Pre-mRNA Secondary Structure And The Regulation of Splicing

Laurent Balvay, Domenico Libri and Marc Y. Fiszman*

Summary

Nuclear pre-mRNAs must be precisely processed to give rise to mature cytoplasmic mRNAs. This maturation process, known as splicing, involves excision of intron sequences and ligation of the exon sequences. One of the major problems in understanding this process is how splice sites, the sequences which form the boundaries between introns and exons, can be accurately selected. A number of studies have defined conserved sequences within introns which were later shown to interact with small nuclear ribonucleoproteins (snRNPs). However, due to the simplicity of these conserved sequences it has become clear that other elements must be involved and a number of studies have indicated the importance of secondary structures within pre-mRNAs. Using various examples, we shall show that such structures can help to specify splice sites by modifying physical distances within introns or by being involved in the definition of exons and, lastly, that they can be part of the regulation of alternative splicing.

Introduction

Most eukaryote genes are partitioned into intron and exon sequences which upon transcription are present on the nuclear premessenger RNAs. Intron sequences must then be precisely excised by RNA splicing to give rise to mature cytoplasmic transcripts. In some cases, pre-mRNAs are alternatively spliced depending on the state of differentiation or the cell type. This type of regulated splicing will lead to the synthesis of different proteins from the same premessenger RNA or can function as an on-off switch by generating RNAs which either lack or possess an open reading frame. This latter case of regulated splicing is at the basis for sex determination in Drosophila[1]. The number of reports on premessenger RNAs which are subjected to regulated splicing has been steadily increasing[2,3]. Understanding how splicing is regulated is, therefore, becoming of fundamental importance.

Pre-mRNA splicing occurs by a two step mechanism which involves transesterification reactions (for review see ref 4). In the first step, the pre-mRNA is cleaved at the 5′ splice site and the phosphorylated guanosine residue at the 5′ end of the intron is covalently joined to the 2′ hydroxyl group of an adenine residue located near the 3′ end of the intron, which is known as the branch site. This step leads to a splicing intermediate in which the intron and the downstream exon are in a lariat configuration. In the second step, the 3′ splice site is cleaved at the same time as the two exons are ligated, thereby generating the spliced mRNA and the intron in a lariat configuration. This lariat is further linearized and subsequently degraded (Fig. 1). Three short sequences are essential, though not sufficient, for the intron to be accurately and efficiently spliced. These are: 1) the 5′ splice site whose consensus sequence is C/AAG-GUA/GAGU, 2) the 3′ splice site whose sequence is $(Y)_nNYAG$-G and 3) the branch point sequence (BPS) which can be more variable but is usually of the kind UNYURAY (N stands for any nucleotide, Y stands for a pyrimidine and R for a purine) and which is located 15-40nt upstream of the 3′ splice site.

The splicing reaction takes place within a large multicomponent complex, designated the spliceosome[5]. These spliceosomes contain the pre-mRNA associated with small nuclear ribonucleoprotein particles (snRNP's) U1, U2, U4, U5 and U6 and non-snRNP splicing factors as well as other protein factors. The function of some spliceosome components have been determined. These include the U1snRNP which binds to the 5′ splice site and the U2snRNP which binds to the BPS; in both cases, this binding involves the formation of complementary base pairs between snRNA and pre-mRNA (for review see refs 4, 6, 7).

Secondary Structure and Splice Site Selection

A central question in the study of both basic splicing and regulated alternative splicing is how specific pairs of 5′ and 3′ splice sites are chosen. The mechanism of selection must involve the establishment of productive factor/splice-site interactions during the process of spliceosome assembly. Accordingly, mutation in the conserved 5′ and 3′ splice site consensus elements will abolish splicing and can lead to the selection of cryptic splice sites. However, these consensus elements are insufficient to specify a splice site unambiguously and it is becoming clear that sequences flanking splice sites contribute to selection. Splice site context could reflect either primary sequence or higher order RNA structure. We will analyze two examples which illustrate how RNA secondary structure can affect splice site selection.

