

## Épissage différentiel des transcrits musculaires

Un même gène peut comporter des exons potentiels qui ne se retrouvent pas tous dans chacun de ses transcrits. Leur utilisation alternative engendre ainsi plusieurs types de messagers et, le plus souvent, plusieurs isoformes protéiques. Ces mécanismes d'épissage alternatif diversifient ainsi la gamme des fonctions que peut contrôler un gène. Parfois associés à l'utilisation de promoteurs et de sites de polyadénylation eux aussi alternatifs, les phénomènes d'épissage différentiels sont très fréquents dans le muscle et le cerveau, tissus constitués de cellules qui ne se divisent pas ; ils permettent ainsi une adaptation des messages de gènes importants à différentes phases de la différenciation terminale et en réponse à des stimulations variées, sans passer par une reprogrammation du génome qui pourrait survenir électivement lors de la réplication de l'ADN, et donc nécessiter au moins un cycle de division cellulaire.

---

**Domenico Libri**  
**Marc Y. Fiszman**

---

### ADRESSE

D. Libri : docteur es sciences. M.Y. Fiszman : docteur es sciences. Unité de biochimie, Cnrs UA 1148 département de biologie moléculaire, Institut Pasteur, 25, rue du Docteur-Roux, 75625 Paris Cedex 15, France.

L'épissage est le mécanisme qui permet de passer de la structure morcelée de la séquence codante des gènes (exons) à la séquence des ARN messagers. Généralement la totalité des exons se retrouve au niveau des ARN messagers ; toutefois, dans un certain nombre de cas, il se produit un choix et certains exons sont utilisés de manière différentielle selon le contexte dans lequel le gène est transcrit. On parle alors d'épissage alternatif. Ce mécanisme, qui permet d'augmenter la capacité de codage des gènes, semble être particulièrement utilisé dans le système

musculaire où il contrôle en grande partie la diversité phénotypique des différents types de fibres. Actuellement, on ne sait pas ce qui différencie un exon constitutif d'un exon alternatif. Toutefois, de plus en plus de données expérimentales indiquent que le choix d'un exon alternatif implique l'action concertée d'éléments en *cis* localisés aussi bien dans les introns que dans les exons et de facteurs agissant en *trans*. Pour ce qui est des premiers, de nombreux exemples suggèrent qu'ils interviendraient par l'intermédiaire de structures secondaires. Quant aux seconds, ils sont généralement spécifiques de sta-

des de développement particuliers ou de tissus.

Une des caractéristiques du tissu musculaire est l'extrême diversité des protéines qui y sont exprimées. En particulier, chaque type de muscle est caractérisé par des protéines isomorphes qui lui sont spécifiques. Se pose alors le problème de savoir comment se crée une telle diversité. Deux hypothèses peuvent être envisagées : (1) la multiplicité des formes protéiques est

corrélée à une multiplicité de gènes ; (2) chaque gène est capable de coder pour plus d'une protéine. Dans ce dernier cas, deux mécanismes sont à considérer : soit un réarrangement de fragments génomiques, soit un épissage alternatif de certains exons. Ce dernier mécanisme semble être très fréquemment utilisé dans les tissus musculaires et dans le système nerveux. Il est à remarquer que ces deux tissus sont constitués de cellules à

durée de vie très longue, qui ont perdu la capacité de synthétiser de l'ADN mais qui sont susceptibles d'être reprogrammées sous l'effet d'effecteurs externes. On voit alors l'avantage que représente le phénomène d'épissage alternatif dans ce processus.

Le but de cette revue est d'analyser les différents types d'épissage alternatif utilisés pour l'expression des gènes codant pour des protéines spécifiques des muscles striés.

