

Myosin Light-Chain 1/3 Gene Alternative Splicing: *cis* Regulation Is Based upon a Hierarchical Compatibility between Splice Sites

MARIA E. GALLEGU AND BERNARDO NADAL-GINARD*

Laboratory of Molecular and Cellular Cardiology, Howard Hughes Medical Institute; Department of Cardiology, Children's Hospital; and Departments of Pediatrics, Physiology and Biophysics, Harvard Medical School, Boston, Massachusetts 02115

Received 9 November 1989/Accepted 10 January 1990

The mechanisms involved in the selective joining of appropriate 5' and 3' splice sites are still poorly understood in both constitutive and alternatively spliced genes. With two promoters associated with different exons, the myosin light-chain 1/3 gene generates two pre-mRNAs that also differ by the use of a pair of internal exons, 3 and 4, that are spliced in a mutually exclusive fashion. When the promoter upstream from exon 1 is used, only exon 4 is included. If the promoter upstream from exon 2 is used, only exon 3 is included. In an attempt to understand the molecular basis for the mutually exclusive behavior of these two exons and the basis of their specific selection, a number of minigene constructs containing exons 3 and 4 were tested in a variety of homologous or heterologous *cis* and *trans* environments. The results demonstrate that the mutually exclusive behavior of myosin light-chain exons 3 and 4 and selection between the two exons are *cis* regulated and are affected by the nature of the flanking sequences. Both exons competed for the common flanking 5' and 3' splice sites. Flanking exons were found that favored inclusion into mature mRNA of exon 3, exon 4, both, or neither, suggesting a specific cooperative interaction between certain 5' and 3' splice sites. Thus, alternative splicing of myosin light-chain 1/3 pre-mRNAs is regulated in *cis* by a hierarchy of compatibilities between pairs of 5' and 3' splice sites.

In higher eucaryotes most genes produce pre-mRNAs containing intervening sequences that have to be removed to generate mature mRNAs. Three sequence elements have been found to be required for splicing: the 5' and 3' splice sites at each end of the intron and the branch point sequence (BPS) commonly located between 20 and 40 nucleotides (nt) upstream of the 3' exon (17, 38, 51). Recent results strongly suggest that the 5' splice site and the BPS and its polypyrimidine tract are the two most important landmarks in the intron (17, 38, 43, 53, 55). (For reasons of simplicity from here on we will refer to the BPS, polypyrimidine tract, and acceptor site together as the 3' splice site.) These sequence elements interact with splicing factors to form a large complex, termed the spliceosome (5, 10, 15), that involves the interaction of small nuclear ribonucleoproteins and a number of other proteins with the pre-mRNA (29, 31).

One of the main unsolved problems in splicing is understanding the mechanism by which the splicing machinery precisely recognizes the correct 5' and 3' splice sites to be joined together. This is so because the sequences conforming to the 5' and 3' splice site consensus sequences are redundant and can be found repeated within introns and exons (37, 50). These cryptic sites, however, are not used unless the bona fide splice site is mutated (17, 38). Thus, questions arise as to how a functional splice site is defined—a consensus sequence is obviously not sufficient—and also as to the mechanism that restricts the pairing of the correct donor and acceptor.

Different sequence elements have been suggested to play a role in splice site selection, including secondary structure of the mRNA (8, 56, 57) and the degree of match to the consensus 5' splice site and BPS (1, 36, 43, 63, 64). These elements, however, by themselves cannot account for the accuracy and efficiency of the splicing process. This problem

becomes even more complex in the case of alternatively spliced transcripts (2, 4, 54). In such cases, proximal functional splice sites are sometimes bypassed in favor of more distal ones. In most cases, this bypass is regulated in a tissue- and/or developmental stage-specific manner. Striking examples of such alternative splicing are exhibited by genes with pairs of mutually exclusive exons. In such pairs, only one of the exons in the pair is included in the mature mRNA, whereas the other one is excluded. For genes with mutually exclusive exon pairs, no cellular mRNAs have yet been detected that incorporate both exons, despite the fact that the introns between them contain functional 5' and 3' splice sites. An understanding of this mode of alternative splicing requires an explanation, first, for the inability of the two exons to be spliced together and, second, for the basis of the cell- and developmental stage-specific selection. In the case of the α -tropomyosin (α -TM) exons 2 and 3 (53), the mutually exclusive behavior is due to the proximity of the BPS upstream of exon 3 to the donor of exon 2, which at 41 nt absolutely prevents splicing complex formation by steric interference. It is not clear, however, whether this is a general mechanism to enforce mutually exclusive behavior or whether it is limited to this pair of exons. Moreover, there are cases of mutually exclusive splicing of pre-mRNAs originating from complex transcription units that have different 5' or 3' sequences, each associated with a particular splicing pattern (2, 40, 45). In these cases each transcript has only one splicing pathway available to it. It is possible that the splicing pattern of these mRNAs is regulated entirely in *cis* through the structure of the pre-mRNA and determined by promoter or poly(A) addition site choice.

We are interested in testing the generality of the mechanisms uncovered for α -TM exons to gain insight into the regulation of mutually exclusive splicing of pre-mRNAs with different primary structures. For this reason we have analyzed the splicing pattern of the two mutually exclusive

* Corresponding author.

exons of the rat skeletal alkali myosin light-chain (MLC) 1/3 gene (MLC1/3). This gene generates two pre-mRNAs differing in length and sequence, through the use of two different promoters (40, 58). The two transcripts have different 5' exons and are spliced in a transcript-specific manner to generate MLC1 or MLC3 mRNAs through the mutually exclusive splicing of exons 4 and 3, respectively.

Here, we show that both the mutually exclusive behavior and the choice between MLC1/3 exons 3 and 4 is *cis* regulated by local sequences in the vicinity of these exons. In contrast to α -TM exons 2 and 3, MLC1/3 exons 3 and 4 are able to splice together in the absence of competition from flanking splice sites. There is a competition between the two exons for the flanking splice sites, and deletion of the preferred exon results in inclusion of the normally excluded one. Interestingly, the nature of the flanking exons plays a role in the selection of exon 3 or 4. Flanking exons were found that favored inclusion into the mature mRNA of exon 3, exon 4, both, or neither, suggesting some form of specific cooperative interaction between certain donor and acceptor sites. These results strongly suggest that alternative splicing of MLC1/3 pre-mRNAs is regulated *in cis* by a hierarchy of compatibilities between specific pairs of 5' and 3' splice sites.

MATERIALS AND METHODS

Plasmid constructions. Plasmids constructs containing MLC1/3 genomic sequences were prepared by standard procedures (28). MLC clones were derived from the genomic clone λ LCH14 (40). pSVpA is a pUC18-based vector that carries the simian virus 40 (SV40) promoter-enhancer element from pr12-SV40-46 (25) linked to the 671-base-pair *Pst*I-*Bam*HI fragment containing the 3' end of the SV40 large T antigen via a unique *Hind*III site. pSVLC2345 is pSVpA containing a 7.5-kilobase *Hind*III fragment from λ LCH14 that includes MLC exons 2, 3, 4, and 5. This fragment includes 800 base pairs of the intron 5' to exon 2, which contains the MLC3 promoter, and 1,700 base pairs of the intron 3' to exon 5. pSVLC1345 is derived from pSVLC2345, in which MLC exon 2 was replaced by MLC exon 1 and 362 nucleotide (nt) of 3' intron fused to 410 nt of intron 5' to MLC exon 3. The MLC1/3 insulin constructs are derived from pr12-SV40-46. They are in a PUC18 vector in which the *Eco*RI site at the 5' end of the promoter was eliminated, leaving a unique *Eco*RI site within intron 2 (pSVIns). pSVIns3, contains MLC exon 3, 410 nt of its 5' intron, and 30 nt of its 3' intron inserted into the *Eco*RI site in pSVIns. pSVIns4 contains on the pSVIns *Eco*RI site MLC exon 4 with 100 nt of 5' intron and 148 nt of 3' intron. pSVIns34 was obtained by cloning a 976-base-pair *Eco*RI fragment containing 410 nt of the intron 5' to exon 3, the complete intron between exons 3 and 4, and 265 nt of the 3' intron. pSVIns43 has MLC exon 4 flanked by 100 nt of 5' intron and 264 nt of 3' intron. Downstream from this exon is MLC exon 3 with 410 nt of 5' intron and 34 nt of 3' intron. pSVTm34 and pSVTm43 contain the same MLC fragments described for the pSVIns constructs into pTMSR. This vector has α -TM exon 1 and 333 nt of 3' intron ligated to α -TM exon 4 and 67 nt of 5' intron through an *Eco*RI site in pSVpA (61). pSVLC234Ins3 and pSVLC234Tm4 were made by replacing exon 5 in pSVLC2345 with insulin exon 3 or α -TM exon 4. Both of them were ligated to MLC exon 4 and 460 nt of 3' intron. Insulin exon 3 included 482 nt of 5' intron, and α -TM exon 4 carried 67 nt of 5' intron. pSVTm1LC345 was made by replacing exon 2 in pSVLC2345 with α -TM exon 1,

containing 333 nt of 3' intron ligated to MLC exon 3 and 420 nt of 5' intron. The construct used *in vitro* contains in pGEM3 an *Nde*I-*Nde*I MLC DNA fragment including 259 nt of intron 5' to exon 3, exon 3, the complete intron between exons 3 and 4, and exon 4 together with 26 nt of 3' intron.