Secondary structure can allow selection of an abnormally located 3′ splice site

Among the various transcripts which arise from the adenovirus E1A transcription unit, there is one which involves the excision of a 216nt intron[8]. One of the peculiarities of this reaction is that it uses as the branch point three A residues located 51, 55 and 59nt upstream of the 3′ splice site. These distances can be considered unusually long with respect to the known consensus distance which is between 18 and 37nt and which has been proposed to be a strong constraint for the splicing machinery. The prediction of RNA folding within the region between the three A's and the 3′ splice site revealed a potential stable hairpin structure the existence of which was demonstrated using site directed mutagenesis. Mutations which stabilize the hairpin structure significantly

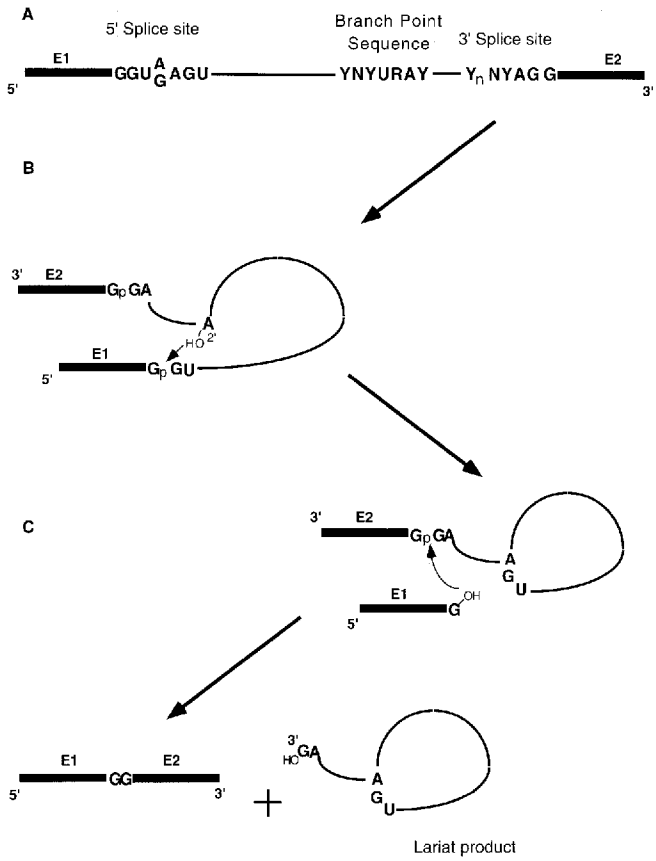

Fig. 1. Splicing of pre-mRNA of higher eukaryotes. A. Structure of a one intron containing pre-mRNA; the consensus sequences of the 5′, 3′ and branch region are indicated. B. The 2′ OH of the branch point nucleotide (A residue) interacts with the 5′ phosphoryl end of the first intron nucleotide (G residue), cleaving off the first exon and forming a lariat intermediate. C. The 3′ OH of the last nucleotide of the intron interacts with the 5′ phosphoryl end of the last nucleotide of exon 1, giving rise to the mature mRNA and the intron as a lariat.

increase the branching efficiency whereas mutations which disrupt the structure abolish the splicing reaction. Lastly, compensatory mutations which restore the secondary structure also restore splicing efficiency[9]. Thus, the formation of a hairpin structure may bring the branch acceptors to an operational distance of 31, 35 and 39nt upstream of the 3′ splice site (Fig. 2).