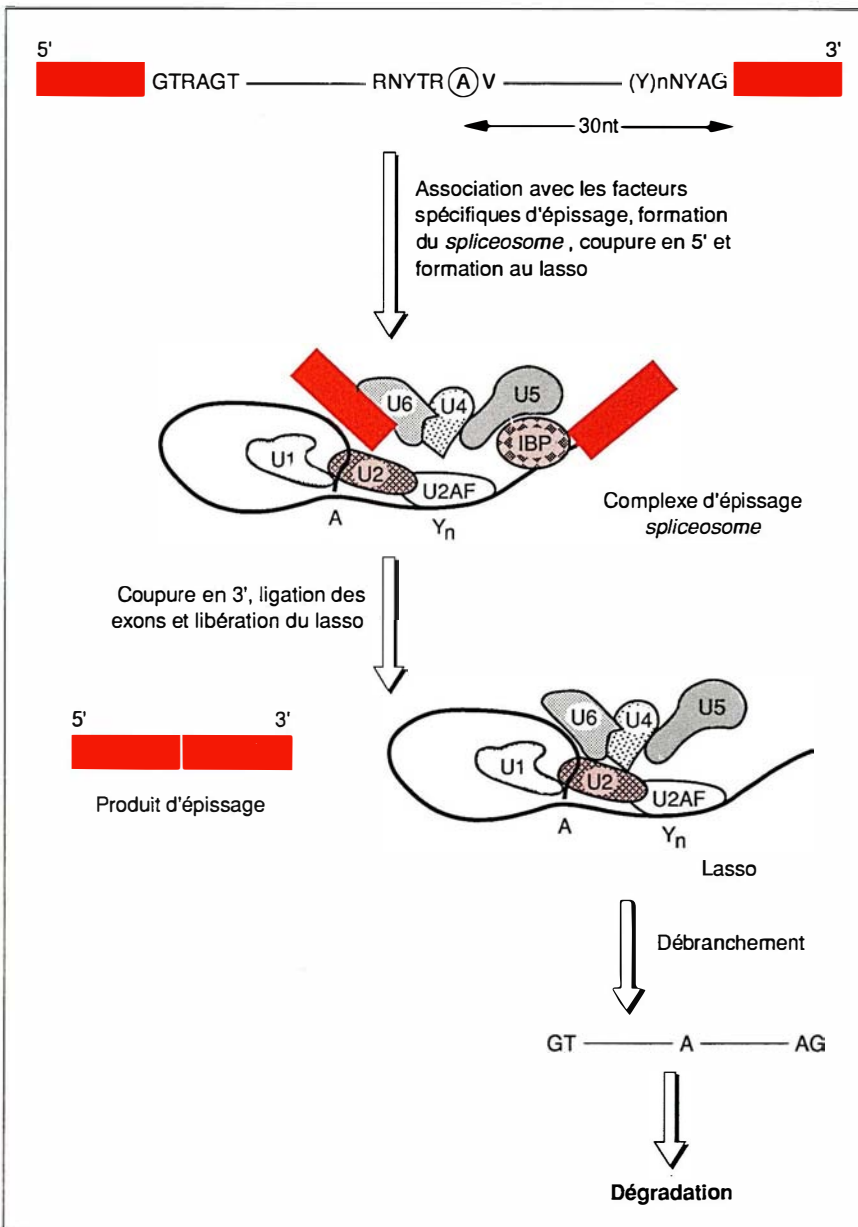


Figure 1. **Épissage du transcrit primaire chez les eucaryotes supérieurs.** Structure d'un fragment de transcrit primaire composé de deux exons séparés par un intron. Au niveau de l'intron sont indiquées les séquences consensus du site donneur (5'), du site accepteur (3') et du point de branchement [1]. Le petit ARN nucléaire U1 sous la forme de particule ribonucléoprotéique s'associe par son extrémité 3' avec la séquence consensus du site donneur. De même, l'association de U2 se fait par appariement avec la séquence consensus du point de branchement en même temps qu'intervient un facteur protéique associé à U2 (U2AF). En revanche, l'association avec U5 met en jeu des interactions ARN-protéine ; quant à U4 et U6, ils se fixent par l'intermédiaire des autres particules ribonucléoprotéiques nucléaires. L'attachement de toutes ces snRNP ainsi que d'autres protéines conduit à la formation du complexe d'épissage (spliceosome). Dans ce complexe, il y a formation d'une liaison 2'-5' entre le A du point de branchement et le G de l'extrémité 5' de l'intron avec libération de l'exon en amont et création d'une forme en lasso. Puis il y a transestérification entre le 3'-OH du dernier nucléotide de l'exon amont et le phosphate en 5' du premier nucléotide de l'exon aval, ce qui provoque la liaison des deux exons et la libération de l'intron sous sa forme en lasso, lequel est ensuite linéarisé avant d'être dégradé. N : nucléotide ; R : purine ; Y : pyrimidine ; IBP : facteur protéique se liant à l'intron (intron binding protein).

## Rappel général concernant le mécanisme de l'épissage

A la différence de leurs homologues bactériens, la plupart des gènes de cellules eucaryotes ont une structure discontinue dans laquelle les séquences que l'on retrouvera dans les ARN messagers (exons) sont interrompues par des séquences qui en seront exclues (introns). Entre la formation des transcrits primaires et l'apparition des messagers dans le cytoplasme vont se dérouler toute une série d'étapes parmi lesquelles l'élimination précise des séquences correspondant aux introns par le phénomène d'épissage. Les introns sont démarqués par des séquences consensus situées aux extrémités 5' et 3' [1]. Ces séquences sont impliquées dans des phénomènes de reconnaissance avec des petits ARN nucléaires intervenant sous la forme de ribonucléoprotéines (snRNP pour *small nuclear ribonucleoprotein*). Au cours du processus d'épissage, l'extrémité 5' de l'intron (site donneur d'épissage) est coupée et une réaction de transestérification se produit entre le résidu guanosyl libéré et un résidu adénosyl (point de branchement) situé dans la partie 3' de l'intron, pour aboutir à la formation d'une structure en lasso (*m/s*, suppl. n° 7, vol. 2, p. 4). Ce point de branchement se trouve au sein d'une séquence conservée et il est situé entre 18 et 40 nucléotides de l'extrémité 3'. Ensuite, une deuxième réaction de transestérification se produit au niveau de l'extrémité 3' (site accepteur d'épissage) de l'intron qui est à son tour coupée tandis que les deux exons sont soudés l'un à l'autre et que l'intron est libéré sous sa forme en lasso (voir *figure 1*). L'ensemble de ces réactions se produit dans une structure complexe qui a reçu le nom de *spliceosome*.

## Les différents modes d'épissage alternatif

Dans la majorité des cas, tous les introns sont éliminés de telle sorte que tous les exons présents au niveau du transcrit primaire sont également présents dans l'ARN messager cytoplasmique. Dans ces conditions, chaque unité de transcription donne

naissance à un seul transcrit stable. Dans certains cas, la même unité de transcription peut donner naissance à plusieurs ARN messagers stables par l'exclusion de certains exons de certains transcrits. On parle alors d'épissage alternatif dont nous allons décrire brièvement les différents modes (*figure 2*). Il va de soi que ces différents modes ne sont pas exclusifs les uns des autres, et bien souvent, plusieurs d'entre eux sont utilisés au niveau d'un même gène. C'est, par exemple, le cas des gènes codant pour les tropomyosines.

**Mode combinatoire :** les exons non constitutifs peuvent être considérés comme des cassettes qui sont aussi bien incluses qu'exclues des ARN messagers. S'il existe plus d'un exon de ce type, le nombre théorique d'ARN messagers possibles est alors de  $2^n$ ,  $n$  étant le nombre de ces exons.

Ce type d'épissage a été décrit pour le gène de la troponine T de muscle squelettique rapide [2, 3] ou celui de la troponine T cardiaque [4]. Dans le premier cas, il existe 5 exons du type cassette, tous localisés en série entre le troisième et le neuvième exon. La particularité de ce système réside dans le fait que plus d'une de ces cassettes peut être présente en même temps sans pour autant changer le cadre de lecture.

**Exons mutuellement exclusifs :** il s'agit d'une paire d'exons de type cassette qui ne peuvent être présents ensemble ni être tous les deux absents. Dans la plupart des cas, l'absence ou la présence simultanée des deux exons provoque un changement de cadre de lecture.

Ce mode d'épissage se retrouve pour les gènes des tropomyosines [5-7] ainsi que pour le gène de la troponine T de muscle squelettique rapide [8].