More specific details can be obtained upon request.

Cell lines and culture conditions. Eucaryotic cell culture media and sera were obtained from GIBCO Laboratories (Chagrim Falls, Ohio). HeLa (48) and COS (14) cells were grown in Dulbecco modified Eagle medium (DMEM) supplemented with 10% fetal calf serum. For mouse skeletal C2C12 myoblasts (62) and rat skeletal L6E9 myoblast cells (35), the growth medium was DMEM supplemented with 20% fetal calf serum; differentiation medium was DMEM with 5% horse serum.

Cells on 100-mm plates were transfected using 10 μ g of DNA by the calcium phosphate coprecipitation method (16), followed by glycerol shock 6 h later (39). The cells were harvested 65 h after glycerol shock.

Isolation and analysis of cellular RNA. Cytoplasmic RNA was isolated by a modification of the hot phenol procedure (9) in the presence of 5 mM vanadyl ribonucleoside complex. RNAs were stored in ethanol at -20°C .

End-labeled probes were oligonucleotides prepared in an automated DNA synthesizer (8600 series; Bioscience Inc., San Rafael, Calif.). Probes were labeled at the 5' end with [γ - ^{32}P]ATP and T4 polynucleotide kinase and at the 3' end with [α - ^{32}P]ddATP and terminal transferase.

S1 nuclease mapping was performed by standard procedures (3). Hybridization of total cytoplasmic RNA (5 to 20 μ g) with a probe (5×10^4 cpm) was in 80% formamide-400 mM NaCl-10 mM piperazine-*N,N'*-bis(2-ethanesulfonic acid)-0.5% sodium dodecyl sulfate-1 mM EDTA. The hybridization mixture was heated for 1 h at 65°C and then allowed to cool to 38°C overnight. S1 digestion was carried out for 1 h at 25°C with 60 U of S1 nuclease. Nuclease-resistant fragments were separated by denaturing polyacrylamide gel electrophoresis and detected by autoradiography.

***In vitro* splicing and RNA analysis.** Transcripts for *in vitro* splicing were prepared with SP6 polymerase as previously described (33, 41, 47). HeLa cell nuclear extracts were prepared by the method of Dignam et al. (6). Splicing reactions were carried out as described by Krainer et al. (22). The amount of RNA used was approximately 20 fmol. RNA species were identified by denaturing polyacrylamide gel electrophoresis. Splicing intermediates and products were identified by the kinetics of their appearance; the lariat forms were identified by their change in mobility in different percentages of acrylamide (47) and after enzymatic debranching in HeLa S-100 fractions (46).

RESULTS

Mutually exclusive splicing of MLC1/3 exons 3 and 4 is determined by *cis*-proximal sequences and is not muscle specific. The MLC1/3 gene generates two different mRNAs through the use of two promoters and the mutually exclusive splicing of exons 3 and 4 (Fig. 1). The two promoters are separated by 10 kilobases, giving rise to two primary transcripts very different in length and sequence (40, 58). When the 5'-proximal promoter-initiating transcription at exon 1 is used, splicing to exon 2 is not possible because it lacks an acceptor site. Instead, exon 1 splices to the acceptor of exon 4, bypassing exons 2 and 3 to give rise to MLC1 mRNA. When the downstream promoter is used, exon 2 is ligated to exon 3; this is joined to exon 5, excluding exon 4, leading to

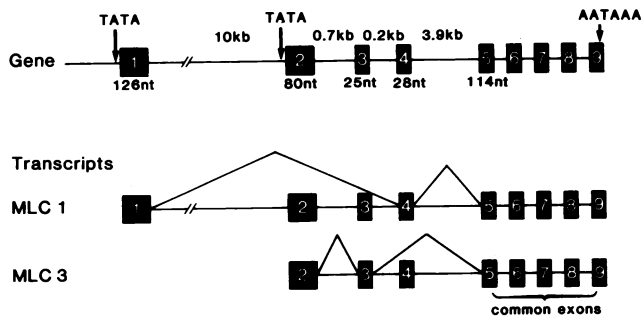


FIG. 1. Schematic diagram of the rat MLC1/3 gene, its primary transcripts, and the two alternative mRNA splicing pathways used to generate MLC1 and MLC3 mRNA, respectively. Boxes represent exons, and horizontal lines represent introns; TATA indicates the position of the two promoters, and AATAAA indicates the polyadenylation signal. The numbers below exons 1 to 5 correspond to exon lengths in nucleotides. Numbers on top represent the sizes of the introns between exons 1 and 5. Diagonal lines in the transcripts, MLC1, and MLC3 represent the alternative splicing pathways. The common exons (5 to 9) are spliced in a constitutive manner.

the production of MLC3 mRNA. MLC1/3 exons 3 and 4 are therefore mutually exclusive. The two promoters are expressed exclusively in skeletal muscle and are developmentally regulated. MLC1 is induced at the late embryonic stage, and MLC3 is induced in the early postnatal stage. Both are expressed simultaneously in adult muscle (40).

Three possible mechanisms could account for the developmental and tissue-specific patterns of splicing of MLC1/3 gene. First, the developmental and tissue-specific switches would be easily understood if each pre-mRNA were committed to a specific splicing pattern, dictated by the primary structure of the transcript. Different RNA folding of these two widely different transcripts could dictate exon selection by differentially masking or exposing certain splice sites. Alternatively, the mutually exclusive behavior of MLC exons 3 and 4 could be locally determined by their flanking sequences, with the choice of one or the other being determined by the nature of the 5' donor splice sites in the transcript. Transcripts with the donor of exon 1 would select exon 4, whereas in the absence of this exon, the donor of exon 2 would select exon 3. According to both hypotheses, the splicing pattern would be transcriptionally regulated. Changes in promoter utilization would be sufficient to produce the observed phenotype. A final possibility is the involvement of muscle-specific *trans*-acting factors enforcing the mutually exclusive behavior. In the absence of these factors the pre-mRNAs would be spliced in a constitutive manner. The fact that MLC1 and MLC3 mRNAs are produced by the same cell strongly suggests that the choice between the alternative exons is regulated in *cis* and not dependent on *trans*-acting factors.

To test these hypotheses, we designed a number of minigene constructs in which different combinations of exons 3 and/or 4 were placed in a variety of homologous and heterologous *cis* and *trans* environments. These constructs were placed under the control of the SV40 enhancer-promoter and poly(A) processing signal sequences to allow expression in muscle and nonmuscle cells. After transfection into various cell lines, the splicing patterns of the transcripts were analyzed by S1 nuclease protection of total cytoplasmic RNA. As probes we used synthetic oligonucleotides that contained the proper diagnostic sequences.

Using this approach, we first sought to determine whether

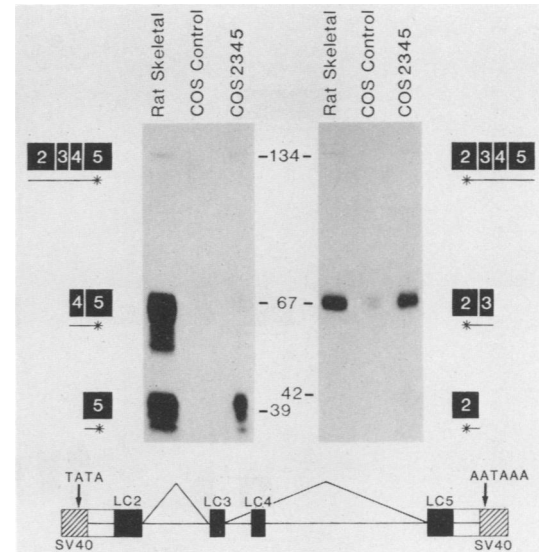


FIG. 2. A minigene construct containing MLC1/3 exons 2, 3, 4, and 5 is spliced according to the MLC3 mRNA pattern. Construct pSVLC2345 and its splicing pattern are diagrammed at the bottom of the figure, using the same symbols as in Fig. 1. In addition, hatched boxes correspond to the SV40 sequences containing the promoter and enhancer sequences. The box with a horizontal line represents MLC intron sequences fused to the SV40. Cytoplasmic RNA was prepared from COS cells (COS control) and after transient expression of pSVLC2345 DNA (COS2345). Rat skeletal RNA was prepared from rat skeletal muscle tissue. The RNAs were analyzed as described in Material and Methods by using a synthetically prepared probe containing sequences corresponding to MLC exons 2, 3, 4, and 5. The position of the ³²P label is indicated by an asterisk. The only RNA detected in the transfected cells contained exons 2, 3, and 5 but not exon 4. No RNAs were detected that lacked exons 3 and 4. The low-intensity band corresponding to the full-length probe represents either undigested probe or RNA containing the two alternatively spliced exons (see the text).

local sequences were sufficient to dictate the mutually exclusive splicing and to ascertain whether muscle-specific factors were required for this behavior. To this end a DNA fragment containing exons 2, 3, 4, and 5 and the intact introns between them was introduced into the SV40 expression vector to generate pSVLC2345. The DNA was transfected into COS cells, and the RNA was analyzed by S1 nuclease protection. A synthetic oligonucleotide probe containing the sequences of the four exons was used to analyze the transcripts. The only mRNA produced in COS cells corresponded to the MLC3 mRNA, in which exons 2, 3, and 5 were spliced together with the exclusion of exon 4 (Fig. 2). No splicing products corresponding to MLC1 mRNA, defined by the splicing together of exons 4 and 5, were produced. The rat skeletal muscle RNA used as a control contained both mRNAs, as expected. No RNAs were detected in which exons 3 and 4 had both been spliced out, either in the muscle or in the transfectants. However, the faint band corresponding to fully protected probe, detected in both muscle and transfectant cell RNA, could be due to the presence of a small amount of RNA containing exons 3 and 4 or to protection of the probe by unspliced pre-mRNA (52).