3′ splice site selection and the rule of the first come, first served

The choice of 3′ splice sites is believed to be dictated by the position of branch formation, the first encountered AG being selected by a 5′-3′ molecular scanning mechanism[10]. Recently, Deshler and his colleagues described a rather unusual situation in the actin gene of *Kluyveromyces lactis*[11]. This gene contains an intron in which a YAG (YAG1), located 73nt downstream of the branch point, is skipped, and a second YAG (YAG2), positioned 122nt downstream of the branch point is chosen as a 3′ splice site. This poses two questions: 1) why is the first AG skipped? and 2) how is the distal AG chosen in spite of the fact that its distance from the branch point had been shown to be inhibitory for the splicing reaction[12]? To answer both questions this group has engineered a genetic screen by fusing the Escherichia coli lacZ gene downstream and in the reading frame of the first AG. This allowed them to monitor the use of this splice site simply by observing colony colour on X-gal-containing plates. Many of the randomly induced cis-acting mutations that activate the silent YAG1 fall within a region surrounding this YAG1 which can hypothetically fold into a hairpin structure. This hypothetical structure was directly proven by site-directed mutagenesis.This secondary structure not only decreases the physical distance between the branch point and YAG2 as in the previous example, but also sequesters the first YAG[13] (Fig. 3). It is interesting to note that this secondary structure seems to be efficient by itself and does not require specific factor(s) to be stabilized since the K. lactis actin gene is properly spliced in S. cerevisiae whose actin intron 3′ end does not include such structure[11].

With these two examples, we can see how RNA folding can overcome a distance problem between a branch point and a 3′ splice site and also explains some exceptions to the scanning mechanism which is supposed to control the choice of 3′ splice site.

Secondary structure and the mechanism of exon recognition

When addressing the problem of splicing of pre-mRNA molecules the question arises as what has to be recognized: sequences in the intron or in the exon? Berget and colleagues have proposed a model in which the unit of assembly for splicing factors is the exon rather than the intron[14].

A careful analysis of the splice sites consensus sequences indicates that this consensus masks a large diversity of indi-

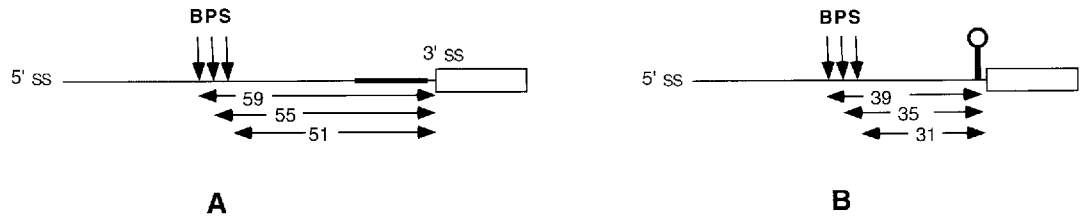

Fig. 2. A secondary structure which decreases the distance between a branch site and the 3′ splice site. A. Linear representation of the 216nt intron of the E1A transcription unit. The numbers indicate the distance in nucleotides between the 3 branch sites and the 3′ end. B. Representation of the same intron with the secondary structure at the 3′ end

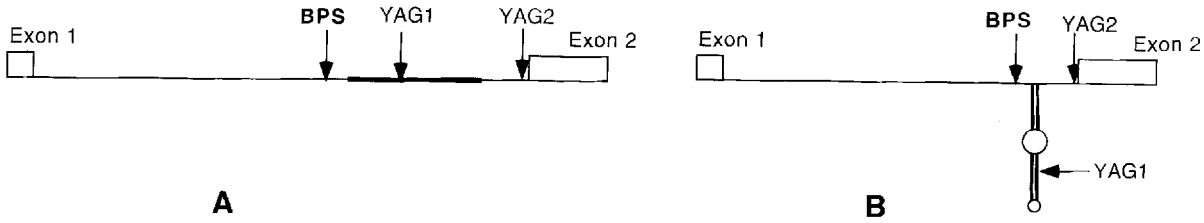

Fig. 3. A secondary structure which allows skipping of a potential 3′ splice site. A. Linear representation of the actin intron of K. Lactis with the position of the two potential 3′ splice sites. B. Representation of the secondary structure with its sequestering effect on the first 3′ splice site.

