon 6. The two adenosine residues appear to be preferred over the uridine. While most studies have demonstrated the use of an adenosine residue as the branch acceptor nucleotide, a number of studies have reported that other nucleotides can participate in branch site formation (56−58). For example, studies of lariat formation in intron A of human growth hormone pre-mRNA demonstrated that branch point utilization occurs mainly at a cytidine residue located 28 nucleotides upstream of the 3' splice site, and to a lesser extent at two uridine residues located 22 and 36 nucleotides upstream from the 3' splice site (58). The splicing of exon 5 to exon 7 (skeletal muscle-type splice) involves the use of three branch residues located 144, 147, and 153 nucleotides upstream from the 3' splice site of exon 7. These results are in contrast to the majority of introns that have been studied in which only a single adenosine residue, located within 18−40 nucleotides from the 3' splice, site is utilized during lariat formation (7−9,18,59,60).

The most striking observation of the present study was that splicing of exon 5 to exon 7 (skeletal muscle-type splice) was associated with the use of one of three A residues located 144, 147, and 153 nucleotides from the 3' splice site of exon 7 (Figure 7). At present it is not known if utilization of these branch points play a role in alternative splice site selection. HeLa cells do not normally carry out this particular splice since the splicing of exon 5 to exon 7 is specific for skeletal muscle (46). However, we previously demonstrated that splicing of exon 5 to exon 7 occurred both *in vitro* and *in vivo* using a HeLa cell system using precursors in which exon 7 was first joined to exon 8 (31). Therefore we believe that splicing of exon 5 to exon 7 in HeLa cells is associated with the use of branch residues located 144−153 nucleotides upstream of the 3' splice site of exon 7. Whether these same branch points are utilized in skeletal muscle cells will require the development of *in vitro* systems derived from skeletal muscle that accurately reflect the *in vivo* splicing observed in this tissue.

The complete sequence of the intron preceding exon 7 is shown in Table I. Inspection of the 3' end of intron 6 reveals that it contains a number of potential branch site sequences which are located close to the 3' splice site of exon 7 but are not utilized during the splicing of exon 5 to exon 7. For example, relatively good branch point consensus sequences are found at positions 29, 74 and 87 upstream of the 3' splice site. The use of branch point sequences located 144−153 nucleotides upstream of the 3' splice of exon 7 raises some intriguing questions regarding the mechanism of branch point selection. As described above, there are a number of potential branch point residues located relatively close to the 3' splice site that are not utilized (Table I). Why these sites are not used is not known. Formation of the branch nucleotide involves the interaction U2 snRNP with the pre-mRNA branch point (13, 20−22). Mutational analyses demonstrate that the sequences most
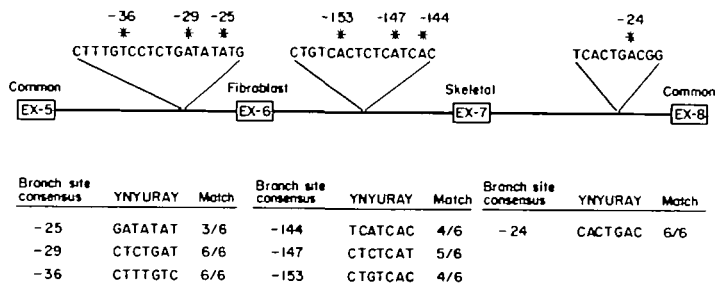
Figure 8. Summary of branch sites utilized in tropomyosin pre-mRNAs. The locations of lariat branch sites used in tropomyosin splicing are indicated with asterisks. The sequences of the branch sites used in introns 5, 6 and 7 are compared to the lariat branch site consensus sequence in which Y =pyrimidine, R =purine, and N indicates any of the four nucleotides. Homology to the consensus branch point sequence is indicated. The distance between each branch site and 3′ splice junction is given.

important for the U2 snRNP-branch point interaction are not the nucleotides in the branch point region itself, but rather the 3′ splice site consensus including the polypyrimidine tract and AG dinucleotide (61−62). Deletions or mutations in the 3′ splice site decrease or abolish U2 snRNP binding and splicing complex formation (62−65). On the other hand, mutations in the branch point sequences generally are not associated with the loss of U2 snRNP binding but instead result in activation of cryptic branch points (19,66). These results demonstrate that the sequences near the 3′ splice site are important for the interaction of U2 snRNP at a given branch point. How sequences in the 3′ splice site of intron 6 contribute to branch site selection remains to be determined. One possibility is that the sequences in intron 6 may establish an RNA conformation that favors the use of particular branch residues, in this case those located 144−153 nucleotides upstream of the 3′ splice site. Alternatively, factors may interact with sequences within the 3′ splice site of exon 7 leading to utilization of a particular branch point. The interaction of factors, such as U2AF, may regulate the binding of U2 snRNP and lead to utilization of a particular branch site (23). Since U2AF requires the AG dinucleotide at the 3′ splice site, this would require the interaction of the U2-snRNP with U2AF over a relatively long distance. In addition, other components may interact with intron 6 that are required for utilization of the 3′ splice site of exon 7.

