

REGULATION DE L'EXPRESSION GENIQUE

QUATRIEME PARTIE

EPISSAGE ALTERNATIF

Le nombre de gènes humains était estimé à plus de 100 000 gènes. Mais depuis 2003 (année de publication de la séquence complète du génome humain), ce chiffre n'est que d'environ 23 000 gènes, un résultat très voisin de celui d'autres espèces comme la souris, le poisson-zèbre et même les nématodes. Le nombre de gènes d'un organisme ne reflète pas sa réelle complexité biologique. Ce paradoxe résulte de la combinaison de plusieurs phénomènes, dont l'épissage alternatif des pré-mRNA.

Lors de sa maturation, le pré-mRNA subit plusieurs modifications, à savoir: l'ajout de la coiffe, l'excision-épissage, le clivage et la polyadénylation.

* Ajout d'une coiffe (cap de capping): la coiffe est une Guanosine méthylée en N7 (7mG) à l'extrémité 5', cette étape débute avant que la transcription ne soit achevée (transcrit de 20 à 30 nucléotides). La coiffe protège les mRNA des phosphatases et des endonucléases et stimulent leur traduction en favorisant le recrutement des ribosomes.

* Clivage et polyadénylation à l'extrémité 3': la polyadénylation est l'addition d'une queue de 80 à 250 résidus adénylate (polyA) à l'extrémité 3' des ARNm. Cette étape est nucléaire et se produit avant l'épissage.

Chez les eucaryotes, ce processus s'effectue bien sûr à la fin de la transcription par l'ARN polymérase II, au niveau d'un signal spécifique de polyadénylation (AAUAAA) sur l'ARN qui se situe dans la partie non codante. Le processus nécessite certains facteurs comme le **CPSF** (*cleavage polyadenylation specificity factor*) qui se fixe au signal AAUAAA pour cliver ARNm au site de démarrage de la polyadénylation. La queue polyA avec les protéines susceptibles de

s'y associer protègent les mRNA d'une éventuelle dégradation enzymatique. La polyadénylation favorise le transport nucléo-cytoplasmique et dans le recrutement du ribosome pour permettre sa traduction. Seuls les ARNm coiffés et polyadénylés sont efficacement traduits.

* Epissage: c'est la modification principale que subie un pré-mRNA. On sait que les gènes humains sont discontinus car les exons sont séparés par des introns. Les exons d'un gène ont une taille d'environ 150 nucléotides, tandis que les introns mesurent environ 10 fois plus. L'épissage consiste en l'élimination des introns et à la mise bout à bout des exons. Le ARNm mature, sera exporté du noyau vers le cytoplasme pour être traduit.

Les introns, ont été découverts en 1977 chez des virus, les bactéries sont dépourvues d'introns. Les études qui ont suivi le séquençage du génome humain ont montré qu'environ 50% des gènes humains produisent au moins deux transcrits. Aujourd'hui il est affirmé que plus de 95% des gènes humains produisent plusieurs variants d'épissage (**fig. 14**). L'épissage alternatif chez l'homme est donc une règle, et non pas une exception.