vidual sequences. For example, from the donor site sequence diversity, it can be predicted that potential donor sites will be present fairly often along the pre-mRNA molecule. One question which arises is thus how does the cell achieve selection between the natural and the cryptic sites? One obvious parameter is the intrinsic strength of the site which reflects its capacity to hybridize with U1snRNA, although recent results also indicate that U5snRNP is also involved in the selection of 5′ and 3′ splice sites[15]. However, this is not sufficient since, in the absence of any other element, all potential splice sites would be used according to their strength, which, obviously, is not the case for cryptic sites. This points to a second parameter which is the sequence context of the splice site and particularly the exon sequences. Indeed, a number of results have shown that scattered exon mutations can interfere with splicing of these exons. Most of these mutations do not allow a general interpretation to be drawn. However, there are some examples in which these mutations have been shown to be part of a secondary structure which seems to be important for the selection of the natural splice site. One such case has been described by Jacob and colleagues[16]. In the E3 pre-mRNA of Adenovirus 2, there is a cryptic donor site which is located 74nt downstream of the natural site. This group demonstrated that the use of the natural site is controlled by exon sequences which are part of a stable stem-loop structure. Exon mutations which retain this structure also maintain selection of the natural site, while mutations which disrupt this structure lead to the use of both sites according to their strength. Also, it should be mentioned that this structure is maintained in the same region of another adenovirus, Adenovirus 5.

More recently, a series of mutations which affect splicing in the human hypoxanthine-guanine phosphoribosyltransferase gene has been analyzed[17]. Some of these mutations, which were found to be in exon 8, resulted in skipping of this exon. To explain this result, the authors postulated the existence of a secondary structure which would be disrupted by some of these mutations.

Although some of these suggestions remain to be proven, it seems rather likely that some exon sequences are involved in the formation of secondary structures which could interact with trans acting factors.

Secondary Structures and Regulated Alternative Splicing

Secondary structures have often been invoked as a possible element for splicing regulation[18,19]. Using model substrates in which an exon was flanked by inverted repeats within introns, it was found that the exon within the loop of the potential structure was omitted during splicing *in vitro*, allowing flanking exons to be spliced together[20]. However, *in vivo*, the effect was much less significant[20]. A systematic analysis of the importance of the length of the complementary sequences revealed that, in vivo, these had to exceed 50nt before significant effects could be detected[21]. A stem

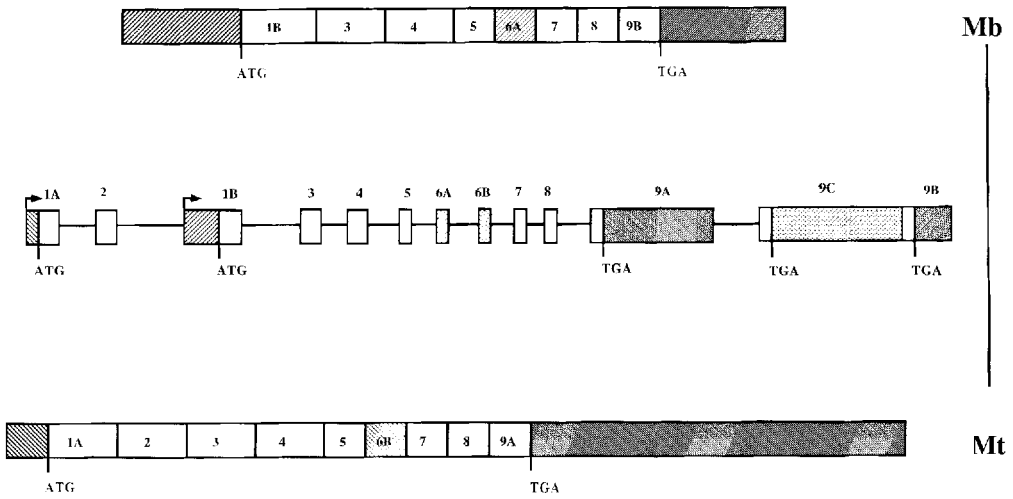

Fig. 4. Structure of the chicken β tropomyosin gene and of transcripts which it generates during the myoblast to myotube transition.