**Sites accepteurs ou donneurs internes :** il existe des exons qui, dans certaines conditions, ne sont pas reconnus dans leur intégralité mais seulement partiellement. Cette reconnaissance partielle correspond à l'existence de sites donneurs ou accepteurs d'épissage au sein de l'exon considéré. Ce mode d'épissage alternatif a été décrit dans le gène codant pour une chaîne légère de la myosine cardiaque (MLC<sub>1c</sub> pour

## RÉFÉRENCES

1. Mount S. A catalogue of splice junction sequences. *Nucleic Acids Res* 1982 ; 10 : 459-72.
2. Breitbart RE, Nadal-Ginard B. Developmentally induced, muscle specific trans factors control the differential splicing of alternative and constitutive troponin T exons. *Cell* 1987 ; 49 : 793-803.
3. Breitbart RE, Nguyen MT, Medford RM, Destrée AT, Mahdavi V, Nadal-Ginard B. Intricate combinatorial patterns of exon splicing generate multiple regulated troponin T isoforms from a single gene. *Cell* 1985 ; 41 : 67-82.
4. Cooper TA, Ordhal CP. A single troponin T gene regulated by different programs in cardiac and skeletal muscle development. *Science* 1984 ; 226 : 976-82.
5. Helfman DM, Cheley S, Kuismanen E, Finn LA, Yamawaki-Kataoka Y. Nonmuscle and muscle tropomyosin isoforms are expressed from a single gene by alternative RNA splicing and polyadenylation. *Mol Cell Biol* 1986 ; 6 : 3582-95.
6. Libri D, Lemonnier M, Meinel T, Fiszman MY. A single gene codes for the  $\beta$  subunit of smooth and skeletal muscle tropomyosin in the chicken. *J Biol Chem* 1989 ; 264 : 2935-44.
7. Ruiz-Opazo N, Nadal-Ginard B.  $\alpha$ -tropomyosin gene organization alternative splicing of duplicated isotype-specific exons accounts for the production of smooth and striated muscle isoforms. *J Biol Chem* 1987 ; 262 : 4755-65.
8. Medford RM, Nguyen HT, Destrée AT, Summers E, Nadal-Ginard B. A novel mechanism of alternative RNA splicing for the developmentally regulated generation of troponin T isoforms from a single gene. *Cell* 1984 ; 38 : 409-21.
9. Nakamura S, Nabeshima YI, Kobayashi H, Nabeshima Y, Nonomura Y, Fujii-Kuriyama Y. Single chicken cardiac myosin alkali light-chain gene generates two different mRNAs by alternative splicing of a complex exon. *J Mol Biol* 1988 ; 203 : 895-904.



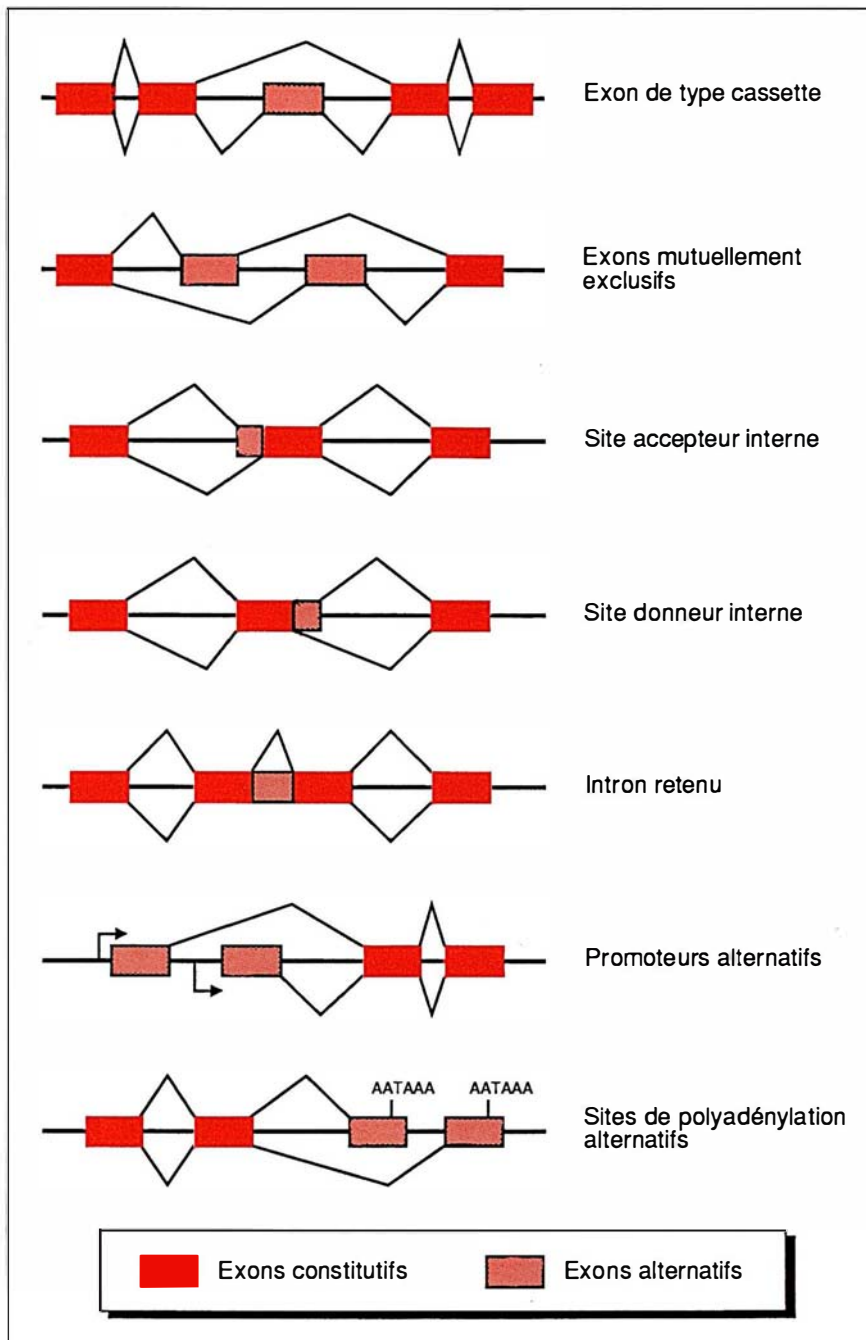


Figure 2. **Les différents modes d'épissage alternatif.** Les flèches indiquent des initiations de transcription.

*myosin light chain 1c*) chez les oiseaux [9].