From this result we conclude that the sequence included in the construct contains the necessary information to specify the mutually exclusive splicing of exons 3 and 4. Sequences downstream from exon 5 are not needed. This process does not require muscle-specific factors and is therefore regulated

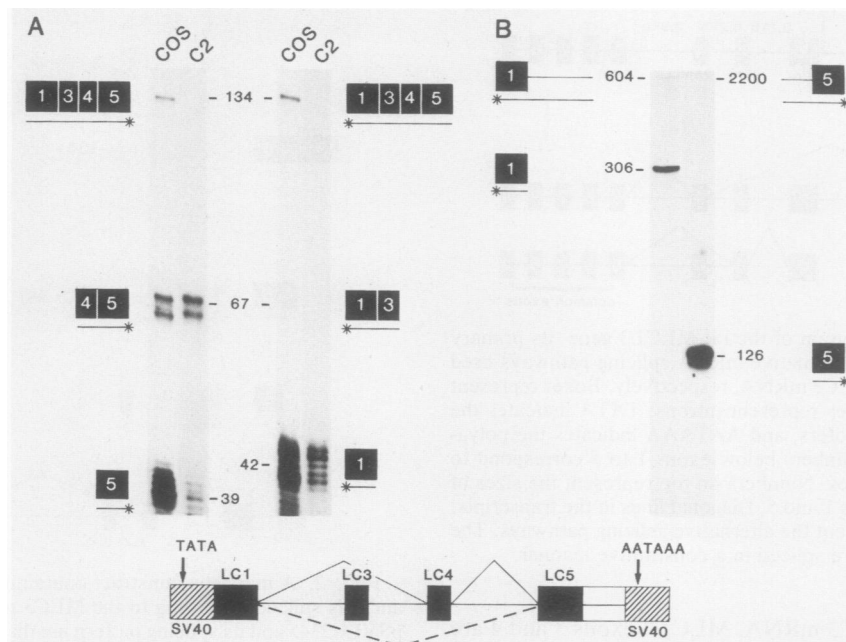


FIG. 3. A minigene construct containing MLC exons 1, 3, 4, and 5 generates mRNA that corresponds to the MLC1 splicing pattern: S1 nuclease analysis of RNAs prepared from COS and C2C12 cells after transient expression of construct pSVLC1345. The probe in panel A is a synthetically made oligonucleotide containing part of MLC exons 1, 3, and 4 and part of exon 5. The other two probes (B) are double-stranded DNA fragments containing exon 1 together with 288 nt of its 3' intron and exon 5 together with 1,219 nt of its 5' intron, respectively. The position of the label and S1 analysis were as in Fig. 2. The diagram at the bottom of the figure shows the structure of the construct and summarizes the splicing patterns. Symbols are as in Fig. 1. The dashed lines represent the smaller percentage of RNAs in which exon 1 is spliced to exon 5 compared with the splicing of exon 1 to exon 4, represented by the solid line. The results obtained with the probes containing intron sequences demonstrated that the lower bands in panel A are not due to the presence of unspliced RNA.

in *cis*. Although COS cells do not normally express the MLC1/3 gene, they are able to splice its transcripts in a manner that is indistinguishable from that of the muscle cells. Moreover, the presence of exon 2, in the absence of exon 1, is sufficient for the inclusion of exon 3 and exclusion of exon 4. This is the pattern used by muscle cells in the production of MLC3 mRNA.

The choice between the mutually exclusive exons 3 and 4 can be determined by the nature of the exon 5' to them. After determining that a truncated transcript with MLC exon 2 at its 5' end spliced according to the MLC3 mRNA pattern, we tested whether the presence of MLC exon 1 and its proximal intron sequences was sufficient to determine the MLC1 mRNA splicing pattern; that is, inclusion of exon 4 and exclusion of exon 3. Exon 2 in pSVLC2345 was replaced by a DNA fragment containing exon 1 and 362 nt of the first intron to generate pSVLC1345 (Fig. 3). As expected from the previous results, together with the MLC1 pre-mRNA pattern of splicing in muscle cells, exon 1 was spliced to exon 4. In contrast to the results obtained with pSVLC2345, no mRNA that included exon 3 was detected in cells transfected with pSVLC1345. Therefore, because the two constructs were identical with the exception of their 5' exons, it is fair to conclude that in this sequence context the nature of the 5' exon determines which of the two mutually exclusive exons will be incorporated into the spliced RNA. In the presence of exon 1, exon 4 was included and exon 3 was excluded, whereas in the presence of exon 2, the pattern was reversed. In addition to the correctly spliced RNA, pSVLC1345 produced an RNA that corresponded to the joining of exon 1 to exon 5, bypassing the two mutually exclusive exons (Fig. 3B). This is not due to the cell type,

since the same RNA was produced in transfected C2C12 myotubes (62), albeit in reduced amounts (Fig. 3A). The same results were obtained with a construct in which exon 5 was replaced by a cDNA fragment containing exons 5, 6, 7, 8, and 9 (data not shown). That these results were not due to protection of the probe by unspliced RNA (52) was further verified by using two DNA probes containing intervening sequences (Fig. 3B). One probe contained exon 1 and 288 nt of its 3' intron, whereas the second contained exon 5 and 1,219 base pairs of its 5' intron.

From these results we can conclude that in nonmuscle as well as in muscle cells the minigene constructs can reproduce the MLC gene pattern of splicing. Furthermore, in this context the selection between the two mutually exclusive exons 3 and 4 is determined by the nature of the common 5' splice site.

To determine whether the results obtained with MLC gene exons 1 and 2 were specific to these exons or part of a more general mechanism of splice site selection, MLC exons 3 and 4 were inserted between heterologous pairs of flanking splice sites (Fig. 4). In pSVIns34, they were inserted into the *EcoRI* site in the second intron of the rat preproinsulin gene (25) and transfected into nonmuscle cells (COS) as well as undifferentiated and differentiated C2C12 myogenic cells (Fig. 4A). In all three cell types, a very significant fraction of the RNA detected corresponded to the splicing of insulin exon 2 to MLC exon 4, with skipping of exon 3 (Fig. 4A). In all cell types, however, a species corresponding to skipping of both exons was detected. In addition, there was also a band corresponding to fully protected probe. Again, this band could represent undigested probe or a small fraction of mRNAs that includes exons 3 and 4 (see below). Identical

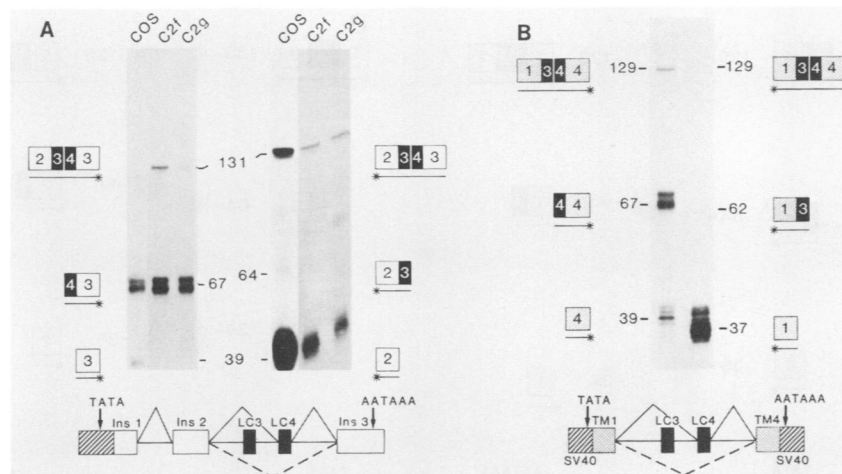


FIG. 4. (A) MLC1/3 exons 3 and 4 are alternatively spliced in a heterologous *cis* environment: S1 nuclease analysis of cytoplasmic RNAs from COS cells, C2 cells, myoblasts (C2g), and myotubes (C2f) transiently transfected with construct pSVIns34. The probe is a synthetically prepared oligonucleotide complementary to 39 nt of insulin exon 2, MLC exons 3 and 4, and 39 nt of insulin exon 3. The sizes of the bands and a diagram of the corresponding probe fragments are shown. A diagram of the constructs and the derived splicing pattern is shown beneath the autoradiogram. Open boxes represent insulin exons, black boxes represent MLC exons, and the striped box represents SV40. The dashed diagonal line represents the splicing pattern of a minor form of the mRNA, in which both MLC exons are excluded. The major pathway represented by the solid diagonal line corresponds to exclusion of MLC exon 3. (B) S1 analysis of RNAs from COS cells transfected with pSVTm34. The probe is a synthetically made oligonucleotide complementary to 37 nt of tropomyosin exon 1, MLC exons 3 and 4, and 39 nt of tropomyosin exon 4. Dotted boxes represent tropomyosin exons, black boxes represent MLC exons, and striped boxes represent SV40.