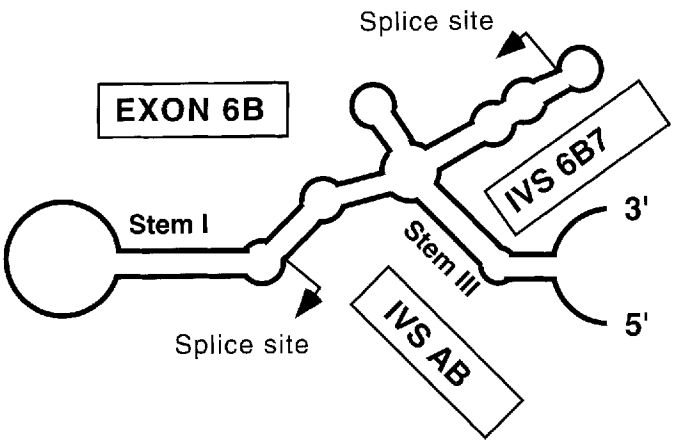

Fig. 5. Schematic representation of the secondary structure around exon 6B. Exon 6B is delineated by the two arrows labelled splice site. IVS B7 indicates the intron between exons 6B and 7; IVS AB indicates the intron between exons 6A and 6B. Stem I and stem III indicate two potential secondary structure which have been tested *in vivo* and *in vitro*.

of such great stability is almost never found in natural pre-mRNAs and it was concluded that the secondary structures of natural mRNA precursors are unlikely to affect splice site selection. However, quite different results were obtained when a potential short stem (17nt) included one of two alternative 5′ splice sites. In this case, all the mRNA which were detected *in vivo* used the unsequestered site[22]. The main difference between the two systems is the size of the loop. In the first case the loop length was 285nt while in the second, the loop length was much smaller. This parameter was analyzed in greater details and it was found that, as long as the loop length is smaller than 50-55nt, any small stem will interfere with splice site selection. With, longer loop length, longer stem structures are required[23]. One interpretation of these results is that following transcription, there is a 'window' within which the RNA is free to fold; the size of this window would be defined by the time taken for protein complexes to bind to the nascent transcript[23]. If such a model is correct, there could be a link between the rate of transcription and the control of alternative splicing.

The β tropomyosin gene

For a number of years, our laboratory has been interested in the regulation of expression of the gene which codes for the β tropomyosin in the chicken. This gene is interesting in several aspects. By the use of two promoters, alternative terminal exons and two mutually exclusive internal exons, this gene codes for three tropomyosin isoforms, two of which are expressed in adult muscle tissues (smooth or skeletal), while the third is expressed in nonmuscle cells and undifferentiated myoblasts[24-26]. The two internal exons (exons 6A and 6B) have a strict tissue specificity: exon 6A is present only in mRNA of smooth muscle, non muscle cells and undifferentiated myoblasts, while exon 6B is used exclusively in skeletal muscle and differentiated myotubes in culture. Similarly, exon 9A and 9B (the 3′-terminal exons) are expressed specifically in skeletal muscle (exon 9A) or in all the other cell types (exon 9B). Recently, a third terminal exon (9C) has been described which is specifically used in brain[27]. Two different transcription initiation sites are used according to whether the gene is expressed in adult smooth and skeletal muscle (exon 1A) or in undifferentiated myoblasts and non muscle cells (exon 1B). The gene organization and its transcripts are shown in Fig 4. The cognate rat gene is very similar to the chicken gene with the exception that the internal promoter is missing and thus, this gene only codes for two tropomyosin isoforms[28]. Using minigene transfection of cultured myogenic cells, we showed that all the necessary information for the tissue specificity as well as the mutually exclusive character of exons 6A and 6B is contained within a 0.9kb genomic fragment containing the two alternative exons flanked by two constitutive exons[29]. Mutagenesis of regions around the splice acceptor site of exon 6B allowed us to identify negatively acting cis-elements which prevent splicing of the muscle specific exon 6B in undifferentiated myoblasts. The 5′ half of the exon and the upstream pyrimidine rich region located between the branch point (position −105) and the AG acceptor contain these negative elements[30,31]. A computer analysis of the sequence from exon 5 to exon 7 led us to propose that a secondary structure could be involved in this regulation[31]; this secondary structure was probed *in vitro* using a combination of enzymatic and chemical methods[32]. Figure 4 is a very schematic representation of this secondary structure around exon 6B. The role of each potential stem was tested by introduction of mutations which should disrupt the structure and compensatory mutations which should allow its reformation. The only stem structure which was unambiguously shown to play a role is stem I. In this case, any change on either one of the strands leads to the use of exon 6B in transfected myoblasts while compensatory mutations on the two strands restore the negative regulation on the use of exon 6B[33]. A further indication that this stem is important for the regulation is the fact that it is conserved in different species (rat,mouse and xenopus) but not in exon 6B of the chicken α tropomyosin gene in spite of the high conservation of the protein sequence in this region. This last point was expected because of the different expression pattern of exon 6B in the α tropomyosin gene[34].

