

Modifications post-transcriptionnelles des ARN

I. Introduction

Les transcrits primaires possèdent « ppp » en 5'. Ils subissent une maturation :

- élimination de séquences 5', 3' et internes excédentaires ;
- modifications des nucléosides ;
- insertion ou délétion de nucléotides = édition.

On obtient un produit mature fonctionnel.

Chez les bactéries, le seul ARN qui ne subit pas de modifications est l'ARNm. Les ARNr sont synthétisés sous forme de transcrit primaire qui renferme les gènes de l'ARN 16S, 23S et 5S. Dans certains opérons codant pour l'ARNr, on trouve des gènes codant pour l'ARNt. Il existe des gènes isolés pour la synthèse d'ARNt mais il y a aussi des transcrits primaires qui possèdent des gènes pour 2 ARNt.

II. Maturation des pré-ARNt

La maturation des pré-ARNt fait intervenir de nombreuses ribonucléases (endo- et exonucléases). Les pré-ARNt possèdent des séquences en excès à leurs 2 extrémités.

A. Maturation des pré-ARNt chez les bactéries

La séquence CCA est déjà présente à l'extrémité 3'. La séquence excédentaire en 3' est reconnue par une ribonucléase E ou F qui coupe juste en amont de l'épingle à cheveux formant ainsi une nouvelle extrémité. Ensuite une exonucléase, la RNase D élimine les nucléotides excédentaires. La RNase P (particule ribonucléoprotéique) coupe le début de la feuille de trèfle formant l'extrémité 5' de l'ARNt mature. On obtient donc un ARNt mature avec une extrémité 5'P et 3'OH.

B. Maturation des pré-ARNt chez les eucaryotes

Le pré-ARNt possède une séquence excédentaire en 5' et un poly-U en 3'. La RNase P catalyse la coupure de la séquence excédentaire en 5' alors que le poly-U est éliminé par une exonucléase 3'-5'. Le codon CCA est ensuite ajouté par une ARNt-nucléotidyl transférase au niveau de l'extrémité 3'.

Dans le cas de la levure, il y a aussi un intron à éliminer, situé entre la boîte A et B du promoteur de ce gène. Il sera éliminé par des enzymes : endonucléases qui le coupe aux bornes, puis une endonucléase ligase ligature les 2 extrémités.

C. Les bases modifiées

Un large spectre de modifications chimiques a été identifié dans les différents transcrits primaires d'ARN : on en connaît au total plus de 50. La plupart s'effectuent directement sur un nucléotide à l'intérieur du transcrit. Beaucoup de ces modifications ont d'abord été identifiées dans les ARNt, dans lesquels environ un nucléotide sur dix est modifié.

Exemple de modifications chimiques :

- La **méthylation** : addition d'un ou plusieurs groupements méthyl sur la base ou le sucre. Par exemple la méthylation de la guanosine donne la 7-méthylguanosine ;
- La **désamination** : élimination d'un groupement aminé (-NH₂) de la base. Par exemple la désamination de l'adénosine donne l'inosine ;
- La **substitution d'un S** : remplacement d'un oxygène par un soufre. Par exemple la 4-thiouridine ;
- L'**isomérisation d'une base** : changement de position des atomes sur le noyau d'une base. Par exemple l'isomérisation de l'uridine donne la pseudo-uridine. ;
- La **saturation d'une double liaison** : transforme une double liaison en simple liaison. Par exemple la conversion de l'uridine en dihydro-uridine ;
- Le **changement de nucléotides** : remplace le nucléotide existant par un nouveau nucléotide. Par exemple la quéosine.

D. Caractéristiques des ARNt

Tous les ARNt possèdent :

- une **extrémité 3'CCA**,
- une **boucle tψC**,
- une **boucle variable**,
- une **boucle de l'anti-codon**,
- une **boucle dihydro-uridine**,
- un **bras accepteur** qui correspond à l'association des 2 extrémités de l'ARNt pour former le site de fixation de l'acide aminé.

Il joue un rôle dans le transport des acides aminés, rythmé par l'ordre des codons portés par le messenger. L'ARNt s'apparie au codon du messenger par l'intermédiaire de son anticodon. Au niveau de l'appariement, la **position de Wobble** induit une certaine flexibilité dans l'appariement de la 1^{ère} position de l'anticodon avec la 3^{ème} position du codon : l'inosine peut s'apparier avec C, A ou U et l'uracile peut s'apparier avec A ou G. Un ARNt peut reconnaître jusqu'à 3 codons différents.