results were obtained when other muscle (L6E9) (35) and nonmuscle (HeLa) cell lines were used (data not shown). These results are not unique to the insulin gene environment. The same splicing pattern was detected when the alternatively spliced exons were placed between the rat α -TM exons 1 and 4 (Fig. 4B). These exons also flank a pair of mutually exclusive exons in the native α -TM gene (53, 61). Therefore, the mutually exclusive behavior of MLC exons 3 and 4 is not dependent on homologous flanking exons. When placed in either the insulin or the α -TM gene environments, the splicing pattern resembled that of MLC1 mRNA. As described above for the homologous constructs, their behavior is independent of cell type and stage of differentiation. Because COS cells mimic muscle cells in the splicing of the MLC exons 3 and 4 and produce significantly larger quantities of transfected RNA, they were used to test the remainder of the constructs to be described below.

MLC exons 3 and 4 are not intrinsically mutually exclusive.

The results presented above demonstrate that the mutually exclusive behavior of MLC exons 3 and 4 is determined by the local sequences around them and is preserved in a variety of *cis* and *trans* environments. In the case of exons 2 and 3 of the α -TM gene (53, 55), the mutually exclusive behavior is determined by the intrinsic inability of the two exons to be spliced together due to the abnormal location of the branch point between them. To directly test whether MLC exons 3 and 4 are intrinsically unable to splice to each other for similar anatomic reasons, the *in vitro* transcription construct shown in Fig. 5 was prepared. RNA was transcribed *in vitro* by using SP6 polymerase and spliced in a HeLa nuclear extract splicing system (6, 22, 33). The results clearly show the appearance of the expected splicing intermediates and products with the expected kinetics. The nature of the bands labeled in Fig. 5 was further ascertained by their electrophoretic mobility in different concentrations of acrylamide and their response to a debranching reaction. The splicing of the two exons was confirmed by testing an *in vivo* expression construct with similar anatomy after trans-

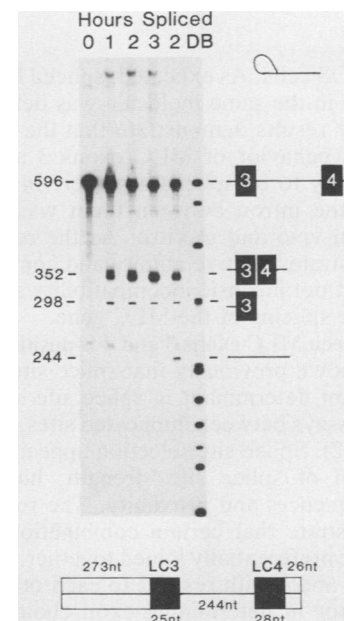


FIG. 5. *In vitro* splicing of MLC1/3 exons 3 and 4. The MLC fragment containing exons 3 and 4 and the intact intron between them was cloned downstream from an SP6 promoter. *In vitro* transcribed RNA was incubated under splicing conditions for the times shown and run directly on an 8% acrylamide denaturing gel or first subjected to a debranching incubation (2DB). The sizes and structures of the observed bands are shown. The intron lariet intermediate is present in low amounts and could not be unambiguously identified in the lanes documenting the time course of the reaction. Unfortunately, it also cannot be identified in the debranched reaction because its size, 298 nt, is identical to the size of the 5' exon. The DNA markers are pBR322 digested with *Hpa*II. At the bottom of the figure, a diagram of the construct with the sizes of exons and introns is shown. Black boxes represent MLC exons, and open boxes with a diagonal line represent MLC intron sequences.

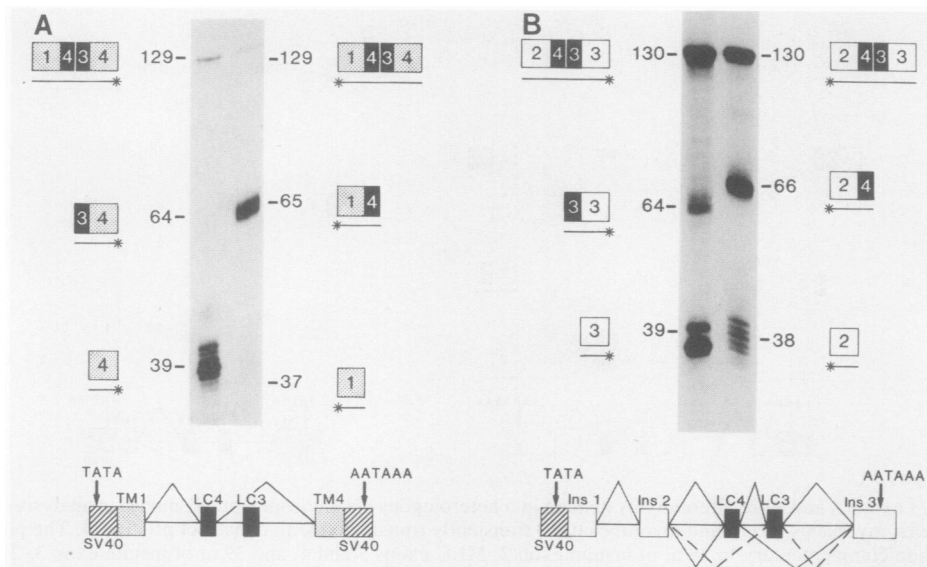


FIG. 6. Analysis of constructs where the position of MLC1/3 exons 3 and 4 has been reversed. (A) S1 nuclease analysis of cytoplasmic RNA from COS cells transiently transfected with pSVTm43. The probe was a synthetically prepared oligonucleotide containing 37 nt of tropomyosin exon 1, MLC exons 4 and 3, and 39 nt of tropomyosin exon 4. (B) Splicing pattern of transcripts from COS cells transfected with pSVIns43. The probe was a synthetically prepared oligonucleotide containing 38 nt of insulin exon 2, MLC exons 4 and 3, and 39 nt of insulin exon 3. A probe containing a tail to distinguish between fully protected and undigested probe was also used (data not shown). The sizes of the fragments in nucleotides and the diagram of the protected probe segment are shown at the sides of the autoradiogram. The diagram at the bottom shows the construct and its splicing pattern. Dotted boxes correspond to tropomyosin exons, open boxes represent insulin exons, black boxes represent MLC exons, and striped boxes represent SV40.

fection into COS cells. As expected, spliced RNA containing exons 3 and 4 in the same molecule was detected (data not shown). These results demonstrate that the basis for mutually exclusive behavior of MLC exons 3 and 4 is not an intrinsic inability to be spliced together. In the absence of competition, the intron between them was effectively removed both *in vivo* and *in vitro*. As the results described below demonstrate, relative affinity and competition among splice sites but not intrinsic incompatibility are the basis for the alternative splicing of the MLC gene.

Choice between MLC exons 3 and 4 is position independent.

It has been shown previously that splice site proximity can be an important determinant in splice site selection. In *cis* competition assays between duplicated sites, the nearest one is preferred (42). Splice site selection appears to result from a combination of splice site strength, homology to the consensus sequence, and proximity. The results presented above demonstrate that certain combinations of 5' and 3' splice sites are preferentially joined together. The position of MLC exons 3 and 4 with respect to each other could be an important factor in determining exon choice. To test this possibility, the positions of these exons in pSVTm34 and pSVIns34 (Fig. 4) were reversed to generate constructs pSVTm43 and pSVIns43, respectively (Fig. 6). pSVTm43 gave rise to a processed RNA that, as in the case of the parental construct (pSVTm34), included exon 4 but excluded exon 3 (Fig. 6A). The relative position of the two exons affected neither their mutually exclusive behavior nor the choice between them. Moreover, it is noteworthy that the integrity of the intron between the two exons was not required for mutual exclusivity.