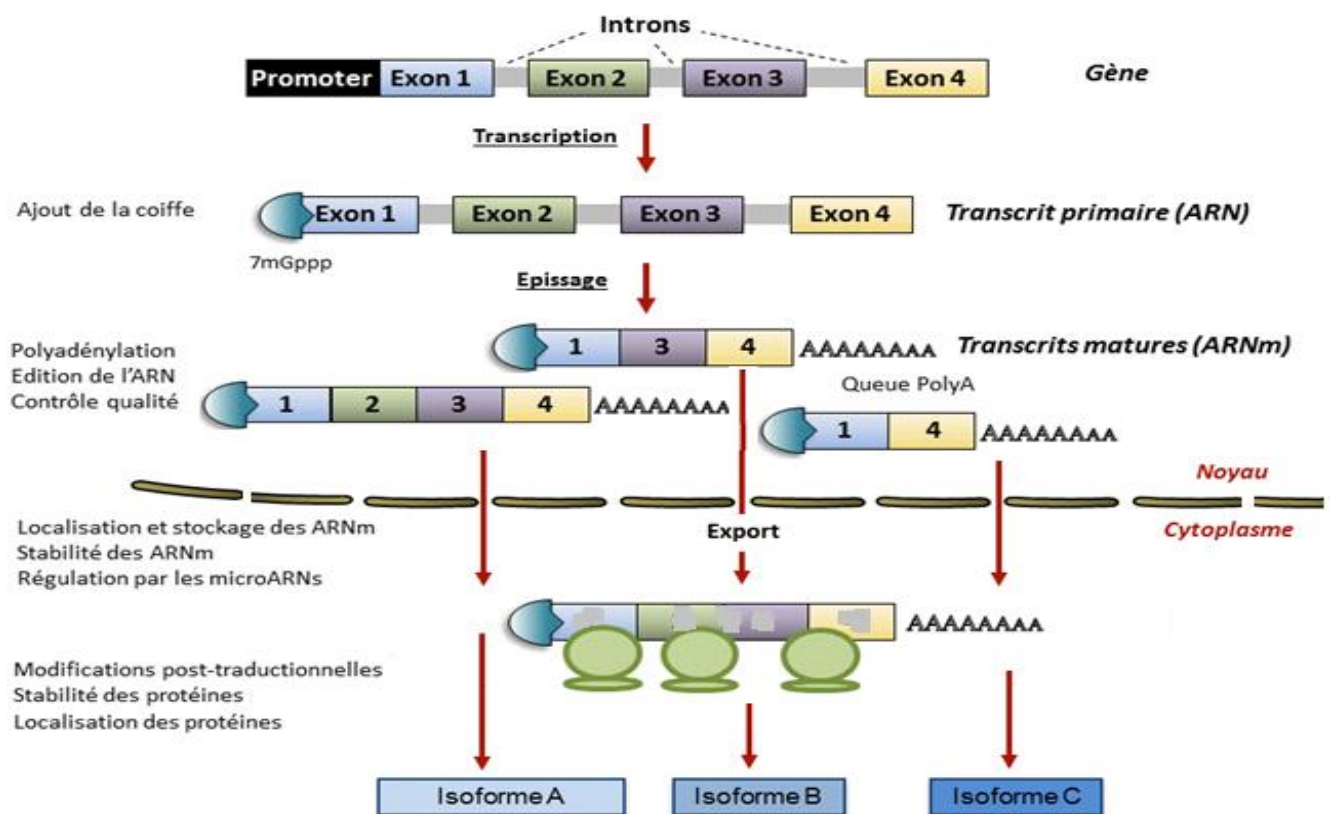


Fig. 14: Différentes étapes de l'expression d'un gène, niveaux de régulation et épissage

Gène ?

Le gène est certainement, le segment d'ADN qui peut être transcrit en RNA. Le pré-ARNm peut être épissé d'une manière alternative, il n'est pas obligatoire que tous les exons d'un gène soient inclus dans le mRNA mature et certains exons sont considérés comme alternatifs. Ainsi, un seul gène peut produire différents mRNA et par conséquent, plusieurs isoformes protéiques.

Dire que l'information génétique du DNA est transcrite sur un seul mRNA qui donnera une seule protéine, n'est pas juste. Il est donc faux de définir un gène, comme étant l'entité qui code pour une seule protéine. En fait, l'unité de l'information génétique n'est plus le gène mais plutôt l'exon.

Le gène pourrait être défini comme étant une succession d'exons sélectionnés alternativement et permettant la production d'un ou de plusieurs mRNA codant autant d'isoformes protéiques. Un gène est donc un message qui peut être lu par la cellule de différentes manières pour assurer différentes fonctions biologiques et répondre aux besoins de la cellule, selon la spécialisation, les conditions physiologiques, le stade de développement ...etc.

1- Mécanisme d'épissage

Le processus excision-épissage est réalisé grâce à deux réactions de transestérification dont le déroulement et la catalyse sont extrêmement complexes. Le mécanisme moléculaire de l'épissage exige une grande précision pour reconnaître les zones qui délimitent les introns des exons dans le pré-mRNAc. La réaction d'épissage est effectuée par le spliceosome (splicing = épissage), un complexe moléculaire formé d'un grand nombre de protéines.

Chez les Eucaryotes, la séquence d'un intron commence par GU et se termine par AG. Dans la majorité des gènes, la séquence signal complète à l'extrémité 5' des introns est: 5'AGGURAGU 3' et celle de l'extrémité 3' est 5' YnNYAGG 3'. Les introns ont aussi un site interne (site de branchement = branchpoint), qui se situe 10 à 40 bases en amont de la séquence signal en 3', sa séquence est NYURAY.

Y = **Pyrimidines** (C/U), R = **Purines** (A/G) et N = n'importe quel nucléotide.

La réaction d'épissage fait appel à des snRNPs (small nuclear ribonucleoproteins), dont certains sont essentiels. En effet, cinq snRNP (U1, U2, U4, U5 et U6) qui reconnaissent les spliceosites 5' et 3', en plus du point de branchement et de la séquence polypyrimidines sont essentiels (**fig. 15**).

Malgré que le fonctionnement du spliceosome ne soit pas bien connu et que les composants responsables de ses activités ne sont pas bien identifiés. Il est connu que le U1 contient une séquence complémentaire du site d'épissage 5' auquel il se fixe et le U2 se lie au point de branchement, le tout constitue le pré-spliceosome. L'entrée des U4/U6 et U5 dans le complexe d'épissage s'accompagne d'une profonde réorganisation du complexe, qui devient mature, permettant un rapprochement physique des sites d'épissage et la réaction catalytique proprement dite (**Fig. 16**).