The use of branch points located a relatively long distance from a 3′ splice site may be an essential feature of some alternatively spliced exons. Alternative splicing of adenovirus E1A RNA was also reported to involve the use of multiple branch points located 51−59 nucleotides from the 3′ splice site (34). However, the function of these unusual branch sites in splice site selection remain to be determined. Recently, studies of alternative splicing of the rat alpha-tropomyosin gene revealed a novel mechanism involving branch point selection (67). Alternative splicing of alpha-tropomyosin pre-mRNA involves mutually exclusive use of exons 2 and 3. Exons 2 and 3 are never spliced together in any cell type. The intron between exon 2 and exon 3 is 218 nucleotides in length. Interestingly it was found that the branch point utilized upstream of exon 3 was located 177 nucleotides from the 3′ splice site of this exon. This branch site is only 42 nucleotides from the 5′ splice site of exon 2. As a result exon 2 is unable to splice to exon 3 presumably because the relatively short distance between the 5′ splice site of exon 2 and the downstream branch

5646

Table I.   Analysis of Intron Sequences Preceding Exon 7

```
-291 gtagtagcctctctaggcctttctaggcaatggcaccttttctcacttca


-241 cgtccctccccagctcagcatctcaggctactgtctctctcctgcggtag

                                           *    *   *
-191 gagggggcgggggcggacgttccaaacagcctgctgtcactctcatcact


-141 ttgctcttctcttttctcctcctcccctcctccactgtgccacccctccc


 -91 cctaaccccaccccctcaccccgtcgtcgcgccaccccactgtctcacct

                                         -1
 -41 cactgtgccctcacgctccatcctgccacacgcccctgcag:   Exon 7
```

Branch Site Consensus	YNYURAY	Match	Use
-22	GCTCCAT	3/6	nd
-29	CCCTCAC	5/6	nd
-40	ACCTCAC	4/6	nd
-45	GTCTCAC	4/6	nd
-53	ACCCCAC	3/6	nd
-58	GCGCCAC	2/6	nd
-74	CCCTCAC	5/6	nd
-82	ACCCCAC	3/6	nd
-87	CCCTAAC	6/6	nd
-100	CTGCCAC	3/6	nd
-108	CCTCCAC	4/6	nd
-144	TCATCAC	4/6	yes
-147	CTCTCAT	5/6	yes
-153	CTGTCAC	4/6	yes

The nucleotide sequence of intron 6 is shown as well as
the locations of the lariat branch sites used in splicing
exon 5 to exon 7 which are indicated with asterisks. The
sequences of branch sites utilized are compared with other
consensus branch point sequences found upstream of the 3'
splice site of exon 7. Y = pyrimidine, R = purine, N
indicates any of the four nucleotides, nd = not detected.

results in steric hindrance that prevent splicing factors from interacting with these splice sites. Such a model cannot alone explain the mutually exclusive use of exon 6 and 7 during alternative splicing of the rat beta-tropomyosin gene. The lariat branch points located upstream of exon 7 are 136−147 nucleotides from the 5' splice site of exon 6. At present the molecular basis for the lack of splicing of exon 6 to exon 7 is unknown.

The regulation of alternative splice site selection will involve the interaction of cellular factors with specific sequences in a pre-mRNA. Accordingly, those sequences located between the 3' splice site of exon 7 (skeletal muscle-type splice) and the upstream branch points may play a role during the regulation of alternative splicing of tropomyosin pre-mRNAs. For example, these sequences may function by providing a binding site for factors which regulate tissue-specific splicing. In this respect it is of interest that the 3' end of intron 6 contains an unusual pyrimidine content (Table I). The region between the first branch point, located at position −144, and the 3' splice site is approximately 80% pyrimidines. In addition, the region between −89 through −143 contains a relatively long pyrimidine-rich tract, containing only 4 purines in this 55 nucleotide long sequence. The high pyrimidine content of this 3' splice site may play a role in splice site selection. For example, the polypyrimidine tract immediately downstream of the branch points used in intron 6 may contribute to the use of these branch point sequences. We have observed in other tropomyosin genes that undergo alternative splicing that the relevant 3' splice sites contain long pyrimidine-rich sequences immediately upstream from the 3' splice site (68−70). These include exons 3, 8, 10 and 11 of the rat alpha-tropomyosin gene (68,69) and exon 8 of the human slow alpha-tropomyosin gene (70). Furthermore, exon 10 (skeletal muscle-type splice) of the rat tropomyosin 1 gene contains a long pyrimidine stretch immediately upstream of the 3' splice site (Erster and Helfman, unpublished observations). These pyrimidine tracts may represent a cis-acting element involved in the choice of alternative exons. This hypothesis is supported by recent studies of SV40 early pre-mRNAs indicating that the pyrimidine content at the 3' splice site had substantial effects on the alternative splicing *in vivo* (27). Experiments are in progress to determine the role of the pyrimidine-rich region upstream of exon 7 in splice site selection and branch site utilization in tropomyosin pre-mRNAs. Further characterization of the factors involved in use of this branch site should lead to a better understanding of the general mechanism of branch point selection and provide clues concerning the regulation of alternative splicing.