**Rétention d'intron :** dans certains cas, un intron peut ne pas être éliminé. Ce mode d'épissage a été décrit de nombreuses fois dans divers systèmes. Toutefois, dans le système musculaire, un des rares exemples semble être celui du gène codant pour l'acétylcholinestérase [10].

**Exons multiples en 5' et en 3' :** dans de nombreux systèmes, les ARN messagers produits par épissage alternatif sont issus de transcrits primaires différents. Cette différence de transcrits primaires est due au fait qu'il existe soit des sites distincts d'initiation de la transcription, soit des signaux multiples pour la polyadénylation des messagers, chacun de ces signaux étant contenu dans des exons séparés.

**Cas des promoteurs multiples :** l'exemple le plus étudié correspondant à cette situation est celui du gène qui code pour les chaînes légères de la myosine du type MLC1/MLC3 [11-13]. On a également décrit plusieurs promoteurs pour le gène qui code pour l'aldolase A [14-16]. Enfin, certains gènes codant pour des tropomyosines ont également plusieurs promoteurs [17, 18].

**Sites multiples de polyadénylation :** cette situation a été essentiellement décrite pour les gènes de drosophile. Chez les vertébrés, l'exemple type est celui des gènes codant pour les tropomyosines [5-7, 17-20].

### Régulation de l'épissage alternatif

L'examen des résultats qui ont été publiés au cours de ces dernières années fait clairement apparaître que des éléments essentiels pour le contrôle de l'épissage alternatif doivent être contenus dans la séquence du transcrit primaire (éléments agissant *en cis*), mais il est non moins clair que dans les cas d'épissage alternatif spécifique d'un tissu ou d'un stade de développement, il faut faire intervenir des éléments additionnels sous la forme de facteurs diffusibles (facteurs agissant *en trans*). En effet, certains gènes produisent plusieurs types de transcrits primaires, c'est le cas des gènes possédant des exons multiples en 5' ou en 3'. Dans ces cas, on peut penser que ces transcrits

## RÉFÉRENCES

10. Sikorav JL, Duval N, Anselmet A, *et al.* Complex alternative splicing of acetylcholinesterase transcripts in Torpedo electric organ; primary structure of the precursor of the glycolipid-anchored dimeric form. *EMBO J* 1988; 7 : 2983-93.
  11. Nabeshima YI, Fujii-Kuriyama Y, Muramatsu M, Ogata K. Molecular cloning and nucleotide sequences of the complementary DNAs to chicken skeletal muscle myosin two alkali light chain mRNAs. *Nucl Acids Res* 1982; 10 : 6099-110.
  12. Periasamy M, Strehler EE, Garfinkel LI, Gubits RM, Ruiz-Opazo N, Nadal-Ginard B. 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* 1984; 259 : 13595-604.
  13. Robert B, Daubas P., Akimenko MA, Cohen A, Guenet JL, Buckingham ME. A single locus in the mouse encodes both myosin light chains 1 and 3, a second locus corresponds to a related pseudogene. *Cell* 1984; 39 : 129-40.
  14. Colbert MC, Ciejek-Baez E. Alternative promoter usage by aldolase A during *in vitro* myogenesis. *Dev Biol* 1988; 130 : 392-6.
  15. Joh K, Arai Y, Mukai T, Hori K. Expression of three mRNA species from a single rat aldolase A gene, differing in their 5' non coding regions. *J Mol Biol* 1986; 190 : 401-10.
  16. Marie P, Gautron S, Hakim V, Groggi C, Mennecier F, Kahn A. Characterization of three optional promoters in the 5' region of the human aldolase A gene. *J Mol Biol* 1987; 197 : 425-38.
  17. Clayton L, Reinach FC, Chumbley GM, Macleod AR. Organization of the hTMNm gene: Implication for the evolution of muscle and non-muscle tropomyosins. *J Mol Biol* 1988; 201 : 507-15.
  18. Lindquester GJ, Flach JE, Fleenor DE, Hickman KH, Devlin RB. Avian tropomyosin gene expression. *Nucleic Acids Res* 1989; 17 : 2099-117.
- contiennent suffisamment d'information en *cis* pour contrôler un profil d'épissage particulier. Dans ces conditions, il n'est pas nécessaire d'impliquer des facteurs agissant en *trans*. En revanche, avec les gènes qui produisent plusieurs messagers à partir d'un transcrite primaire unique, il faut obligatoirement faire intervenir des facteurs agissant en *trans*.
- Sur le plan des mécanismes de contrôle, on peut diviser l'épissage alternatif en deux catégories selon que le choix des sites d'épissage est gouverné exclusivement par des éléments agissant en *cis* ou par une combinaison d'éléments agissant en *cis* et de facteurs agissant en *trans*. Dans ce dernier cas, les éléments en *cis* peuvent être particuliers au mode d'épissage alternatif envisagé ou faire partie des éléments consensus de l'épissage en général.
- Régulation par des éléments en *cis*.** Le meilleur exemple de ce type de régulation est fourni par le gène qui code pour les chaînes légères de la myosine de type MLC1/MLC3. Ainsi que le montre la figure 3, ce gène est constitué par 9 exons. Les deux premiers sont commandés chacun par un promoteur différent et sont suivis par deux exons alternatifs tels que l'exon 3 est toujours associé à l'exon 2 et l'exon 4 à l'exon 1. Ce gène produit deux transcrits. L'un, qui code pour la chaîne légère MLC1, est présent dans tous les muscles embryonnaires et dans les muscles adultes de type rapide, alors que l'autre code pour la chaîne légère MLC3 et n'est essentiellement présent que dans les muscles adultes de type rapide.
- Pour analyser les mécanismes de régulation, la méthode de choix consiste à fabriquer un minigène constitué par un promoteur ubiquitaire (le plus souvent, on utilise un des promoteurs du virus SV40) suivi par les exons dont on veut étudier le mode d'épissage. Le tout est terminé par un exon portant un signal de polyadénylation afin de stabiliser les transcrits. Ce minigène est ensuite introduit dans une cellule donnée et la nature des produits d'épissage est ensuite analysée. Si on réalise une telle expérience avec les exons du gène MLC1/MLC3, on constate que si le minigène comporte les exons 1, 3 et 4, l'exon 1 sera associé à
- l'exon 4. En revanche, si le minigène comporte les exons 2, 3 et 4, alors l'exon 2 sera associé à l'exon 3 et l'exon 4 sera toujours exclu [21]. Par ailleurs, si on introduit le couple des exons 3-4 entre deux exons provenant d'un gène hétérologue (par exemple, deux exons du gène de l'insuline), on constate que c'est systématiquement l'exon 4 qui est utilisé. On voit donc, dans ce système, que tout se passe comme si, une fois l'exon initial choisi (exon 1 ou exon 2), le choix du deuxième exon était entièrement déterminé. Par ailleurs, l'expérience avec le minigène hétérologue indique que les exons 3 et 4 ont peu ou pas d'affinité l'un pour l'autre et que l'exon 4 est dominant sur l'exon 3 vis-à-vis d'un site donneur hétérologue. Pour expliquer l'ensemble de ces données, on peut imaginer que l'utilisation de l'un ou de l'autre promoteur entraîne la formation, au niveau du transcrite primaire, d'une structure secondaire particulière qui va favoriser l'utilisation préférentielle soit de l'exon 4 avec le promoteur 1, soit de l'exon 3 avec le promoteur 2. Nous reviendrons plus en détail sur ce problème de structure secondaire dans un de nos prochains exemples.
- Régulation combinée par des éléments en *cis* et des facteurs agissant en *trans*.** Cette catégorie regroupe la majeure partie des transcrits soumis à épissage alternatif. Nous limiterons cette partie à l'analyse de deux exemples : d'une part, celui du gène qui code pour la  $\beta$  tropomyosine de muscle squelettique et, d'autre part, celui qui code pour la troponine T de cœur.
- La figure 4 présente la structure du gène  $\beta$  tropomyosine de poulet [6]. Ce gène possède deux promoteurs; le plus en 5' étant utilisé dans les tissus musculaires, squelettique et lisse, alors que le promoteur interne est utilisé exclusivement dans le tissu non musculaire. Outre ces deux promoteurs, on note la présence de deux exons en 3' : l'exon 9A est exclusivement utilisé dans le muscle du squelette alors que l'exon 9B est utilisé aussi bien dans le muscle lisse que dans le tissu non musculaire. Enfin, il existe deux exons internes soumis à épissage alternatif : l'exon 6A, utilisé dans le muscle lisse et le tissu non musculaire, et l'exon 6B, utilisé uniquement dans le muscle du