The results obtained with pSVIns43 were more complex (Fig. 6B). As in pSVIns34, the preferred exon continued to be exon 4, but, in addition, all other possible patterns of splicing were detected. In particular, there was a large

amount of RNA containing both MLC exons 4 and 3. The RNA was further tested by using a probe that contained a nonhybridizing tail, so full protection could be distinguished from nondigested probe (data not shown). This result supported the notion, obtained with pSVIns34, that in the insulin gene context exon 4 is preferred over exon 3. In addition, it demonstrated that when splice sites in a transcript are rearranged among different sets of pairs, their relative strengths can be affected significantly. In this particular construct, most of the mutually exclusive behavior was lost. Since this was not the case for pSVTm43, which contained identical MLC gene sequences, it is clear that the nature of the flanking splice sites plays a role in determining the pattern of splicing.

The mutually exclusive behavior of MLC exons 3 and 4 is due to competition between them for the flanking exons. Exons 3 and 4 can be spliced to each other in the absence but not in the presence of other functional splice sites. The choice between the exons is determined, at least in part, by the flanking exons. Taken together these two observations suggest that there are different levels of compatibilities between splice sites. Different donor sites appear to discriminate between different acceptor sites and vice versa. The exclusion of one alternative exon from the spliced RNAs in the constructs described so far could be due to competition for the flanking common exons. This hypothesis predicts that deletion of the preferred exon from the minigene construct would result in the inclusion of the otherwise excluded exon. To test its validity, two derivatives of pSVIns34 containing only exon 3 (pSVIns3) or exon 4 (pSVIns4) were tested in COS and C2C12 myogenic cells. Exon 4 was efficiently spliced into the RNA when exon 3 had been deleted (Fig. 7). This is the expected result, since exon 4 was also the preferred exon in the presence of exon 3. In contrast, MLC exon 3, which was completely excluded in

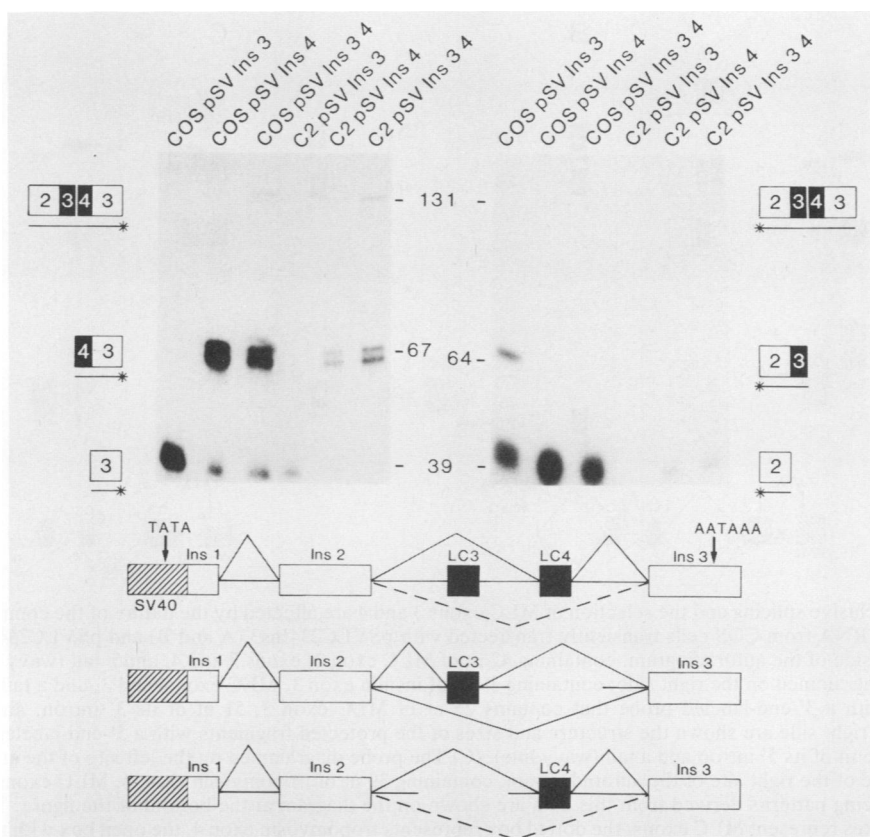


FIG. 7. MLC1/3 exons 3 and 4 compete for the common flanking splice sites. Constructs pSVIns34, pSVIns3, and pSVIns4 were transfected into COS and C2 myogenic cells. Cytoplasmic RNA was prepared after transient expression and analyzed with S1 nuclease. The protected fragments obtained with an oligonucleotide probe complementary to 39 nt of insulin exon 2, MLC exons 3 and 4, and 39 nt of insulin exon 3 were separated in a 6% acrylamide denaturing gel. The diagrams at the bottom of the figure show the splicing pattern corresponding to the three different constructs. Open boxes correspond to insulin exons, black boxes represent MLC exons, and striped boxes represent SV40. The diagonal solid lines represent the major pathways of splicing, and the dashed lines indicate the minor ones.

the presence of exon 4, was spliced in when the competition was deleted. The efficiency of splicing of exon 3, even in the absence of competition, was significantly lower than that of exon 4, because at best it constituted less than 50% of total spliced RNA. This result clearly demonstrates the existence of competition between exons 3 and 4 for the common flanking splice sites. In the context of the insulin gene, exon 4 was preferred and outcompeted exon 3. It is interesting to note, however, that pSVIns3 and pSVIns4 and all of the heterologous constructs tested produced spliced RNAs that contained none of the MLC exons. This result suggests that the compatibility between the two splice sites of the flanking exons for each other is higher than that for either of the MLC exons.

The results presented here indicate that the mutually exclusive splicing of MLC exons 3 and 4 is, at least in part, kinetically determined and due to a competition between the two exons for the common flanking sites. This competition is regulated in *cis*. In the context of the insulin gene, exon 4 was preferred, leading to the exclusion of exon 3, in both muscle and nonmuscle cells. In both cell types, however, exon 3 was recognized at low efficiency, and its inclusion was dramatically increased when exon 4 was deleted. It appears, therefore, that the constitutive splicing machinery is sufficient to govern selection between these two exons.

The pattern of splicing of MLC exons 3 and 4, including their mutually exclusive behavior, is affected by the flanking 3'

and 5' common exons. The results presented above demonstrate that the mutually exclusive behavior of MLC exons 3 and 4 is regulated in *cis*. The choice between the two MLC exons is determined, at least in part, by the nature of the 5' common exon (Fig. 2 and 3). Together with the fact that MLC1 and MLC3 pre-mRNAs differ at their 5' ends, it suggests that this is the site that determines the choice between alternative exons. To further investigate this issue and test the role of the 3' common exon, an additional set of constructs was designed. pSVLC234Ins3 and pSVLC234Tm4 were constructed by replacing exon 5 in pSVLC2345 (Fig. 2) with insulin exon 3 and α -Tm exon 4, respectively (Fig. 8A and B). As previously described, pSVLC2345 was spliced according to the MLC3 mRNA pattern, including exon 3 and excluding exon 4. However, the two constructs with a 3' exon replacement exhibited a very different pattern of splicing. Both constructs produced RNAs containing all possible combinations of MLC exons 3 and 4. RNAs containing only exon 4 and excluding exon 3 represented a small percentage of the total. RNAs including both exons 3 and 4, only exon 3, or neither were each produced at similar levels. Therefore, the mutually exclusive behavior and the pattern of exon choice was dramatically altered by substituting the 3' common exon. To ascertain that the protected bands were produced by spliced RNA and were not an artifact caused by looping out of the intron in unspliced RNA (52), the same RNAs were probed with two