Deux réactions de transesterification successives assurent l'élimination de l'intron et la ligation des 2 exons. Pratiquement, l'épissage se fait en deux phases:

Phase 1: une fois la liaison phosphodiester du G de **GU** est excisée, l'extrémité 5' de l'intron est libre avec son phosphate, elle sera attachée au 2'OH de **A** du branchpoint. L'intron avec une extrémité détachée donne un lasso, une forme d'épingle à cheveux.

Phase 2: l'extrémité 3' de l'intron est clivée après G du **AG**. Le lasso est libéré puis dégradé dans le noyau et les 2 exons seront reliés avec une liaison phosphodiester.

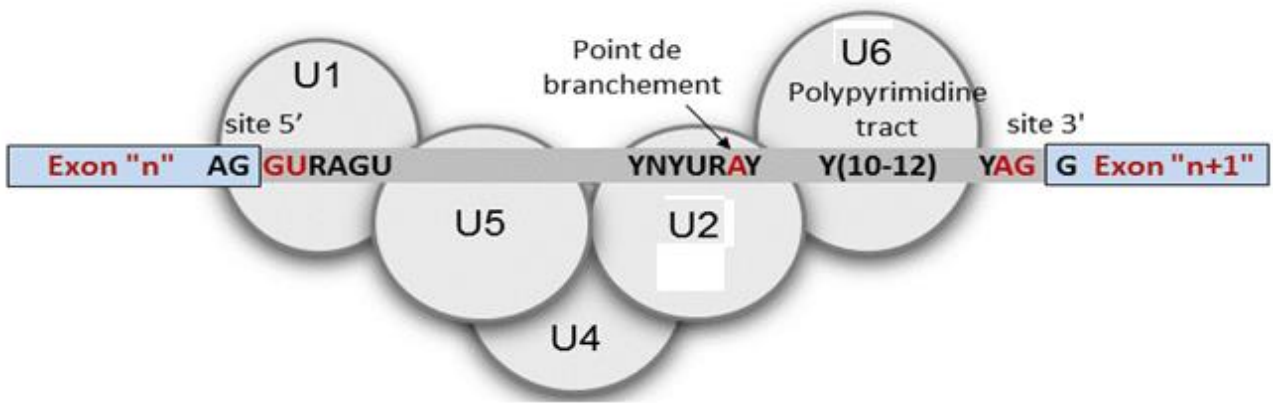


Fig. 15 : Formation du spliceosome

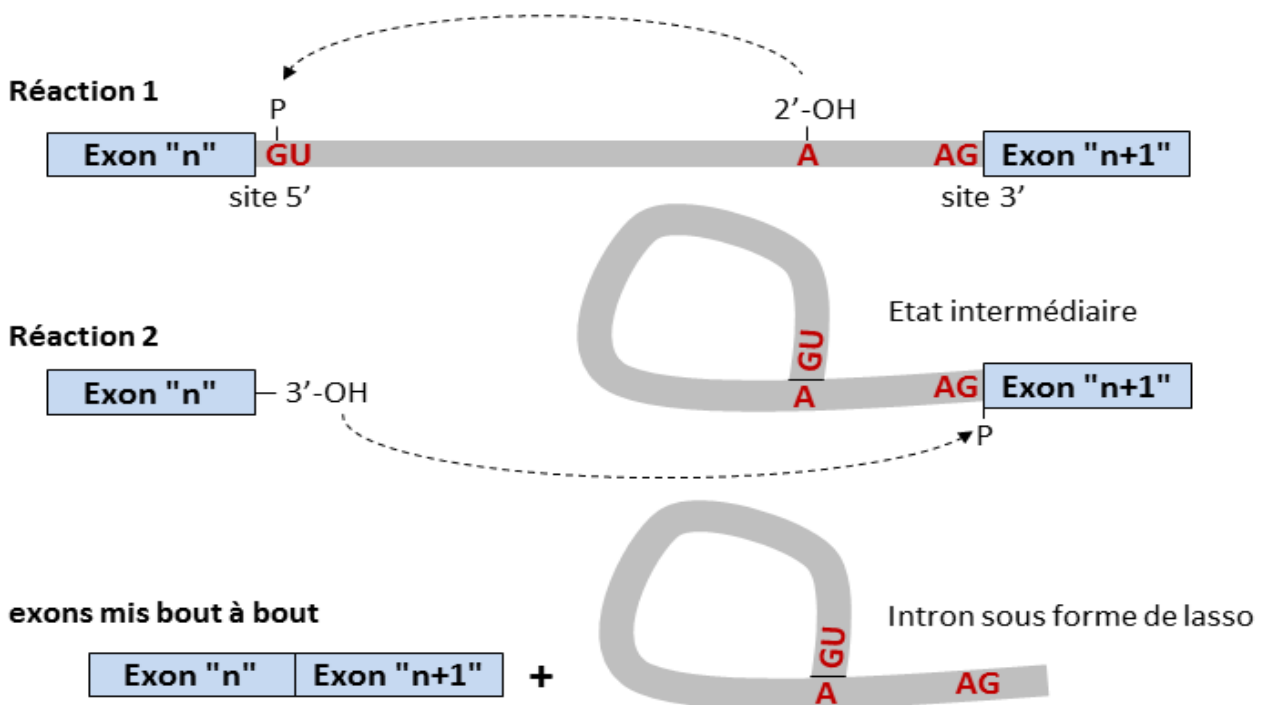


Fig. 16: mécanisme d'excision-épissage

2. Différents types d'épissage alternatif

Les événements de l'épissage alternatif se produisent avec maintien du cadre de lecture, résultant en l'expression de différentes isoformes protéiques. Les exemples d'épissage alternatif peuvent être classés en cinq modèles différents (**Fig.17**).