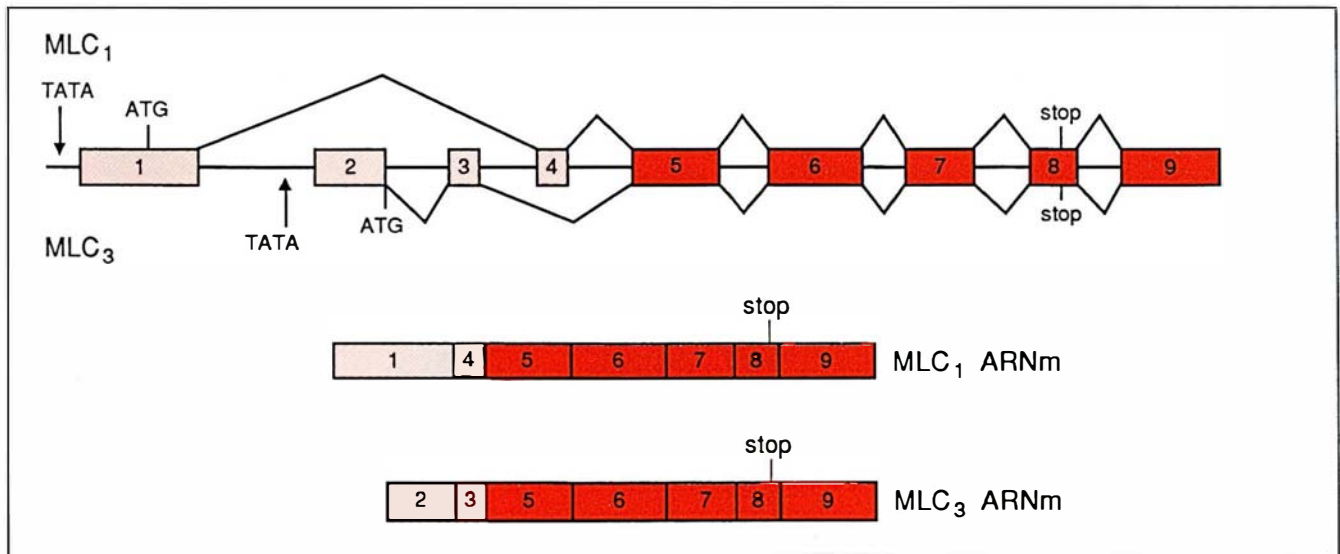


Figure 3. **Organisation du gène codant pour les chaînes légères de la myosine de muscle squelettique rapide (MLC1/MLC3).** La partie supérieure représente la structure simplifiée du gène. Les flèches représentent les deux initiations de la transcription. L'indication « stop » sur l'exon 8 marque la fin de la séquence codante. La partie inférieure représente les deux ARN messagers.

squelette. Ce gène produit trois transcrits codant pour trois protéines : la  $\beta$  tropomyosine de muscle squelettique, la  $\beta$  tropomyosine de muscle lisse et une tropomyosine non musculaire. Bien qu'il y ait deux sites d'initiation de la transcription, il est évident qu'ils n'influencent pas l'utilisation des exons 6A et 6B, puisque le même choix se produit au niveau des transcrits présents dans le muscle lisse (exon 6A) et le muscle du squelette (exon 6B) qui sont initiés au même promoteur. En revanche, la question peut se poser pour les extrémités 3', puisque l'exon 6A est toujours associé à l'exon 9B et l'exon 6B à l'exon 9A. Pour résoudre ce problème, nous avons construit un minigène contenant le promoteur de SV40 associé aux exons 5, 6A, 6B et 7. Ce minigène a été introduit dans des cellules musculaires et les transcrits ont été analysés en fonction de l'état de différenciation des cellules. Nous avons observé que dans les cellules à l'état myoblaste, la construction donne un transcrit qui contient les exons 5, 6A et 7 : alors qu'à l'état myotube, ce