When the other stem structures were tested, the only mutations which were able to disrupt the control were all located in the IVS AB and none could be found either in the 3′ half of exon 6B or in the downstream intron (IVS B7)[35]. These experiments would argue against the hypothesis that the large secondary structure,with the exception of stem I, is involved in the control of alternative splicing of exon 6B[36]. It would appear that the sequences in IVS AB which have been defined as regulatory elements must inhibit splicing of exon 6B not by forming a particular secondary structure, but probably by interacting with some trans acting factors, as has been proposed for the rat β tropomyosin gene[37].

Conclusions

In conclusion, RNA secondary structure has been implicated

in a variety of regulatory mechanisms such as transcriptional attenuation[38-40] and initiation of translation[41]. As far as splicing is concerned, such secondary structures are very important for the splicing of type I and II introns[42] as well as for trans-splicing in plants[43]. As far as nuclear pre-mRNAs are concerned, it is becoming clear that RNA folding participates in the regulation of splicing either by specifying splice sites[16] or by interfering with splice site selection[9] or as part of negatively cis-acting elements[33]. In most cases, however, the implication that secondary structures are involved derives from the demonstration of their existence in solution and on experiments using *in vitro* splicing. It is important to stress that in the absence of *in vivo* experiments or of *in vitro* systems where transcription and splicing are coupled, all these conclusions about the functional significance of secondary structures should be taken as tentative ones.

Acknowledgements

The work reported in this review has been supported by grants from the AFM, ARC, CNRS, INSERM, Fondation Médicale pour la Recherche.