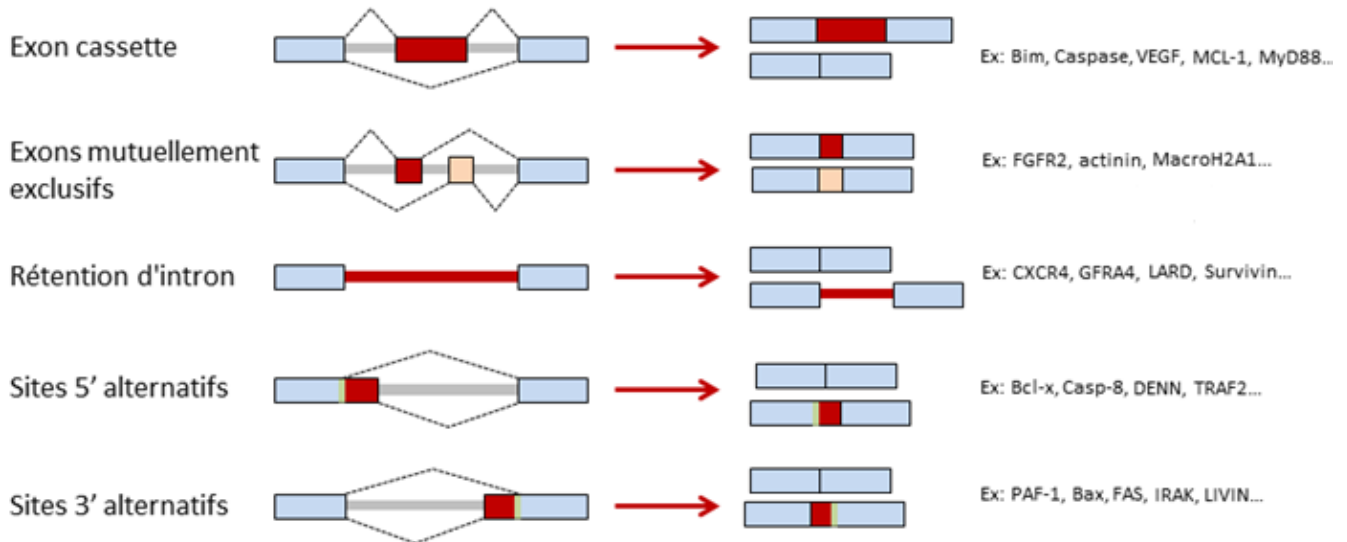


Figure 17 : Différents modèles d'épissage alternatif

1. Exon cassette (saut d'exon): C'est l'épissage alternatif le plus simple, qui prédomine chez les mammifères, donnant lieu à des transcrits plus ou moins longs. L'exon cassette est l'évènement où un exon (compris entre 2 exons du mRNA mature) pourrait être inclus dans le transcrit mature, ou éliminé pour générer deux isoformes protéiques différentes.
2. Exons mutuellement exclusifs: l'un des deux exons (généralement de tailles similaires), est retenu dans le mRNA, mais jamais les deux ensemble. Cet évènement change peu la taille du transcrit mature mais modifie les propriétés de la protéine.
3. Rétention d'intron: un segment de pré-mRNA peut être épissé comme un intron ou bien retenu. Si l'intron ne modifie pas le cadre de lecture, une protéine d'un poids moléculaire supérieur sera produite. Dans le cas contraire, il entraîne l'apparition d'un codon stop prématuré, ce qui provoque la dégradation du transcrit mature. Ce type d'épissage est très rare chez les mammifères.

4. Site d'épissage alternatif en 5': ce variant d'épissage modifie la borne de l'exon dans sa partie aval.
5. Site d'épissage alternatif en 3': cet événement modifie la borne de l'exon dans sa partie amont, incluant ainsi un segment de l'extrémité 3' de l'intron épissé.

3. Epissage alternatif et fonctions protéiques

L'épissage alternatif permet à un gène de générer plusieurs mRNA et par conséquent plusieurs isoformes protéiques **de structure et de fonctions différents**. La majorité des gènes possèdent plusieurs exons, les possibilités d'exclusion ou d'inclusion des exons alternatifs sont donc très grandes. Ainsi, chaque gène peut produire potentiellement des dizaines, voire des centaines de transcrits différents (**Fig.18**).