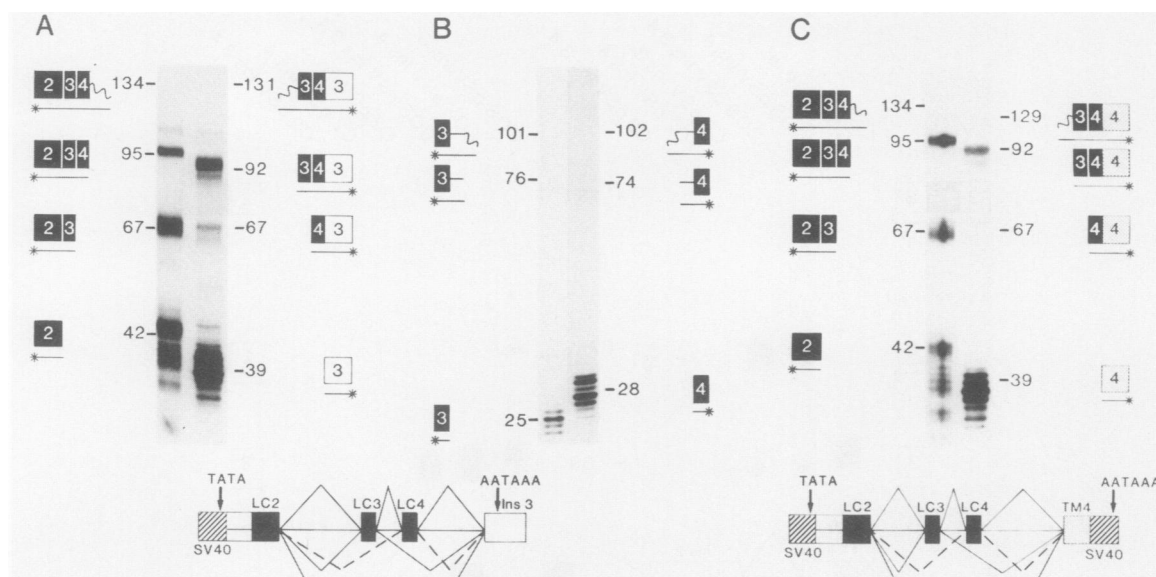


FIG. 8. Mutually exclusive splicing and the selection of MLC exons 3 and 4 are affected by the nature of the common 3' exon: S1 nuclease analysis of cytoplasmic RNA from COS cells transiently transfected with pSVLC234Ins3 (A and B) and pSVLC234Tm4 (C). (A) The probe diagrammed on the left side of the autoradiogram, containing 42 nt of MLC exon 2, exons 3 and 4, and a tail (wavy line), was 3' end labeled (asterisk). The probe diagrammed on the right side, containing 39 nt of insulin exon 3, MLC exon 4 and 3, and a tail, was 5' end labeled. (B) S1 nuclease analysis with a 3'-end-labeled probe that contains 25 nt of MLC exon 3, 51 nt of its 3' intron, and a tail (left side of the autoradiogram). On the right side are shown the structure and sizes of the protected fragments with a 5'-end-labeled probe containing MLC exon 4 sequences and 46 nt of its 5' intron and a tail (wavy line). (C) The probe diagrammed on the left side of the autoradiogram is the same as in panel A. The probe of the right side of the autoradiogram, containing 39 nt of tropomyosin exon 4, MLC exons 4 and 3, and a tail, was 3' end labeled. The splicing patterns derived from this data are shown on the diagram at the bottom of the figure. The open box represents insulin exon 3, black boxes represent MLC exons, the dotted box represents tropomyosin exon 4, the open box with a diagonal line represents MLC intron sequences 5' to exon 2 fused to the SV40 sequences, and striped boxes represent SV40. Dashed diagonal lines represent the minor splicing pathway.

oligonucleotides containing the corresponding exons, part of their intron, and a nonhybridizing tail (Fig. 8C). The S1 nuclease protection products were 28 and 25 nt long with the 5'- or 3'-labeled probe, respectively. This result excluded the presence of unspliced RNA as the source of the protected bands and demonstrated that MLC exon 3 had been spliced to MLC exon 4.

That the 5' common exons are equally important is demonstrated by the results shown in Fig. 3 and the construct shown in Fig. 9. This construct (pSVTm1LC345), like the two previous ones, was also derived from pSVLC2345, but in this case MLC exon 2 was replaced by the α -TM exon 1. In contrast to the result obtained with the parent construct, pSVTm1LC345 resulted in the preferential splicing of MLC exon 4 instead of exon 3, with an increase in the amount of RNA from which both exons have been excluded. No mRNA containing exon 3, either alone or in combination with exon 4, was detected. This result is very similar to that obtained with pSVLC1345. Since these two constructs exhibited similar splicing patterns, the results suggest that in this context the two 5' exons are equivalent but that each one is qualitatively different from MLC exon 2.

The results presented herein demonstrate that the mutually exclusive behavior of MLC exons 3 and 4 and their selection are significantly influenced by the nature of the 5'- and 3'-flanking common exons (Fig. 10). A given 5' or 3' splice site exhibited dramatically different efficiencies of splicing, depending on the nature of its partner. These efficiencies cannot be directly predicted from the strength of the splice site detected in any particular combination. Sites that appeared to be strong—that is, they outcompeted the

others—in one combination were weak—they were outcompeted—in another that differed from the first in a single splice site.

DISCUSSION

The results obtained with the minigene constructs of the MLC1/3 gene reported here demonstrate that alternative splicing of exons 3 and 4 is regulated in *cis*. The nature of the 5' or 3' common exons is sufficient to determine exon choice. Thus, in the natural transcripts, the splicing pattern is in fact transcriptionally determined. Use of the upstream promoter includes exon 1 in the transcript, which is spliced to exon 4, excluding exons 2 and 3. Transcription from the downstream promoter excludes exon 1 from the transcript, and exon 2 is spliced to exon 3 with the exclusion of exon 4. The choice between exons 3 and 4 is determined by the flanking common 5' and 3' splice sites in both homologous and heterologous *cis* and *trans* environments. Using a truncated rat MLC1/3 gene, Garfinkel and Davidson (13) also found that correctly spliced mRNA was produced after transfer of this gene into both myogenic and nonmyogenic cells. Furthermore, there are similarities between MLC and immunoglobulin κ gene splicing. The immunoglobulin κ gene contains several 5' splice sites available for pairing with a common 3' splice site. Only the most 5' exon is joined to the constant region, bypassing all of the 5' splice sites between them in a *cis*-regulated manner that is independent of cell type (20, 21, 26, 27). This gene, therefore, represents a natural example of *cis*-competition assay in which 5' splice sites are available to a single 3' site. The case of MLC exons

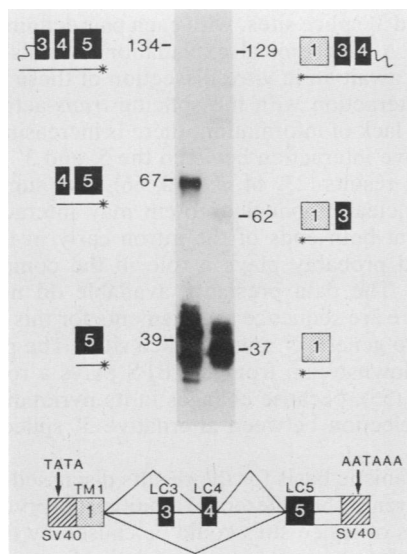


FIG. 9. The selection of MLC exons 3 and 4 is affected by the nature of the common 5' exon. Transfection of COS cells with plasmid pSVTm1LC345. Cytoplasmic RNA was prepared after transient expression and analyzed with S1 nuclease. Two synthetically prepared oligonucleotide probes were used. The probe diagrammed on the left side, containing 39 nt of MLC exon 5, MLC exons 4 and 3, and a tail, was 5' end labeled. The probe diagrammed on the right side, containing 37 nt of tropomyosin exon 1, MLC exons 3 and 4, and a tail, was 3' end labeled. The diagram at the bottom shows the observed pattern of splicing. The dotted box represents tropomyosin exon 1, black boxes represent MLC exons, and the striped boxes represent SV40.

3 and 4, in the context studied here, is more complex because there are several 5' sites available to each 3' site and vice versa.

Contrary to the result obtained with α -TM exons 2 and 3 (53), the mutual exclusivity of the MLC1/3 exons is not due to steric hindrance problems. However, the underlying principle in the regulated splicing of the two genes appears to be similar. It is based on a lower efficiency of splicing between the two splice sites in the intron than between each of them and the common flanking splice sites. This low efficiency is absolute in the case of the α -TM gene and only relative in the MLC1/3 gene. In both cases, the choice of spliced exon is the result of competition between the two exons for the common flanking sites. Surprisingly, for the MLC1/3 gene, this competition does not require the presence of the intact natural intron between the mutually exclusive exons. It persists when the position of the exons is reversed and, therefore, the intron is disrupted. The lack of steric hindrance as a mechanism to enforce the mutually exclusive splicing of two exons is not unique to MLC1/3. The BPSs between exons 6A and 6B of the chicken β -TM (24; E. Brody, personal communication), between exons 6 and 7 of the rat β -TM gene (19), and between exons 7 and 8 of the rat α -TM gene (Sengupta and Nadal-Ginard, unpublished data) have been located far upstream from the acceptor AG. None of these, however, is close enough to the 5' splice site to block splicing. Moreover, except for the rat β -TM, these exons are also able to be spliced together in the absence of other competing splice sites.