References

1 Baker, B. S. (1989). Sex in flies: the splice of life. *Nature* **340**, 521-524.
2 Smith, C. W. J., Patton, J. G. and Nadal-Ginard, B. (1989). Alternative splicing in the control of gene expression. *Annu. Rev. Genet.* **23**, 527-577.
3 Andreadis, A., Gallego, M. E. and Nadal-Ginard, B. (1987). Generation of protein isoform diversity by alternative splicing: mechanistic and biological implications. *Annu. Rev. Cell. Biol.* **3**, 207-242.
4 Green, M. R. (1991). Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu. Rev. Cell Biol.* **7**, 559-599.
5 Brody, E. and Abelson, J. (1985). The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* **288**, 963-967.
6 Ruby, S. W. and Abelson, J. (1991). Pre-mRNA splicing in yeast. *Trends Genet.* **7**, 79-85.
7 Rosbash, M. and Seraphin, B. (1991). Who's on first? The U1 snRNP-5′ splice site interaction and splicing. *Trends Biochem. Sci.* **16**, 187-190.
8 Gattoni, R., Schmitt, P. and Stevenin, J. (1988). In vitro splicing of adenovirus E1A transcripts: characterization of novel reactions and of multiple branch points abnormally far from the 3′ splice site. *Nucl. Acids Res.* **16**, 2389-2409.
9 Chebli, K., Gattoni, R., Schmitt, P., Hildwein, G. and Stevenin, J. (1989). The 216-nucleotide intron of the E1A pre-mRNA contains a hairpin structure that permits utilization of unusually distant branch acceptor. *Mol. Cell. Biol.* **9**, 4852-4861.
10 Smith, C. W. J., Porro, E. B., Patton, J. G. and Nadal-Ginard, B. (1989). Scanning from an independently specified branch point defines the 3′ splice site of mammalian introns. *Nature* **342**, 243-247.
11 Deshler, J. O., Larson, G. P. and Rossi, J. J. (1989). Kluyveromyces lactis maintains Saccharomyces cerevisiae intron-encoded splicing signals. *Mol. Cell. Biol.* **9**, 2208-2213.
12 Cellini, A., Felder, E. and Rossi, J. J. (1986). Yeast pre-messenger RNA splicing efficiency depends on critical spacing requirements between the branch point and 3′ splice site. *EMBO J.* **5**, 1023-1030.
13 Deshler, J. O. and Rossi, J. J. (1991). Unexpected point mutations activate cryptic 3′ splice sites in perturbing a natural secondary structure within a yeast intron. *Genes Dev.* **5**, 1252-1263.
14 Robberson, B. L., Cote, G. J. and Berget, S. M. (1990). Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* **10**, 84-94.
15 Newman, A. J. and Norman, C. (1992). U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. *Cell* **68**, 743-754.
16 Domenjoud, L., Gallinaro, H., Kister, L., Meyer, S. and Jacob, M. (1991). Identification of a specific exon sequence that is a major determinant in the selection between a natural and a cryptic 5′ splice site. *Mol. Cell. Biol.* **11**, 4581-4590.
17 Steingrimsdottir, H., Rowley, G., Dorado, G., Cole, J. and Lehmann, A. R. (1992). Mutations which alter splicing in human hypoxanthine-guanine phosphoribosyltransferase gene. *Nucl. Acids Res.* **20**, 1201-1208.
18 Munroe, S. H. (1984). Secondary structure of splice sites in adenovirus mRNA precursors. *Nucl. Acids Res.* **12**, 8437-8456.
19 Wollenzien, P. L., Goswami, P., Teare, J., Szeberenyi, R. and Goldenberg, C. J. (1987). The secondary structure of a messenger RNA precursor probed with psoralen is melted in a in vitro splicing reaction. *Nucl. Acids Res.* **15**, 9279-9298.
20 Solnick, D. (1985). Alternative splicing caused by RNA secondary structure. *Cell* **43**, 667-676.
21 Solnick, D. and Lee, S. I. (1987). Amount of RNA secondary structure required to induce an alternative splice. *Mol. Cell. Biol.* **7**, 3194-3198.
22 Eperon, L. P., Estibeiro, J. P. and Eperon, I. C. (1986). The role of nucleotidesequences in splice site selection in eukaryotic pre-messenger RNA. *Nature* **324**, 280-282.
23 Eperon, L. P., Graham, I. R., Griffiths, A. D. and Eperon, I. C. (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.
24 Libri, D. and Fiszman, M. Y. (1988). Alternative splicing of the chicken β tropomyosin gene may be controlled by secondary structure of the primary transcript. In *Sarcomeric and non sarcomeric muscles basic and applied research prospects for the 90's* (ed. Ugo Carraro), pp. 67-72. Unipress Padova, Padova.