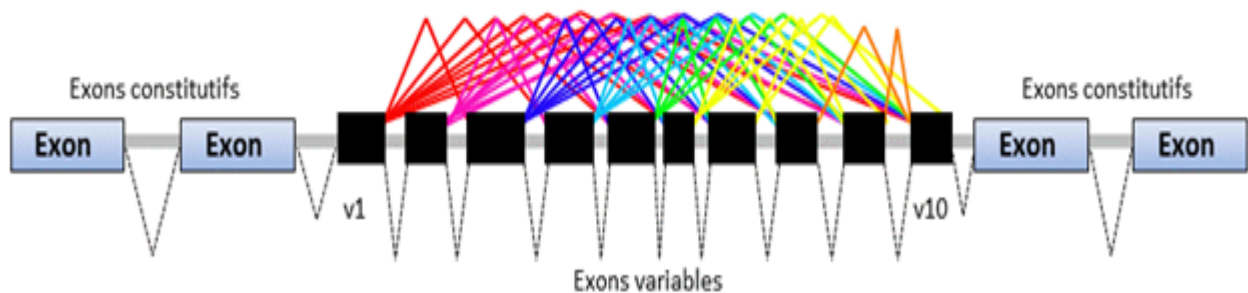


Figure 18: Exons constitutifs et exons variables et la possibilité de formation d'un grand nombre de transcrits différents.

Il est à noter que les sites d'épissage sont incapables, à eux seuls, d'initier l'assemblage du spliceosome. Pratiquement, d'autres séquences régulatrices, agissant en *cis*, doivent intervenir. Ces séquences, qui peuvent se trouver dans les segments des exons ou des introns sont capables d'activer ou d'inhiber la reconnaissance d'un exon à l'aide des facteurs d'épissage (régulateurs *trans*).

Il existe deux grandes familles de régulateurs *trans* se liant aux séquences régulatrices *cis* à savoir les protéines riches en **serine** et **arginine**, nommées protéines **SR** et les **hnRNP** (**heterogeneous nuclear ribonucleoproteins**).

Les protéines SR, en se fixant sur les séquences activatrices (SA), favorisent la reconnaissance des sites d'épissage et facilitent le recrutement des **snRNP** U1 et U2 (**small nuclear ribonucleoproteins**) (**Fig.19**). Par contre, les **hnRNP** se lient sur les

séquences inhibitrices (SI), jouant le rôle de répresseur en masquant les sites d'épissage (**Fig.20**).

L'inclusion ou l'exclusion d'un exon dépend donc de la liaison compétitive de ces différents facteurs à leurs séquences-cibles.