Par exemple au niveau du messenger, le codon 3'-UAU-5' peut s'apparier avec l'anticodon 5'-AUU-3' (Ile) ou avec l'anticodon 5'-AUG-3' (Met). Si on modifie le codon en 3'-UAI-5', il ne pourra s'apparier qu'avec l'anticodon 5'-AUU-3' car I ne reconnaît pas G.

La présence d'une pseudo-uracile à la place d'une uracile ne permettra la reconnaissance que d'un seul codon alors que l'uracile peut s'apparier à deux nucléotides.

III. Maturation des pré-ARN ribosomiques

A. Maturation des pré-ARNr bactériens

Un opéron code pour 3 ARNr différents appelés 5S, 16S et 23S et certains pré-ARNr codent aussi pour des ARNt. Il existe en général plusieurs copies du transcrite l^{re} (7 pour *E. coli*).

Ces ARNr sont séparés dans le transcrite par des **espaceurs**, il y a une séquence excédentaire en 5' : **séquence leader** et une séquence excédentaire en 3' : **traîne**.

Un transcrit primaire unique contient donc les copies des trois ARNr. Il est donc nécessaire de le couper pour libérer les ARNr matures. Ces coupures sont faites par des ribonucléases : la **RNase III**, qui est une endonucléase, coupe le pré-ARNr en 5' et 3' des séquences de l'ARNr 16S et 23S et les **RNases P et F** coupent respectivement en 5' et 3' de la séquence de l'ARNr 5S. Ensuite des **exonucléases** éliminent les parties excédentaires pour produire les ARNr matures.

Au cours de la synthèse du pré-ARNr, des protéines vont se fixer pour empêcher la RNase III de couper.

B. Maturation des pré-ARNr eucaryotes

Chez les eucaryotes il y a 4 ARNr. L'un d'eux, l'ARNr 5S, est transcrit par l'ARN polymérase III et ne subit pas de maturation. Les trois autres (5,8S, 18S et 28S) sont transcrits par l'ARN polymérase I en un seul transcrit primaire qui, comme chez les bactéries, doit être coupé et dont les extrémités doivent être retaillées.

La maturation a lieu **dans le nucléole** et fait intervenir beaucoup de molécules comme les **snoRNP** (small nucleolar RiboNucleoprotein Particle) qui, associés à des protéines, jouent le rôle de guide pour la coupure du pré-ARNr par les endonucléases ; et des exonucléases qui éliminent les séquences excédentaires (exosomes).

Ce mécanisme est beaucoup plus complexe que chez les bactéries.

C. Modifications des bases des ARNr

Les ARNr subissent deux types de modifications : **addition d'un groupement méthyl** surtout à l'hydroxyl en 2' du sucre, et **conversion de l'uridine en pseudo-uridine**. Les transcrits primaires d'ARNr humains, par exemple, subissent 106 2'O-méthylations du sucre, 10 méthylations de base et 95 pseudo-uridinylations.

La méthylation du sucre est réalisée par une méthyltransférase dont l'action est assurée par des snoRNP qui servent de guide. En effet, des snoRNP permettent d'indiquer la place des méthylations sur le transcrit primaire par des appariements de base dans la région concernée. Ces appariements n'impliquent que quelques nucléotides, mais ils sont toujours situés immédiatement en aval d'une séquence du snoRNA. La modification a donc lieu au niveau de la partie hybridée.

IV. Maturation des pré-messagers

- Addition d'une coiffe
- Epissage (splicing)
 - o Le prémessager est colinéaire au gène ;
 - o Un messenger n'est plus colinéaire au gène.
 Il y a élimination des introns.
- Polyadénylation : coupure du transcrit primaire en 3' et addition d'une queue poly-A.

1. Addition d'une coiffe

L'addition de la coiffe est couplée à la transcription, cette réaction est catalysée par la guanylyl transférase et la phosphohydrolase qui sont apportées par le CTD de la polymérase II.

Il y a tout d'abord élimination du γ P en 5' du transcrit primaire par la *phosphohydrolase*, ensuite il y a transfert de la GMP au niveau de ce 5' par la *guanylyl transférase* avec libération de pyrophosphate (PPi). La *guanine-7-méthyl transférase* ajoute un méthyle au niveau du N7 de la guanine. Et pour finir il y a addition d'un méthyle sur le 2'-O des sucres des deux bases suivantes par une *2'-O méthyl transférase*.

Cette coiffe constitue le site de fixation de protéines qui reconnaissent le m7G dans le noyau.