25 Libri, D., Lemonnier, M., Meinnel, T. and Fiszman, M. Y. (1989). A single gene codes for the β subunit of smooth and skeletal muscle tropomyosin in the chicken. *J. Biol. Chem.* **264**, 2935-2944.
26 Forry-Schaudies, S., Maihle, N. J. and Hughes, S. H. (1990). Generation of skeletal, smooth and low molecular weight non-muscle tropomyosin isoforms from the chicken tropomyosin 1 gene. *J. Mol. Biol.* **211**, 321-330.
27 Forry-Schaudies, S. and Hughes, S. H. (1991). The chicken tropomyosin 1 gene generates nine mRNAs by alternative splicing. *J. Biol. Chem.* **266**, 13821-13827.
28 Helfman, D. M., Cheley, S., Kuismanen, E., Finn, L. A. and Yamawaki-Kataoka, Y. (1987). Non-muscle and muscle tropomyosin isoforms are expressed from a single gene by alternative RNA splicing and polyadenylation. *Mol. Cell Biol.*, **6**, 3582-3595.
29 Libri, D., Marie, J., Brody, E. and Fiszman, M. Y. (1989). A subfragment of the β tropomyosin gene is alternatively spliced when transfected into differentiating muscle cells. *Nucl. Acids Res.* **17**, 6449-6462.
30 Goux-Pelletan, M., Libri, D., d'Aubenton-Carafa, Y., Fiszman, M., Brody, E. and Marie, J. (1990). *In vitro* splicing of mutually exclusive exons from the chicken b-tropomyosin gene: role of the branch point location and very long pyrimidine stretch. *EMBO J.* **9**, 241-249.
31 Libri, D., Goux-Pelletan, M., Brody, E. and Fiszman, M. Y. (1990). Exon as well as intron sequences are cis-regulating elements for the mutually exclusive alternative splicing of the β tropomyosin gene. *Mol. Cell. Biol.* **10**, 5036-5046.
32 Clouet-d'Orval, B., d'Aubenton-Carafa, Y., Marie, J. and Brody, E. (1991). Determination of an RNA structure involved in splicing inhibition of a muscle specific exon. *J.. Mol. Biol.* **221**, 837-856.
33 Libri, D., Piseri, A. and Fiszman, M. Y. (1991). Tissue specific splicing in vivo of the β-tropomyosin gene: dependence on an RNA secondary structure. *Science* **252**, 1842-1845.
34 Lemonnier, M., Balvay, L., Mouly, V., Libri, D. and Fiszman, M. Y. (1991). The chicken gene encoding the α isoform of fast-twitch muscle fibers: organization, expression and identification of the major proteins synthesized. *Gene* **107**, 229-240.
35 Libri, D., Balvay, L. and Fiszman, M. Y. (1992). In vivo splicing of the β tropomyosin pre-mRNA: a role for branch point and donor site competition. *Mol. Cell. Biol.* **12**, 3204-3215.
36 Clouet d'Orval, B., d'Aubenton Carafa, Y., Sirand-Pugnet, P., Gallego, M., Brody, E. and Marie, J. (1991). RNA secondary structure repression of a muscle specific exon in Hela cell nuclear extracts. *Science* **252**, 1823-1828.
37 Guo, W., Mulligan, G. J., Wormsley, S. and Helfman, D. M. (1991). Alternative splicing of β-tropomyosin pre-mRNA: cis acting elements and cellular factors that block the use of a skeletal muscle exon in non muscle cells. *Genes Dev.* **5**, 2096-2107.
38 Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature* **289**, 751-758.
39 Friedman, D. I., Imperiale, M. J. and Adhya, S. I. (1987). RNA 3′ end formation in the control of gene expression. *Annu. Rev. Genet.* **21**, 453-488.
40 Aloni, Y. and Hay, N. (1985). Attenuation may regulate gene expression in animal viruses and cells. *Crit. Rev. Biochem.* **18**, 327-383.
41 Schulz, V. P. and Reznikoff, W. S. (1990). In vitro secondary structure analysis of mRNA from lacZ translation initiation mutants. *J. Mol. Biol.* **211**, 427-445.
42 Jacquier, A. (1990). Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *Trends Biochem. Sci.* **15**, 351-354.
43 Wissinger, B., Schuster, W. and Brennicke, A. (1991). Trans splicing in oenothera mitochondria: nad1 mRNAs are edited in exon and trans-splicing group II intron sequences. *Cell* **65**, 473-482.

Laurent Balvay, Domenico Libri and Marc Y. Fiszman are at the Department of Molecular Biology, Pasteur Institute, 25 rue du Dr Roux, 75624 Paris Cedex 15, FRANCE. *To whom correspondence should be sent.