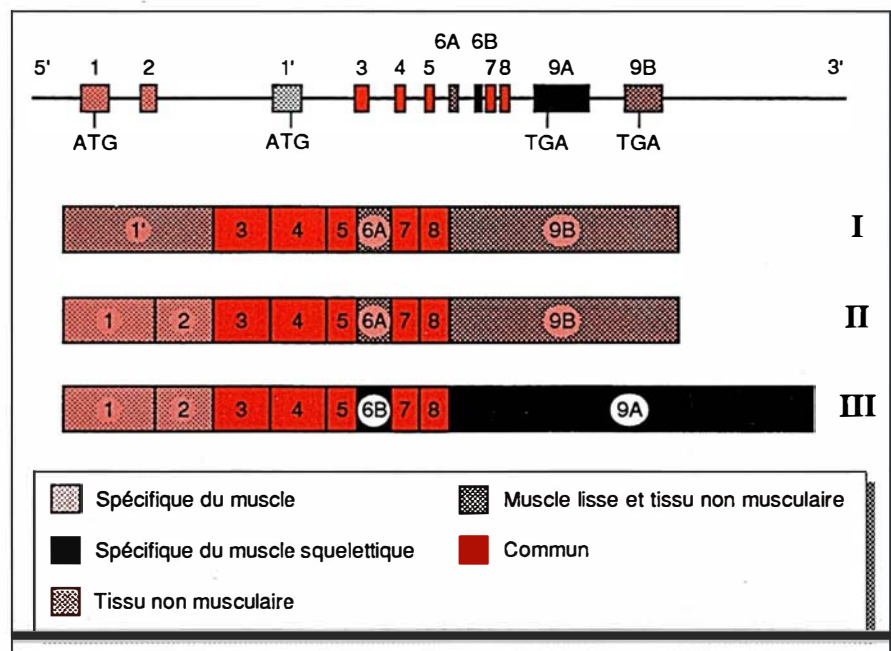


Figure 4. **Organisation du gène codant pour la  $\beta$  tropomyosine.** La partie supérieure représente l'organisation du gène. L'indication « ATG » représente le début de la séquence codante alors que les « TGA » marquent la fin de la séquence codante. Les trois ARN messagers ARN<sub>m</sub> sont représentés en dessous. I : ARN<sub>m</sub> dans le tissu non musculaire ; II : ARN<sub>m</sub> dans le muscle lisse ; III : ARN<sub>m</sub> dans le muscle squelettique.



## RÉFÉRENCES

19. Helfman DM, Cheley S, Kuismanen E, Finn LA, Yamawaki-Kataoka Y. Non muscle and muscle tropomyosin isoforms are expressed from a single gene by alternative RNA splicing and polyadenylation. *Mol Cell Biol* 1987 ; 6 : 3582-95.
20. Wicczorek DF, Smith CWJ, Nadal-Ginard B. The rat  $\alpha$  tropomyosin gene generates a minimum of six different mRNAs coding for striated, smooth and nonmuscle isoforms by alternative splicing. *Mol Cell Biol* 1988 ; 8 : 679-94.
21. Andreadis A, Gallego ME, Nadal-Ginard B. Generation of protein isoform diversity by alternative splicing : mechanistic and biological implications. *Ann Rev Cell Biol* 1987 ; 3 : 207-42.
22. Libri D, Marie J, Brody E, Fiszman MY. A subfragment of the  $\alpha$  tropomyosin gene is alternatively spliced when transfected into differentiating muscle cells. *Nucleic Acids Res* 1989 ; 17 : 6449-62.
23. Smith CWJ, Nadal-Ginard B. Mutually exclusive splicing of  $\beta$  tropomyosin exons enforced by an unusual lariat branch point location : implication for constitutive splicing. *Cell* 1989 ; 56 : 749-58.
24. Cooper TA, Cardone MH, Ordahl CP. Cis requirements for alternative splicing of the cardiac troponin T pre-mRNA. *Nucleic Acids Res* 1988 ; 17 : 8443-65.
25. Cooper TA, Ordahl CP. Nucleotide substitutions within the cardiac troponin T alternative exon disrupt pre-mRNA alternative splicing. *Nucleic Acids Res* 1989 ; 17 : 7905-21.
26. Gower HJ, Barton CH, Elsom VL, et al. Alternative splicing generates a secreted form of N-CAM in muscle and brain. *Cell* 1988 ; 55 : 955-64.
27. Lenz S, Lohse P, Seidel U, Arnold HH. The alkali light chains of human smooth and non muscle myosins are encoded by a single gene : tissue-specific expression by alternative splicing pathways. *J Biol Chem* 1989 ; 264 : 9009-15.
- transcrit contient les exons 5, 6B et 7. Trois conclusions s'imposent immédiatement : premièrement, l'épissage alternatif des exons 6A et 6B est indépendant des séquences en 5' et en 3' du gène ; deuxièmement, les éléments de régulation en *cis* sont localisés sur le fragment génomique compris entre les exons 5 et 7 et, troisièmement, des facteurs agissant en *trans* sont nécessaires puisque les transcrits primaires sont identiques dans les myoblastes et les myotubes [22].
- Afin de comprendre la nature des éléments en *cis*, nous avons, d'une part, comparé l'intron alternatif (celui qui sépare les exons 6A et 6B) à des introns constitutifs et, d'autre part, analysé la séquence du transcrit primaire entre l'exon 5 et l'exon 7 en utilisant un programme informatique permettant la prédiction de structures secondaires dues à des appariements de bases.
- L'examen de l'intron alternatif a fait apparaître une distribution particulière des nucléotides. En effet, l'extrémité 3' de l'intron est particulièrement riche en pyrimidines et cette suite de pyrimidines s'étend sur une centaine de nucléotides alors que, dans un intron constitutif, elle ne dépasse pas la dizaine. L'existence d'une région riche en pyrimidines n'est pas spécifique du gène  $\beta$  tropomyosine, ni même des gènes musculaires, mais sa présence est systématiquement associée au fait que l'exon qui suit est impliqué dans un processus d'épissage alternatif. Cette séquence riche en pyrimidines pourrait interagir avec des facteurs d'épissage leur permettant de jouer un rôle régulateur.
- L'analyse informatique du transcrit primaire entre les exons 5 et 7 a fait apparaître des structures secondaires impliquant de manière alternative les exons 6A et 6B. Comme on peut le voir sur la *figure 5*, l'exon 6B pourrait être bloqué dans une structure de type tige-boucle, la tige étant constituée en partie par l'appariement des séquences exoniques en 5' avec des séquences exoniques situées au milieu de l'exon et par l'appariement des séquences exoniques en 3' avec des séquences situées en 3' de l'intron qui sépare les exons 6A et 6B. La réalité de la structure secondaire a été testée en introduisant des changements de la séquence nucléotidique dans les régions supposées appariées, en particulier dans la séquence 5' de l'exon, dans la séquence 3' de l'exon et dans la séquence 3' de l'intron, et en réintroduisant ces constructions dans des cellules musculaires. Sans entrer dans les détails, les résultats obtenus peuvent se résumer de la manière suivante : (1) la structure théorique telle qu'elle est présentée sur la *figure 5* n'est vraisemblablement pas la structure réelle, mais les résultats obtenus semblent pourtant faire intervenir une structure secondaire qui reste encore à définir ; (2) les régions en 3' de l'intron et en 5' de l'exon jouent un rôle capital dans le contrôle de l'épissage alternatif, dans la mesure où toute modification de l'une comme de l'autre entraîne l'utilisation constitutive de l'exon 6B, quel que soit l'état de différenciation de la cellule musculaire. En conclusion, on peut proposer que dans une cellule musculaire non différenciée, l'exon 6B se trouve dans une structure secondaire qui le rend invisible à la machinerie d'épissage ; dans ces conditions, l'exon 6A est utilisé. Au cours de la transition entre myoblaste et myotube, il apparaît un/des facteur(s) qui va/vont reconnaître la région 3' de l'intron et/ou la région 5' de l'exon, empêchant la structure secondaire de se former et permettant ainsi l'utilisation de l'exon 6B (*figure 5*). Inutile de dire que la nature de ce(s) facteur(s) est complètement inconnue de même qu'est inconnu le mécanisme qui assure le caractère mutuellement exclusif des deux exons. A ce sujet, il est intéressant de signaler qu'il existe un cas d'épissage alternatif où ce mécanisme a été élucidé. C'est celui de l'épissage alternatif des exons 2A et 2B du gène  $\alpha$  tropomyosine. Dans ce gène, on s'est aperçu que le point de branchement de l'intron qui sépare les deux exons est situé à 42 nucléotides de l'extrémité 5' de l'intron, ce qui inhibe la première étape de la réaction d'épissage [23]. Un mécanisme très simple peut donc contrôler le caractère mutuellement exclusif de deux exons ; toutefois un tel mécanisme n'est pas général.
- Le deuxième exemple que je souhaiterais mentionner est celui du gène