ACKNOWLEDGEMENTS

REFERENCES

1. Padgett, R.A., Grabowski, P.J., Konarska, M., Seiler, S. and Sharp, P.A. (1986). Annu. Rev. Biochem. 55, 1119−1150.
2. Green, M.R. (1986). Ann. Rev. genet. 20, 671−708.
3. Krainer, A.K., and Maniatis, T. (1988) *In:* Transcription and Splicing (Eds. Hames, B.D. and Glover, D.M.) p. 131−206.

4. Hernandez, N., and Keller, W. (1983). Cell 35, 89−99.
5. Hardy, S.F., Grabowski, P.J., Padgett, R.A., and Sharp, P.A. (1984). Nature 308, 375−377.
6. Krainer, A.R., Maniatis, T., Ruskin, B., and Green, M R. (1984). Cell 36, 993−1005.
7. Grabowski, P.J., Padgett, R.A., and Sharp, P.A. (1984). Cell 37, 415−427.
8. Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984). Cell 38, 317−331.
9. Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984). Science 225, 898−903.
10. Breathnach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349−383.
11. Mount, S. (1982). Nucl. Acids. Res. 10, 459−472.
12. Mount, S.M , Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J.A. (1983). Cell 33, 509−518.
13. Black, D.L., Chabot, B., and Steitz, J.A. (1985). Cell 42, 737−750.
14. Zhuang, Y., and Weiner, A.M. (1986) Cell 46, 827−835.
15. Gerke, V., amd Steitz, J.A. (1986). Cell 47, 973−984.
16. Tazi, J., Albert, C., Temsamani, J., Reveillaud, I., Cathala, G., Brunel, C., and Jeanteur, P. (1986). Cell 47, 755−766.
17. Keller, E.B., and Noon, W.A. (1985). Nucl. Acids. Res. 13, 4971−4981.
18. Reed, R., and Maniatis, T. (1985). Cell 41, 95−105.
19. Ruskin, B., Greene, J.M., and Green, M.R. (1985). Cell 41, 833−844.
20. Konarska, M.M., and Sharp, P.A. (1986). Cell 46, 845−855.
21. Parker, R., Siliciano, P., and Guthrie, C. (1987) Cell 49, 229−239.
22. Bindereif, A., and Green, M.R. (1987) EMBO J. 6, 2415−2424.
23. Ruskin, B., Zamore, P.D., and Green, M.R. (1988). Cell 52, 207−219.
24. Breitbart, R.E., Andreadis, A., and Nadal-Ginard, B. (1987). Ann. Rev. Biochem. 56, 467−495.
25. Fu, X-Y, and Manley, J.L. (1987). Mol. Cell. Biol. 7, 738−748.
26. Zhuang, Z., Leung, H., and Weiner, A. (1987) Mol. Cell. Biol. 7, 3018−3020.
27. Fu, Xin-Yuan, Ge, H., and Manley, J.L. (1988). EMBO J. 7, 809−817.
28. Somasekhar, M.B., and Mertz, J.E. (1985). Nuc. Acids Res. 13, 5591−5609.
29. Reed, R., and Maniatis, T. (1986). Cell 46, 681−690.
30. Mardon, H.J., Sebastio, G., and Baralle, F.E. (1987). Nuc. Acids Res. 15, 7725−7733.
31. Helfman, D.M., Ricci, W.M., and Finn, L.A. (1988). Genes and Dev. 2, 1627−1638.
32. Noble, J.C.S., Pan, Z. -Q., Prives, C., and Manley, J.L. (1987) Cell 50, 227−236.
33. Noble, J.C.S., Prives, C., Manley, J.L. (1988). Genes and Dev. 2, 1460−1475.
34. Gattoni, R., Schmitt, P., and Stevenin, J. (1988). Nuc. Acids Res. 16, 2389−2409.
35. Khoury, G., Gruss, P., Dhar, R., and Lai, C. (1979). Cell 18, 85−92.
36. Munroe, S.H. (1984). Nucl. Acids. Res. 12, 8437−8455.
37. Solnick, D. (1985). Cell 43, 667−676.
38. Eperon, L.P., Estibeiro, J.P., and Eperon, I.C. (1986). Nature 324, 280−282.
39. Eperon, L.P., Graham, I.R., Griffiths, A.D., and Eperon, I.C. (1988). Cell 54, 393−401.
40. Leff, S.E., Evans, R.M., and Rosenfeld, M.G. (1987). Cell 48, 517−524.
41. Edlind, T.D., Cooley, T.E., and Ihler, G.M. (1987). Nucl. Acids Res. 15, 8566-
42. Breitbart, R.E., and Nadal-Ginard, B. (1987). Cell 49, 783−803.
43. Schmitt, P., Gattoni, R., Keohavong, P., and Stevenin, J. (1987). Cell 50, 31−39.
44. Stephens, C., and Harlow, E. (1987) EMBO J. 6, 2027−2036.
45. Ulfendahl, P.J., Linder, S., Kreivi, J.P., Nordqvist, K., Svensson, C., Hultberg, H., and Akusjarvi, G. (1987) EMBO J. 6, 2037−2044.
46. Helfman, D.M., Cheley, S., Kuismanen, E., Finn, L.A., and Yamawaki-Kataoka, Y. (1986). Mol. Cell. Biol. 6, 3582−3595.
47. Yamawaki-Kataoka, Y., and Helfman, D.M. (1985). J. Biol. Chem. 260, 14440−14445.
48. Eghtedarzadeh, M.K., and Henikoff, S. (1986). Nuc. Acids Res. 14, 5115.
49. Konarska, M.M., Padgett, R.A., and Sharp, P.A. (1984). Cell 38, 731−736.
50. Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Nuc. Acids Res. 11, 1475−1489.
51. Ruskin, B., and Green, M.R. (1985). Science 229, 135−140.
52. Ghosh, P.K., Reddy, V.B., Swinscoe, J., Lemowitz, P., and Weissman, S.M. (1978). J. Mol. Biol. 126, 813−846.
53. Erster, S.H., Finn, L.A., Frendewey, D.A., and Helfman (1988). Nuc. Acids Res. 16, 5999−6014.
54. Maxam, A.M., and Gilbert, W. (1980). Methods Enzymol. 65, 499−560.
55. Sanger, F., Nicklen, S., and Coulson, A.R. (1977). Proc. Natl. Acad. Sci. USA 74, 5463−5467.
56. Hornig, H., Aebi, M., and Weissmann, C. (1986). Nature 324, 589−591.