2. Epissage des introns

Exon = séquence d'ARN maintenue dans l'ARN mature, attention elle n'est pas forcément codante ;

Intron = séquence d'ARN éliminée au cours de l'épissage du transcrit primaire. Ils peuvent servir dans la synthèse des ARN nucléolaires.

La taille des exons est en moyenne de 100 à 500 nucléotides, ils codent pour le domaine d'une protéine ; les introns peuvent faire jusqu'à 500 000 nucléotides.

Les séquences sont peu conservées au niveau du site de coupure : un pré-messager peut donner naissance à plusieurs messagers. L'épissage doit être très précis à cause du cadre de lecture qui risquerait d'être décalé.

Chez les mammifères il y a un **épissage alternatif** à cause des séquences peu conservées aux sites de coupure ce qui augmente la capacité codante du génome.

Chez les levures, seulement 4% des gènes possèdent des introns, ils sont de petite taille et situés en 5' du pré-messager. Dans leur cas il n'y a donc pas d'épissage alternatif.

Mise en évidence des introns dans le gène de l'ovalbumine de poulet :

Une expérience d'hybridation, entre de l'ADN génomique et le messager mature, a été réalisée. Le transcrit primaire du gène de l'ovalbumine s'étend sur environ 7 700 nucléotides ; cependant, l'ARNm qui sera traduit ultérieurement n'est long que de 1 900 nucléotides environ. En général, les transcrits primaires sont quatre à six fois plus longs que l'ARN traduit. Les transcrits primaires – également appelés collectivement ARN hétérogènes nucléaires ou hnRNA – comportent souvent des séquences qui sont perdues au cours de la maturation de l'ARN : ce sont les introns. Dans l'hybride ADN/ARN, les boucles présentes correspondent donc aux introns présents dans le génome qui ne peuvent plus s'hybrider. Le nombre de boucles observées est égal au nombre d'introns et le nombre de parties appariées est égal au nombre d'exons.

Utilisation des enzymes de restriction pour mettre en évidence la présence d'introns :

On utilise de l'ADNc double brin parfaitement colinéaire au messager mature. S'il possède 2 sites de restriction pour les enzymes E1 et E1 on obtient 3 fragments d'ADN qui migrent différemment.

Lorsqu'on coupe le gène avec ces mêmes enzymes, on obtient le même résultat lorsqu'il ne possède pas d'introns ; mais si celui-ci en possède le profil de digestion ne sera pas le même et le segment contenant l'intron sera plus long et migrera moins vite.

Caractéristiques des introns des pré-messagers :

Les comparaisons de séquences nucléotidiques d'introns du génome permettent d'établir des **séquences consensus**. Au sein des introns on retrouve toujours les deux dinucléotides GT en 5' de l'intron (site donneur d'épissage) et AG en 3' de l'intron alors que le reste de la séquence est plus ou moins aléatoire. Un **résidu adénosine** au centre de l'intron constitue le **point de branchement** et joue un rôle important dans le processus d'épissage.

Chez la levure, les séquences sont plus conservées et le point de branchement est constitué des nucléotides ATAAC.

Réactions de trans-estérification :

L'épissage commence par une **attaque nucléophile** de la liaison phosphodiester du donneur d'épissage de l'extrémité 5' de l'intron par l'hydroxyl en position 2' du résidu adénosyl conservé du point de branchement. Il se forme alors un **intermédiaire cyclique** ressemblant à un lasso et l'extrémité de l'exon 1 est libre. Ce groupement 3'-OH libre **attaque le site accepteur** d'épissage de l'extrémité 3' de l'intron ce qui libère l'ARN de l'intron et joint les extrémités des deux exons en une molécule continue.

Pour les gènes eucaryotiques, ces réactions s'effectuent le plus souvent dans un complexe appelé **spliceosome** qui permet le rapprochement des extrémités 5' et 3' de l'intron grâce aux ARN nucléaires. Le spliceosome est très dynamique et peut changer de conformation. Il contient également des hélicases et de l'ATP. Attention, il ne faut pas oublier que les snARN sont associés à des protéines pour former des snRNP.

Des expériences d'*extension d'amorce* ont été effectuées : comme pour le séquençage on hybride une amorce qui est utilisée par la transcriptase inverse (ADN polymérase ARN dépendante) pour synthétiser un ADN complémentaire à l'intron. Mais la taille de l'ADNc obtenu est très courte car les enzymes sont arrêtées par le lasso au point de branchement.