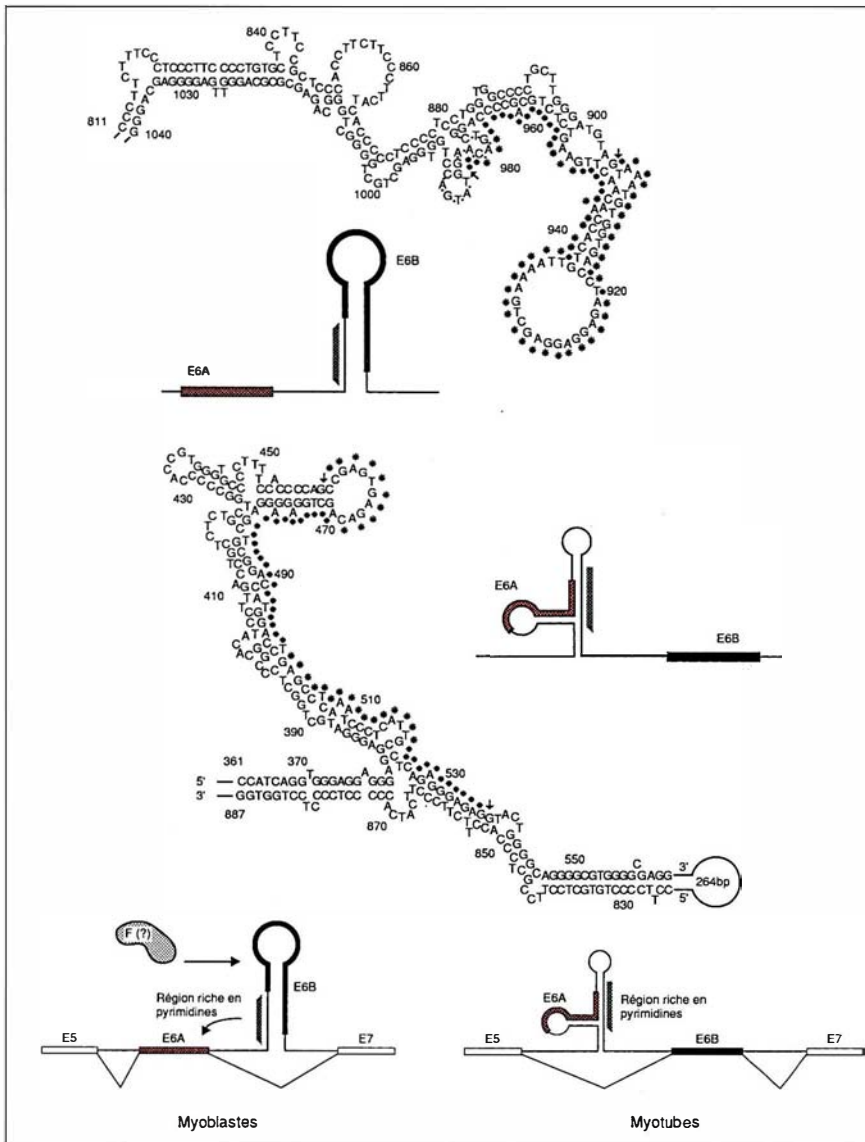


Figure 5. **Exemple de structure alternative du transcrit primaire autour des exons 6A et 6B du gène  $\beta$  tropomyosine.** La partie supérieure représente la structure autour de l'exon 6B ainsi que sa représentation schématique. En dessous est représentée la structure alternative autour de l'exon 6A. Dans les deux cas, les nucléotides associés à des astérisques sont dans l'exon. Les flèches indiquent les bordures 5' et 3' des exons. Enfin, la région hachurée représente la région riche en pyrimidines. La partie inférieure est une représentation schématique du changement dans le mode d'épissage qui se produit au cours de la myogenèse. Dans les myoblastes, l'exon 6B est séquestré par la structure secondaire et l'épissage se fait entre l'exon 6A et l'exon 7. Au cours de la transition myoblastes/myotubes, il apparaît un facteur permettant à l'exon 6B d'être reconnu par la machinerie d'épissage, ce qui provoquerait un changement d'appariement de la région riche en pyrimidines et formation de la structure autour de l'exon 6A. Dans ces conditions, l'épissage se fait entre les exons 5, 6B et 7.