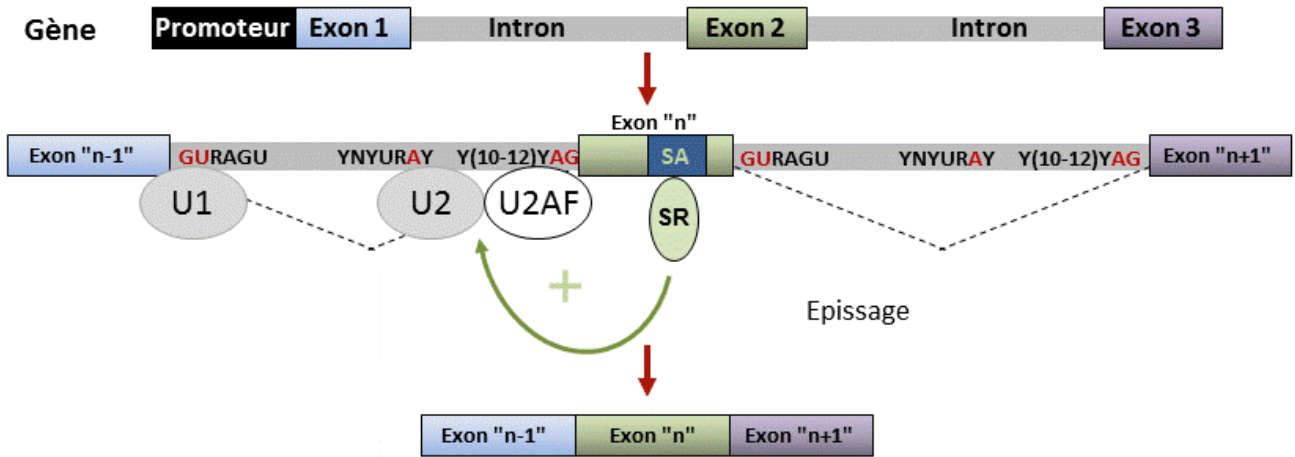


Figure 19: Action positive

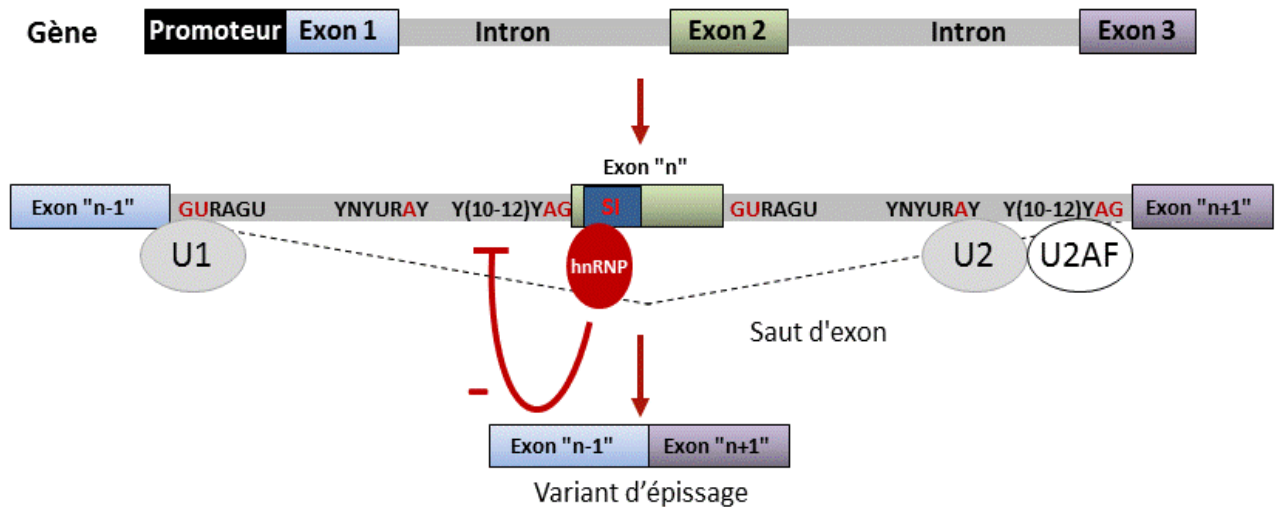


Figure 20: Action négative

Parfois, la diversité fonctionnelle offerte par l'épissage alternatif peut affecter des processus biologiques complètement différents. Par exemple, dans la glande thyroïdienne, les pré-mRNA du gène CALCA produisent, par épissage alternatif, la **calcitonine** (hormone du métabolisme du phosphore et du calcium). Par contre, dans le système nerveux central, les transcrits CALCA produisent le peptide **CGRP**, qui est un vasodilatateur et un médiateur de la douleur (**Fig.21**).

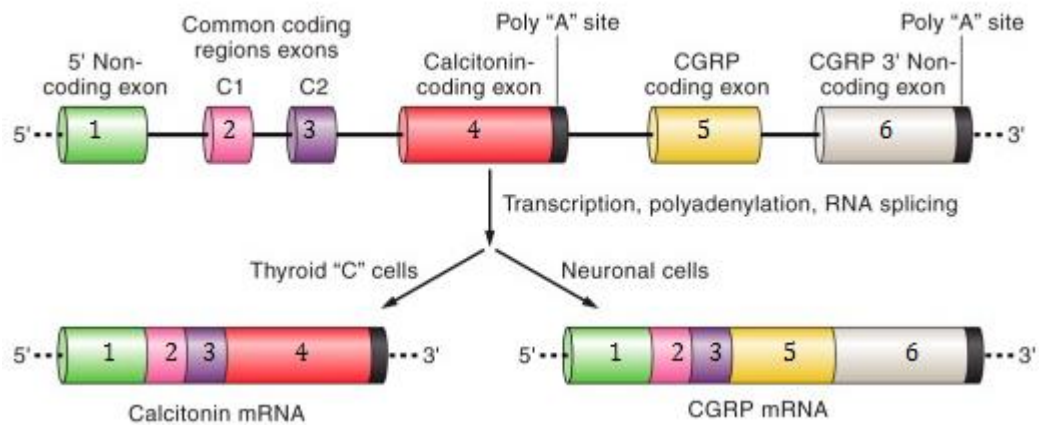


Figure 21 : Gène CALCA. Dans la glande thyroïde il donne la calcitonine (hormone hypocalcémiant...) et dans les neurones il donne CGRP (neuromédiateur de la douleur)

Une autre conséquence possible de l'épissage alternatif pourrait être la localisation des isoformes protéiques produites. Par exemple, un récepteur transmembranaire contient un domaine protéique qui lui permet de s'insérer à la membrane cellulaire, et jouer son rôle. Si un épissage alternatif prive le mRNA de la région codant pour le signal de localisation transmembranaire, la protéine résultante n'est plus transmembranaire et sa fonction est bloquée (pas signalisation par ex).

La fonction d'un gène ne peut être réellement déterminée qu'une fois la fonction de chacun de tous ses variants d'épissage sera connue.

4. Régulation au niveau de la chromatine (Remodelage de la chromatine)

Les histones qui sont des protéines basiques très conservées s'associent étroitement avec l'ADN pour former une structure dense limitant l'accès des protéines à l'ADN (y compris les protéines de la transcription) et réprimant ainsi l'expression des gènes.