The most surprising aspect of the results presented here is that the behavior of a given splice site is not intrinsic to the site but determined in a significant manner by the other

MUTUALLY EXCLUSIVE SPLICING MLC 1/3

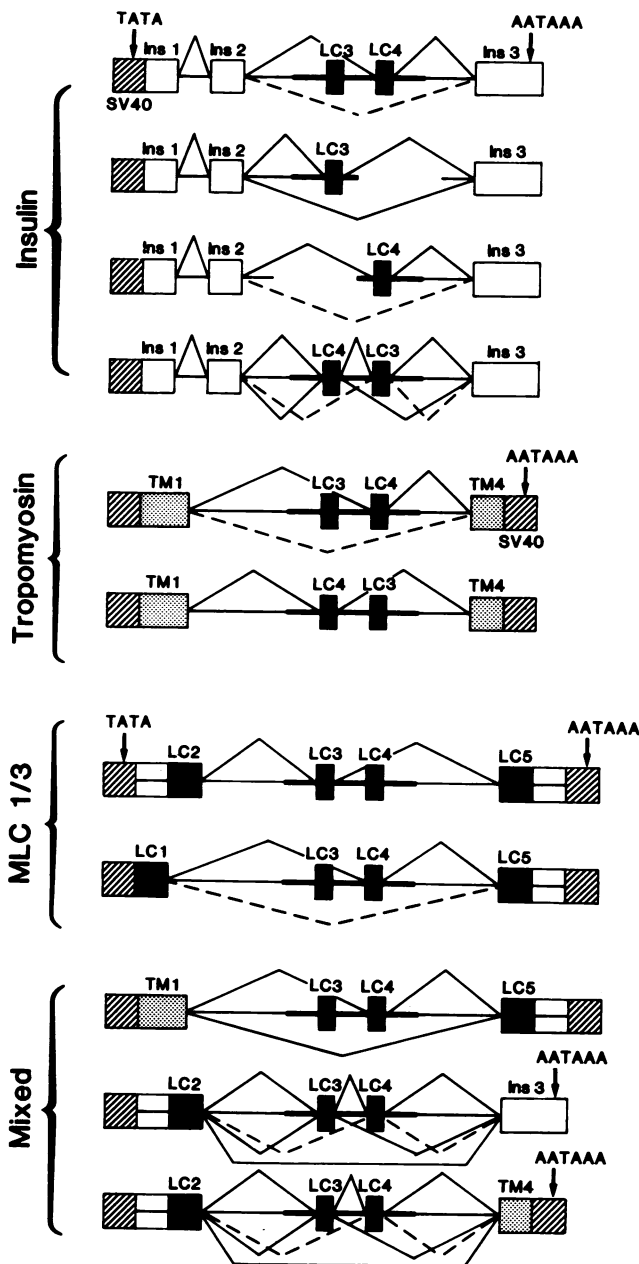


FIG. 10. Summary of the in vivo patterns of splicing observed with the different constructs.

splice sites in the pre-mRNA. Sites that are capable of being spliced to each other in single-intron constructs, such as those between exons 3 and 4, will not do so if other sites are available. Splice sites that are outcompeted in one combination, such as the ones flanking MLC1/3 exon 3 in most of the constructs tested, are the strongest in another that differs from the first by a single splice site, such as exon 3 in the presence of the 5' splice site of MLC exon 2. Thus, changing a single 5' or 3' splice site in these three intron constructs can have a profound effect on the utilization of splice sites located at great distance from them.

During the past several years, significant progress has

been made in defining the sequence requirements of the 5' splice site. The 5' end of the U1 RNA base pairs with the 5' splice site (62). Surprisingly, however, only 3.6% of splice sites have perfect complementarity to U1 (34). Moreover, the best match to the consensus sequence is frequently but not always the site used (79%) (37). These facts are surprising and raise the question of whether a match to the consensus sequence is the main determinant of 5' splice site performance. If this were the case, it would be expected that through evolution most splice sites would have evolved a perfect complementarity to the U1 hybridizing sequence, as is the case in *Saccharomyces cerevisiae*. Yet, a significant body of experimental data indicates that degree of match to the consensus sequence is directly correlated with both efficiency of U1 binding and strength in *cis*-competition assays, as determined by their ability to successfully compete for other sites (7, 32, 36, 42, 65). Sequence context and proximity to the 3' splice site are also important determinants of 5' splice site performance (36, 42). When sites are placed in similar contexts, the site with the highest homology to the consensus sequence and/or most proximal to the 3' end of the intron is the one preferentially used. The results obtained with the MLC1/3 exons can not be easily explained from what is currently known about the determinants of splice site selection. Either there are additional *cis* elements or there is additional information in the known ones that is involved in determining splice site behavior. Homology to the consensus sequence does not explain the results observed. The 5' splice site of exon 3 is the closest to the consensus sequence (eight out of nine residues are homologous). However, it is outcompeted by each of the five 5' exons tested in its ability to splice to exon 4 or to any other exon. A similar observation can be made about the behavior of the 5' and 3' splice sites of exon 4. In all but one of the constructs tested the splice sites of this exon are stronger—that is, are preferred—than those of exon 3. Yet in most constructs this exon is outcompeted by the flanking common exons that have splice sites with less homology to the consensus sequence and are located more distally. In the presence of MLC exon 2, it is even weaker than exon 3. Therefore, there are clear preferences in splice-site pairing. A given 5' splice site will select among several acceptor sites, and vice versa, in a manner that is determined neither by proximity nor by homology to the consensus sequence. These results uncover the existence of unsuspected cooperative interactions between 5' and 3' splice sites that appear to be site specific.

A number of reports have demonstrated that exon sequences other than the 5' and 3' splice sites can affect efficiency (12, 18, 30, 42, 59). Although it is possible that the results reported here are due to sequences other than the canonical *cis* elements known to be involved in intron removal, this seems unlikely for several reasons. First, there is no significant homology between the sequences contained in the DNA fragments used as the 5' exon from α -TM, MLC1, and insulin. These sequences were randomly chosen, yet all behaved in a similar manner when tested with the same downstream sequences. Second, the changes in splicing pattern observed with different 3' terminal exons were not directly correlated with the exon per se but were dependent on the exon at the 5' end. Third, our preliminary results indicate that the behavior of the constructs can be modified by substituting one 5' splice site sequence for another (unpublished observations).

The results obtained with the MLC1/3 constructs are better explained by the existence of a hierarchy of affinities

among 5' and 3' splice sites, with each pair defining a specific interaction. A biochemical explanation for this specificity will have to await an *in vitro* dissection of these splice sites and their interaction with the splicing *trans*-acting factors. Despite this lack of information, there is increasing evidence of cooperative interaction between the 5' and 3' splice sites (60). Recent results (23, 44, 49, 60, 66) have suggested that U1 small nuclear ribonucleoprotein may interact with the splice sites at both ends of the intron early in the splicing reaction and probably plays a role in the commitment of splice sites. The data presently available do not indicate whether there are sequence requirements for this interaction in addition to generic 5' and 3' splice sites. The polypyrimidine track downstream from the BPS plays a role in these interactions (55), because changes in its pyrimidine content affect the selection between alternative 5' splice sites (11, 59).

The mechanistic basis for the results discussed here is not readily apparent. The differential affinities observed between specific pairs of splice sites could be caused by the variable nature of the interactions between splicing factors and splice sites, which in turn could be due to the loose conservation of the consensus splicing elements. The formation of spliceosomes involving particular pairs of splice sites may therefore be the result of higher-affinity interactions between the splicing factors initially bound at these pairs of sites. The basis for such cooperative interactions between specific splice site pairs remains to be determined. However, it should be remembered that our understanding of the basic mechanisms of splicing has been derived mainly from simple pre-mRNA substrates, in which the kind of complex interactions observed here would not have been detected.

ACKNOWLEDGMENTS

We thank Steve Fish for technical assistance, Michelle Emfinger for help with tissue culture work, and Emily Flynn for the graphic artwork. We are grateful to Christopher Smith, James Patton, and Dola Sengupta for careful and critical reading of the manuscript as well as for helpful advice and discussions.

This work was supported in part by grants from the National Institutes of Health and from the Muscular Dystrophy Association of America. M.E.G. was supported by a postdoctoral fellowship from Consejo Superior de Investigaciones Científicas of Spain.

LITERATURE CITED

1. Aebi, M., H. Hornig, and C. Weissmann. 1987. Cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved GU. *Cell* 50:237-246.
2. Andreadis, A., M. E. Gallego, and B. Nadal-Ginard. 1987. Generation of protein isoform diversity by alternative splicing. *Annu. Rev. Cell Biol.* 3:207-242.
3. Berk, A. G., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12:721-732.
4. Breitbart, R. E., A. Andreadis, and B. Nadal-Ginard. 1987. Alternative splicing; a ubiquitous mechanism for the generation of multiple protein isoforms from single genes. *Annu. Rev. Biochem.* 56:467-495.
5. Brody, E., and J. Abelson. 1985. The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing dependent reaction. *Science* 228:963-967.
6. Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11:1475-1489.
7. Eperon, L. P., J. P. Estibeiro, and I. C. Eperon. 1986. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. *Nature (London)* 324:280-282.
8. Eperon, L. P., I. R. Graham, A. D. Griffiths, and I. C. Eperon.