qui code pour la troponine T du muscle cardiaque. Ce gène produit deux transcrits : le plus important correspond à la troponine T cardiaque et est synthétisé aussi bien chez l'embryon que chez l'adulte. Le second est un transcrit mineur qui n'est présent que chez l'embryon dans les muscles squelettiques. Les deux transcrits diffèrent entre eux par l'exclusion de l'exon 5 (forme cardiaque) ou son inclusion (forme squelettique) [4]. L'utilisation de minigène réintroduit dans des cellules en culture a montré qu'aucune séquence en 5' ou 3' n'est requise pour contrôler l'utilisation de l'exon 5 [24]. En revanche, quel que soit le contexte cellulaire dans lequel il s'exprime, le minigène produit toujours les deux transcrits dans un rapport constant, ce qui traduit un phénomène de compétition pour l'utilisation de l'exon 5. Pour expliquer la régulation dans l'organisme, il faut donc faire intervenir des facteurs agissant en *trans*, mais il suffit qu'ils reconnaissent de manière différentielle l'un des éléments constitutifs de l'épissage pour influencer cette compétition et pour que l'exon 5 soit ou non inclus. A l'appui de cette proposition, on peut remarquer qu'il suffit de changer un nucléotide (remplacement d'un T en G) en position +5 dans l'intron (donc dans la séquence du site donneur d'épissage) qui suit l'exon 5 pour que cet exon devienne constitutif [25].

### Implications biologiques de l'épissage alternatif

Les effets fonctionnels et les implications biologiques de l'épissage alternatif sont souvent obscurs mais, du fait de la multiplication des cas d'épissage alternatif, on commence à pouvoir dégager certaines corrélations. Ainsi, l'épissage alternatif va permettre, à partir d'une quantité limitée d'information génétique, d'engendrer des familles plus ou moins complexes de protéines isomorphes. De cette manière, on va pouvoir soit introduire une nouvelle activité, soit moduler une activité existante. C'est ce que nous avons vu avec l'exemple des tropomyosines ou celui de la troponine T. De même, l'épissage alternatif peut entraîner des



modifications dans la localisation cellulaire d'une protéine. C'est ce qui se produit avec la protéine d'adhésion cellulaire N-CAM, pour laquelle l'inclusion d'un exon de type « cassette » permet la formation d'un polypeptide dépourvu du domaine de liaison à la membrane, qui sera ainsi sécrété au lieu d'être exposé à la surface [26]. Parallèlement, l'épissage alternatif peut également être utilisé pour contrôler l'activité d'un gène à un niveau post-transcriptionnel ou même à un niveau purement traductionnel. En effet, l'utilisation d'exons alternatifs correspondant à des régions 5' ou 3' non traduites pourrait permettre d'introduire des signaux contrôlant de manière différentielle la traductibilité ou la stabilité des ARN messagers. De même, l'utilisation d'un exon alternatif peut provoquer un changement du cadre de lecture qui va entraîner l'arrêt de la traduction du transcrite du fait de l'introduction d'un codon stop. C'est par exemple ce qui se produit dans le cas du gène qui code pour une chaîne légère de myosine non musculaire chez l'homme [27]. Ce gène, soumis à épissage alternatif dans sa région 3', code pour une chaîne légère de tissu non musculaire et pour une chaîne légère de la myosine de muscle lisse. De plus, dans le muscle squelettique, il y a utilisation d'un autre exon alternatif, localisé dans la partie 5' du gène, qui change le cadre de lecture et provoque une terminaison anticipée de la traduction du transcrite. Ainsi, dans le muscle squelettique, la transcription du gène n'est pas modifiée, mais il n'y a pas de protéine synthétisée. Ce mode de régulation pourrait être très général et très important et les exemples de son utilisation devraient rapidement s'accumuler.

En conclusion, on voit l'importance que peut avoir l'épissage alternatif tant pour engendrer de la diversité phénotypique que pour contrôler l'activité des gènes. Cela est particulièrement important dans le cas du système musculaire, puisque toute cette régulation peut être réalisée sans reprogrammation transcriptionnelle de l'activité des gènes. En revanche, cette multiplicité de potentialités peut représenter un point de faiblesse, puisque des mutations ponctuelles pourront avoir des répercussions dramati-

ques dans plusieurs tissus (ou plusieurs types de fibres dans le cas du système musculaire). En effet, toute mutation dans un exon commun provoquera des anomalies de l'ensemble des protéines isomorphes et pourra avoir des répercussions dans des tissus différents. Jusqu'à présent, aucune maladie musculaire n'a été corrélée à un défaut d'épissage, mais — compte tenu de leur taille ou de leur complexité — les gènes musculaires sont des candidats susceptibles de présenter de telles altérations ■

## Summary

### Differential splicing of muscle specific transcripts

RNA splicing is an ubiquitous phenomenon which allows the dispersed coding information of eukaryotic genes to be reassembled in the mature RNA molecules. In some cases not all the exons are incorporated into mature mRNAs. This differential incorporation of exon is often under developmental and/or tissue-specific regulation and allows single genes to produce a variety of protein isoforms. This process of alternative splicing is widely used in muscle cells to generate the phenotypic diversity of the various muscle fibers. At the present time, little is known about the mechanisms which distinguish between constitutively *versus* alternatively spliced exons. Evidence is accumulating in support of the hypothesis that *cis*-acting sequences, both in introns and in exons, are involved in the control of this process together with *trans*-acting factors. In some cases, it has been proposed that the *cis*-acting elements could be secondary structures of the primary transcript, although these structures remain to be proven. The *trans*-acting factors are usually developmental stage and/or tissue-specific but nothing is known about their nature.

## TIRÉS A PART

M.Y. Fiszman.