57. Freyer, G.A., Arenas, J., Perkins, K.K., Furneaux, L., Pick, B., Young, B., Roberts, R.J., and Hurwitz, J. (1987). J. Biol. Chem. 262, 4267–4273.
58. Hartmuth, K., and Barta, A. (1988). Mol. Cell. Biol. 8, 2011–2020.
59. Zeitlin, S., and Efstratiadis, A. (1984). Cell 39, 589–602.
60. Konarska, M.M., Grabowski, P.J., Padgett, R.A., and Sharp, P.A. (1985). Nature (London) 313, 552–557.
61. Ruskin, B., and Green, M.R. (1985). Cell 43, 131–142.
62. Chabot, B., and steitz, J.A. (1987). Mol. Cell., Biol. 17, 281–293.
63. Ruskin, B., and Green, M.R. (1985). Nature 317, 732–734.
64. Frendewey, D., and Keller, W. (1985). Cell 42, 355–367.
65. Bindereif, A , and Green, M.R. (1986). Mol. Cell. Biol. 6, 2582–2592.
66. Padgett, R.A., Konarska, M.M., Aebi, H., Hornig, H, Weissmann, C., and Sharp, P.A. (1985). Proc. Natl. Acad. Sci. 82, 8349–8353.
67. Smith, C.W.J., and Nadal-Ginard, B. (1988). Cell 56, 749–758.
68. Ruiz-Opazo, N., and Nadal-Ginard, B. (1987). J. Biol. Chem. 262, 4755–4765.
69. Wieczorek, D.F., Smith, C.W.J., and Nadal-Ginard, B. (1988). Mol. Cell. Biol. 8, 679–694.
70. Clayton, L., Reinach, F.C., Chumbley, G.M., and MacLeod, A.R. (1988) J. Mol. Biol. 210, 507–515.

> **This article, submitted on disc, has been automatically converted into this typeset format by the publisher.**