Le spliceosome est un ribozyme :

C'est un énorme complexe constitué de plusieurs snRNP (5) : le **snRNP-U1** renferme le snRNA-U1 impliqué dans la reconnaissance de la borne 5'-intron, le **snRNA-U2** s'apparie à la région contenant le point de branchement en présence de la protéine **U2AF** qui se fixe sur la région riche en pyrimidine située en 3'-intron devant le dinucléotide AG. Les **snRNP-U4-U5-U6** forment une triparticule qui intervient dans la mise en place du spliceosome : le

snRNP-U4 joue le rôle de transporteur uniquement, le **snRNP-U6** en association avec le **snRNP-U2** constitué le site catalytique (spliceosome actif), le **snRNP-U5** permet le maintien des extrémités 5' et 3' de l'intron à proximité l'une de l'autre.

Des réarrangements et changements de conformations ont lieu en présence d'ATP et de protéines au cours de la mise en place du spliceosome actif. Les **protéines SR** riches en sérine et arginine sont présentes dans le site de fixation à l'ARN et jouent un rôle de « ciment ».

Les **UsnRNA** présents au sein des UsnRNP sont des ARN riches en U. 5 d'entre eux sont impliqués dans le splicing : snRNA-U1, snRNA-U2, snRNA-U4, snRNA-U5 sont transcrits par l'ARN polymérase II et portent une coiffe en 5' différentes des messagers car elle est constituée d'un triméthyl-G. Ils sont associés à des protéines sm envers lesquelles des anticorps existent. Le snRNA-U6 est transcrit par la polymérase III et est coiffé d'un γ -monométhyl-phosphate.

Mise en place du spliceosome :

Le **snRNP-U1** reconnaît la borne 5'-intron et il y a appariement de base entre l'extrémité 5' du snRNP-U1 et l'extrémité 5'-intron. Deux nucléotides de l'exon sont aussi compris dans l'appariement.

La **protéine U2AF** se fixe au niveau de la région riche en pyrimidine en amont du dinucléotide AG.

Le **snRNP-U2** reconnaît le point de branchement et s'apparie dans cette région. Le résidu adénosyl n'est pas apparié et reste accessible pour la formation de la structure en lasso.

La **triparticule snRNP-U4-U5-U6** induit un changement de conformation au sein du spliceosome. Les bornes des introns se rapprochent dans l'espace et forment un spliceosome inactif. Le spliceosome devient actif lorsque le snRNP-U1 et le snRNP-U4 sont éliminés par un autre changement de conformation. Lorsque le snRNP-U1 se désapparie, le snRNP-U6 prend sa place au niveau de l'extrémité 5'-intron et s'apparie au niveau du snRNP-U2 qui est apparié avec le point de branchement.

Le cœur catalytique est donc activé lorsque le dinucléotide GU est apparié avec le snRNP-U6 qui interagit avec le snRNP-U2 qui est lui-même apparié au point de branchement.

L'intron éliminé sous forme de lasso sera débranché et dégradé. Certains introns portent des snoRNA et ne seront pas dégradés grâce à des protéines qui se fixent sur l'ARN d'intérêt.

Epissage alternatif :

Il est soumis à régulation et mis en place dans des conditions précises. Ça augmente la capacité codante du génome : 60% des pré-messagers humains sont épissés de façon alternative.

Soit les exons sont contigus, soit il y a une utilisation alternative des sites donneurs et accepteurs différents de l'ARN primaire qui produit un ARNm contenant ou ne contenant pas l'exon 2. L'utilisation de sites d'épissage internes à un exon conduit à d'autres protéines. De cette manière, quatre ARNm différents et donc quatre protéines, sont issus d'une seule population d'hnRNA.

Sélection différentielle des sites d'épissage :

Il existe des **signaux** au niveau des introns et des exons : ce sont des séquences activatrices ou inhibitrices.

- **Activateurs** = séquences qui permettent de reconnaître le site d'épissage en présence de protéines SR ;
- **Inhibiteurs** = séquences qui empêchent la reconnaissance des bornes 3' ou 5' d'un intron par les hnRNP.

Ce sont les protéines les plus importantes et abondantes qui induisent le mécanisme : c'est une balance entre les concentrations en protéines SR et hnRNP.

Si il y a plus de protéines SR que de hnRNP on aura un signal activateur et inversement.