1988. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell* **54**:393-401.
9. Favoloro, J., R. Freisman, and R. Kamen. 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. *Methods Enzymol.* **65**:718-749.
 10. Frendeway, D., and W. Keller. 1985. The stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. *Cell* **42**:355-367.
 11. Fu, X., H. Ge, and J. L. Manley. 1988. The role of the polypyrimidine stretch at the SV40 early pre-mRNA 3' splice site in alternative splicing. *EMBO J.* **7**:809-817.
 12. Furdon, P. J., and R. Kole. 1986. Inhibition of splicing but not cleavage at the 5' splice site by truncating human β -globin pre-mRNA. *Proc. Natl. Acad. Sci. USA* **83**:927-931.
 13. Garfinkel, L. I., and N. Davidson. 1987. Developmentally regulated expression of a truncated myosin light-chain 1/3 gene. *Mol. Cell. Biol.* **7**:3826-3829.
 14. Gluzman, Y. 1981. SV40 transformed simian cells support the replication of early SV40 mutants. *Cell* **23**:175-182.
 15. Grabowski, P. J., S. R. Sellar, and P. A. Sharp. 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. *Cell* **42**:345-353.
 16. Graham, F. L., and A. J. van der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**:456-467.
 17. Green, M. R. 1986. Pre-mRNA splicing. *Annu. Rev. Genet.* **20**:671-708.
 18. Hampson, K. R., L. La Follette, and M. F. Rottman. 1989. Alternative processing of bovine growth hormone mRNA is influenced by downstream exon sequences. *Mol. Cell. Biol.* **9**:1604-1610.
 19. Helfman, D., and W. M. Ricci. 1985. Branch point selection in alternative splicing of tropomyosin pre-mRNAs. *Nucleic Acids Res.* **17**:5633-5650.
 20. Kedes, D. H., and J. A. Steitz. 1987. Accurate 5' splice-site selection in mouse κ immunoglobulin light chain premessenger RNAs is not cell type specific. *Proc. Natl. Acad. Sci. USA* **84**:7928-7932.
 21. Kedes, D. H., and J. A. Steitz. 1988. Correct in vivo splicing of the mouse immunoglobulin κ light chain pre-mRNA is dependent on 5' splice-site position even in the absence of transcription. *Genes Dev.* **2**:1448-1459.
 22. Krainer, A. R., T. Maniatis, B. Ruskin, and M. R. Green. 1984. Normal and mutant human β -globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell* **36**:993-1005.
 23. Lamond, A. I., M. M. Konarska, and P. A. Sharp. 1987. A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. *Genes Dev.* **1**:532-543.
 24. Libri, D., J. Marie, E. Brody, and Y. M. Fisman. 1989. A subfragment of the β tropomyosin gene is alternatively spliced when transfected into differentiating muscle cells. *Nucleic Acids Res.* **17**:6449-6462.
 25. Lomedico, P. T. 1982. Use of recombinant DNA technology to program eukaryotic cells to synthesize rat proinsulin: a rapid expression assay for cloned genes. *Proc. Natl. Acad. Sci. USA* **79**:5798-5802.
 26. Lowery, D. E., and B. G. Van Ness. 1987. In vitro splicing of kappa immunoglobulin precursor mRNA. *Mol. Cell. Biol.* **7**:1346-1351.
 27. Lowery, D. E., and B. G. Van Ness. 1988. Comparison of in vitro and in vivo splice site selection in κ -immunoglobulin precursor mRNA. *Mol. Cell. Biol.* **8**:2610-2619.
 28. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 29. Maniatis, T., and R. Reed. 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature (London)* **325**:673-678.
 30. Mardon, H. J., G. Sebastio, and F. E. Baralle. 1987. A role for exon sequences in alternative splicing of the human fibronectin gene. *Nucleic Acids Res.* **15**:7725-7733.
 31. Mattaj, W. I. 1989. A binding consensus: RNA-protein interactions in splicing, snRNPs, and sex. *Cell* **57**:1-3.
 32. Mayeda, A., and Y. Oshima. 1988. Short donor site sequences inserted within the intron of β -globin pre-mRNA serve for splicing in vitro. *Mol. Cell. Biol.* **8**:4484-4491.
 33. Melton, D. A., P. A. Krieg, M. R. Reblaggiati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**:7035-7056.
 34. Mount, S. M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**:459-472.
 35. Nadal-Ginard, B. 1978. Commitment, fusion, and biochemical differentiation of a myogenic cell line in the absence of DNA synthesis. *Cell* **15**:855-864.
 36. Nelson, K. K., and M. R. Green. 1988. Splice site selection and ribonucleoprotein complex assembly during in vitro pre-mRNA splicing. *Genes Dev.* **2**:319-329.
 37. Oshima, Y., and Y. Gotoh. 1987. Signals for the selection of a splice site in pre-mRNA. *J. Mol. Biol.* **195**:247-259.
 38. Padgett, R. A., P. J. Grobowski, M. M. Konarska, S. Seiler, and P. A. Sharp. 1986. Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* **55**:1119-1150.
 39. Parker, B. A., and G. R. Stark. 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA present early in infections by virus or virus DNA. *J. Virol.* **31**:360-369.
 40. Periasamy, M., E. E. Strehler, L. I. Garfinkel, R. M. Gubits, N. Ruiz-Opazo, and B. Nadal-Ginard. 1984. Fast skeletal muscle myosin light chains 1 and 3 are produced from a single gene by a combined process of differential RNA transcription and splicing. *J. Biol. Chem.* **259**:13595-13604.
 41. Reed, R., and T. Maniatis. 1985. Intron sequences involved in lariat formation during pre-mRNA splicing. *Cell* **41**:95-105.
 42. Reed, R., and T. Maniatis. 1986. A role for exon sequences and splice site proximity in splice site selection. *Cell* **46**:681-690.
 43. Reed, R., and T. Maniatis. 1988. The role of the mammalian branchpoint sequence in pre-mRNA splicing. *Genes Dev.* **2**:1268-1276.
 44. Ruby, W. S., and J. Abelson. 1988. An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* **242**:1028-1035.
 45. Ruiz-Opazo, N., and B. Nadal-Ginard. 1987. Tropomyosin gene organization. *J. Biol. Chem.* **261**:4755-4765.
 46. Ruskin, B., and M. R. Green. 1985. An RNA processing activity that debranches RNA lariats. *Science* **229**:135-140.
 47. Ruskin, B., A. R. Krainer, T. Maniatis, and M. R. Green. 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell* **38**:317-331.
 48. Scherer, W. F., J. T. Syverton, and G. O. Gey. 1953. Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. *J. Exp. Med.* **97**:695-710.
 49. Seraphin, B., and M. Rosbash. 1989. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* **59**:349-358.
 50. Shapiro, M. B., and P. Senapathy. 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* **15**:7155-7174.
 51. Sharp, P. 1987. Splicing of messenger RNA precursors. *Science* **235**:766-771.
 52. Sisodia, S. S., W. D. Cleveland, and B. Sollner-Webb. 1987. A combination of RNase H and S1 nuclease circumvents an artefact inherent to conventional S1 analysis of RNA splicing. *Nucleic Acids Res.* **15**:1995-2011.
 53. Smith, C. W. J., and B. Nadal-Ginard. 1989. Mutually exclusive splicing of α -tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. *Cell* **56**:749-758.
 54. Smith, C. W. J., G. J. Patton, and B. Nadal-Ginard. 1989. Alternative splicing in the control of gene expression. *Annu. Rev. Genet.* **23**:527-577.

55. **Smith, C. W. J., E. B. Porro, J. G. Patton, and B. Nadal-Ginard.** 1989. The 3' splice site of mammalian introns is defined by scanning from an independently specified branch point. *Nature (London)* **342**:243-247.
56. **Solnick, D.** 1985. Alternative RNA splicing caused by RNA secondary structure. *Cell* **43**:667-676.
57. **Solnick, D., and S. I. Lee.** 1987. Amount of RNA secondary structure required to induce an alternative splice. *Mol. Cell. Biol.* **7**:3194-3198.
58. **Strehler, E. E., M. Periasamy, M.-E. Strehler-Page, and B. Nadal-Ginard.** 1985. Myosin light chain 1 and 3 gene has two structurally distinct and differentially regulated promoters evolving at different rates. *Mol. Cell. Biol.* **5**:3168-3182.
59. **Streuli, M., and H. Saito.** 1989. Regulation of tissue specific alternative splicing: exon specific cis-elements govern the splicing of leukocyte common antigen pre-mRNA. *EMBO J.* **8**:787-796.
60. **Ulfendahl, J. P., J. Kreivi, and G. Akusjarvi.** 1989. Role of the branch/3'-splice site region in adenovirus-2 E1A pre-mRNA alternative splicing: evidence for 5'- and 3'-splice site cooperation. *Nucleic Acids Res.* **17**:925-938.
61. **Wieczoreck, D., C. W. J. Smith, and B. Nadal-Ginard.** 1988. The rat α -tropomyosin gene generates a minimum of six different mRNAs coding for striated, smooth, and nonmuscle isoforms by alternative splicing. *Mol. Cell. Biol.* **8**:679-694.
62. **Yaffe, D., and O. Saxel.** 1977. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature (London)* **270**:725-727.
63. **Zhuang, Y., M. A. Goldstein, and M. A. Weiner.** 1989. UAC UAAC is the preferred branch site for mammalian mRNA splicing. *Proc. Natl. Acad. Sci. USA* **86**:2752-2756.
64. **Zhuang, Y., H. Leung, and A. M. Weiner.** 1987. The natural 5' splice site of simian virus 40 large T antigen can be improved by increasing the base complementarity to U1 RNA. *Mol. Cell. Biol.* **7**:3018-3020.
65. **Zhuang, Y., and A. M. Weiner.** 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* **46**:827-835.
66. **Zillman, M., S. D. Rose, and S. M. Berget.** 1987. U1 small nuclear ribonucleoproteins are required early during spliceosome assembly. *Mol. Cell. Biol.* **7**:2877-2883.