Le remodelage de la chromatine est nécessaire à l'initiation de la transcription. En effet, pour que les facteurs d'initiation puissent prendre place au niveau du promoteur et permettre l'ouverture du double brin d'ADN et la formation du TIC, il faut une modification de la structure de la fibre chromatine et des nucléosomes. Cette modification doit être: **rapide, temporaire, réversible et ponctuelle.**

La fixation des protéines de liaison à l'ADN engendre le recrutement de complexes protéiques de remodelage de la chromatine, faisant intervenir une machinerie cellulaire spécialisée comportant des enzymes capables de modifier les histones par **acétylation, phosphorylation ou méthylation**, rompant ainsi les liaisons ADN-histones et permettant la **translocation** des nucléosomes.

4.1. Acétylation des histones

L'acétylation (CO-CH₃) permet de modifier l'affinité des histones pour le DNA et de relâcher leurs interactions moléculaires, donc le DNA devient plus disponible pour accueillir les protéines de transcription.

Les histones H3 et H4 peuvent être acétylés au niveau des résidus lysine situés dans leur partie N-terminale. Les interactions ADN-histones semblent être déstabilisées, ce qui conduit à une forme moins condensée de la chromatine donnant des nucléosomes un peu déstructurés. L'acétylation des histones est un processus réversible catalysé par les enzymes HAT (*histone acetyl-transferase*). A l'inverse, les HDAC (*histone de-acetylase, comme HDAC1, HDAC2, HDAC3, HDAC3A, HDAC3B...etc.*) par leur activité désacétylase permettent à la chromatine acétylée de revenir à son état initial plus condensé.

L'acétylation cible les régions des gènes qui doivent être exprimés, rendant ainsi l'ADN accessible aux facteurs de transcription. De l'autre côté, l'activité transcriptionnelle est aussi contrôlée par le biais des régulateurs qui désacétylent les histones et qui, en conséquence, influencent négativement la liaison des facteurs de transcription à l'ADN.

4.2. Phosphorylation des histones

La phosphorylation du résidu sérine en position 10 (Ser10) de l'histone H3 est associée à l'activation de la transcription chez la levure, la drosophile et les mammifères. Il semble exister une intime corrélation entre la phosphorylation de la Ser10 et l'acétylation des résidus Lys adjacents. En effet, l'histone H3 présent au niveau des loci transcriptionnellement actifs est doublement modifiée: en plus de la phosphorylation de la Ser10, il est acétylé sur la Lys9 et/ou 14.

La phosphorylation de l'histone H1 semble aussi être couplée au déroulement de la transcription. Cette phosphorylation semble maintenir la chromatine dans un état décondensé afin de favoriser la transcription.

4.3. Méthylation des histones

Selon qu'elles ciblent les Lys ou Arg, les histone methyl-transferase (HMT) sont classées en deux catégories.

La première HMT identifiée fut SUV39H1 (levure et drosophile) est capable de méthyler *in vitro* la Lys9 de l'histone H3 mais cette méthylation est inhibée par la phosphorylation préalable de la Ser10 adjacente. *In vivo*, cette modification permet le recrutement de la protéine HP1, qui joue le rôle de répresseur transcriptionnel.

Il existe par ailleurs deux familles de HMT agissant sur des résidus arginine: CARM1 (*co-activator-associated arginine methyltransferase1*) et PRMT1 (*protein arginine methy-ltransferase 1*). CARM1 est un co-activateur capable de méthyler *in vitro* les histones H3 et H2B sur plusieurs résidus arginine, son activité enzymatique est essentielle à l'activation de gènes cibles. PRMT1 pour sa part est une méthyl-transférase spécifique au résidu arginine en position 3 de l'histone H4, à la fois *in vitro* et *in vivo*. La méthylation de l'Arg3 facilite l'acétylation de la Lys8 de l'histone H4 par p300/CBP, mais à l'inverse, l'acétylation de Lys 8 et Lys12 du H4 inhibe sa méthylation en Arg3 par PRMT1 (**fig.22**)

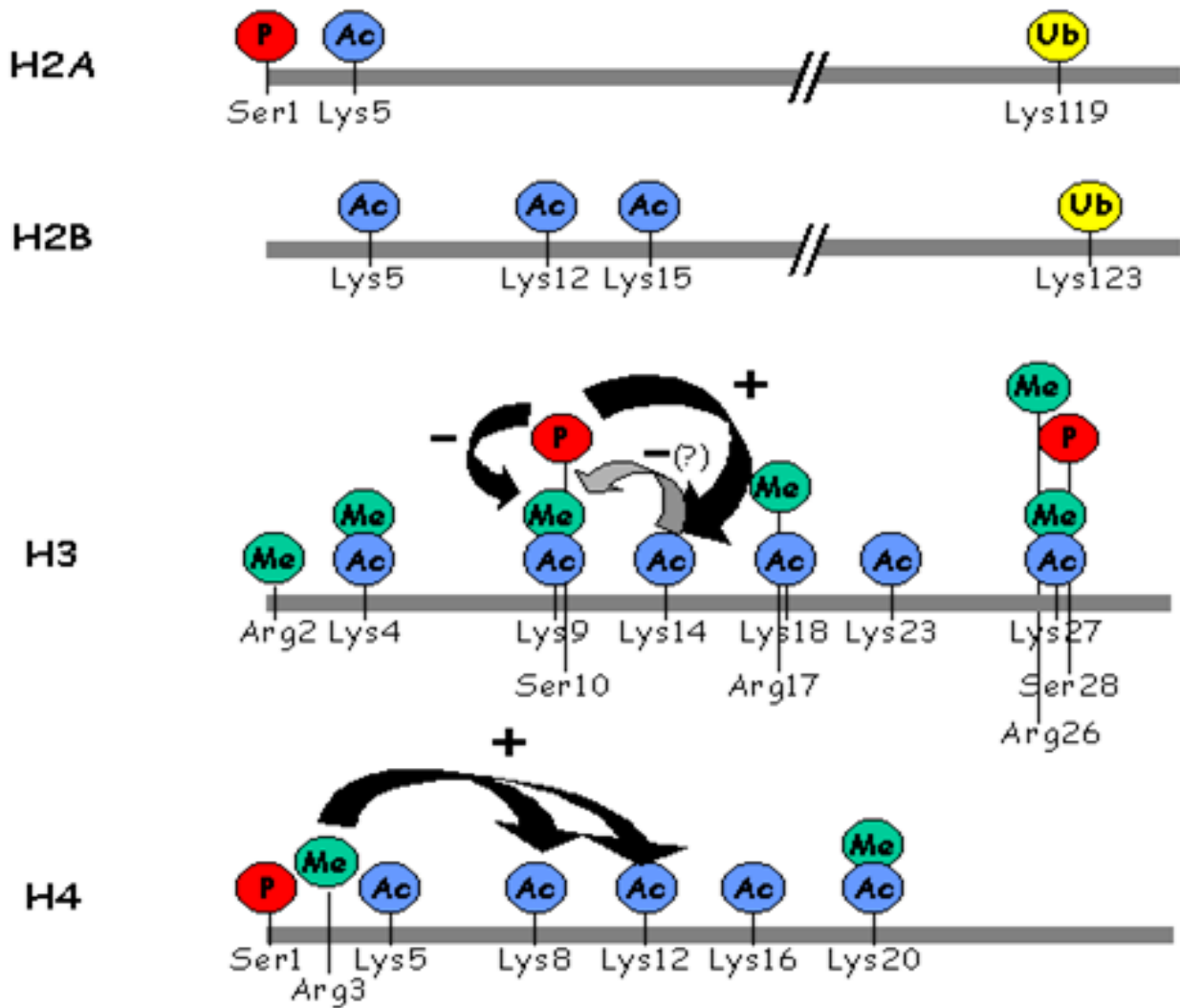


Figure 22 : Modifications covalentes des histones. Représentation schématique des sites possibles d'acétylation (AC), de méthylation (Me) de phosphorylation (P) ou d'ubiquitinylation (Ub) des histones *in vitro*. Les connexions identifiées entre les différentes modifications sont représentées par des flèches, qui indiquent à la fois les résidus et le type de modification concernés.

4.4. Ubiquitinylation des histones

L'ubiquitinylation (modification biochimique post-traductionnelle qui permet la fixation covalente d'une ou de plusieurs ubiquitines -peptide de 76aa se trouvant dans toutes les cellules des organismes-) des histones est une autre forme de modification de la chromatine, qui pourrait affecter la transcription. Les formes ubiquitinyllées des histones H2B et H2A sont effectivement associées à la fraction transcriptionnellement active de la chromatine. La modification de l'histone H2A, a lieu sur le résidu Lys119 et celle du H2B sur le résidu Lys123. L'ubiquitinylation de l'histone H2B diminue à la fois son

affinité de liaison à l'ADN, l'association des histones H2A et H2B au nucléosome et la liaison de l'histone H1, favorisant ainsi la transcription en maintenant la chromatine dans un état relâché accessible.

4.5. Régulation par choix du promoteur

Certains gènes possèdent plusieurs promoteurs de force inégale et plusieurs sites d'initiation de la transcription.

Le choix du promoteur adéquat va dépendre de la présence et de l'activité de facteurs de transcriptions qui peuvent n'être présents que dans certains tissus.

Les mRNA produits seront alors différents, selon les types de cellules dans lesquelles elles sont produites.

Exemple: le gène de l' α -amylase comprend 2 promoteurs alternatifs P1 et P2. Le P1 est très actif, il est fonctionnel dans les glandes salivaires, par contre le P2 est peu actif, il est exprimé dans le foie.

4.6. Régulation de la durée de vie des mRNA

C'est une régulation post-transcriptionnelle. Il est connu que des séquences particulières présentes dans la région 3'UTR des ARNm sont responsables de la stabilité des mRNA. (1h à 20h). L'augmentation du degré de stabilité signifie une augmentation de la quantité d'ARNm et par conséquent la quantité de protéines produites et inversement.