Détermination du sexe chez la drosophile :

- *Synthèse de la protéine SXL*
Par un épissage alternatif, le pré-messager SXL donne naissance à 2 messagers différents.
Chez la femelle l'exon-3 qui portait le codon « stop » est éliminé, la protéine sera donc fonctionnelle car elle sera synthétisée dans son intégralité.
Chez le mâle cet exon existe toujours donc la protéine SXL synthétisée ne sera pas fonctionnelle.
- *Influence de la protéine SXL sur l'épissage du pré-messager codant pour la protéine transformer (tra)*
Chez la femelle : la protéine SXL peut se fixer par compétition en 3' du premier intron ce qui empêche la fixation du facteur U2AF (l'affinité de la protéine SXL pour ce domaine est plus grande que l'affinité du facteur U2AF). La borne 3' reconnue dans l'épissage est donc la borne 3' de l'intron suivant (l'intron 2) et l'exon 2 qui porte le codon « stop » est éliminé. Le messager mature est complètement synthétisé et traduit en protéine tra fonctionnelle.
Cette protéine s'associe avec la protéine tra2 pour former un complexe qui se fixe en 3' de l'intron 3 et qui est stimulé par des protéines SR. La fixation de ce complexe facilite la fixation du facteur U2AF qui permet la reconnaissance et donc l'épissage de l'intron 3. Le messager produit contient donc l'exon 4 qui porte le codon « stop » et est traduit en une petite protéine dsx qui réprime les gènes mâles spécifiques.
Chez le mâle la protéine SXL n'étant pas fonctionnelle, elle n'entraîne pas la formation de la protéine tra. Le facteur U2AF se fixe alors préférentiellement au niveau de la borne 3' de l'intron 4 et produit un messager traduit en une protéine plus longue que chez la femelle. Cette protéine dsx bloque la différenciation des gènes femelles spécifiques.

Ceci est un bel exemple de mécanisme d'épissage qui induit la reconnaissance d'un site ou non.

Maladie génétique due à un défaut d'épissage :

L'altération est due à un polymorphisme dans une région riche en pyrimidine, au niveau de l'intron entre les exons 8 et 9.

En 3' de cette région, lorsqu'il y a 9 fois le dinucléotide « UG » associé à 7 ou 9 résidus uridilique (U) il y a un épissage normal et l'exon 9 est maintenu dans l'ARN mature. Chez un malade, lorsque dans cette région il y a 13 fois le dinucléotide « UG » associé à 5 résidus uridiliques, il y a une perturbation de l'épissage qui induit la perte de l'exon 9. Cette perturbation est due à la fixation d'une protéine qui masque la borne 3' de l'intron.

3. Polyadénylation

Tous les pré-messagers présentent en amont de leur extrémité 3' le **signal de polyadénylation** AAUAAA situé à 15 ou 18 nucléotides en amont du site de coupure ; et en aval une région riche en dinucléotide « GU ». Sur chacun de ces sites un facteur se fixe : **CPSF** se fixe sur la région AAUAAA et **CstF** sur la région en aval du site de coupure et stimule le clivage.

Pour cliver l'ARN, il y a des facteurs de clivage tels que **CFI** et **CFII** qui se fixent au niveau des facteurs CPSF et CstF et coupent les messagers. Ensuite la poly-A polymérase (**PAP**) se fixe sur le complexe et ajoute un segment de poly-A d'environ 100 à 200 nucléotides. Des poly-A binding protein (**PAB**) se fixent sur le poly-A et jouent un rôle important dans la protection du messenger.

Cas de polyadénylation alternative :

Dans le cas des IgM par exemple, il existe 2 formes : la forme sécrétée la plus courte qui ne possède pas de domaine C-terminal et la forme membranaire qui possède le domaine C-term.

A partir d'un transcrit primaire qui renferme plusieurs sites de polyadénylation on peut synthétiser 2 protéines différentes par polyadénylation alternative.



Au niveau du pré-messager, il y a un site de polyadénylation avant et un site après l'exon qui code pour la partie C-term.

Dans les **cellules B**, les fonctions de polyadénylation se fixent préférentiellement sur le site de polyadénylation 2 parce que le facteur CstF est limitant et a plus d'affinité pour l'extrémité 3' de l'intron. Il y a donc une stimulation du facteur de polyadénylation à cette extrémité.

Dans le cas des **cellules circulantes ou plasmatisques**, le facteur CstF n'est plus limitant et se fixe aussi bien sur le site de polyadénylation 1 que sur le 2. Il y aura donc une coupure du pré-messager en amont du site qui code pour l'extrémité C-term.

Le poly-A est important pour la traduction : des expériences ont été réalisées chez le Xénope qui possède des messagers dormants dus à un poly-A raccourci. En effet, ce messager ne sera pas traduit mais il sera stocké en association avec des protéines. Au cours du développement, ces messagers se réveillent grâce à une polyadénylation cytoplasmique avec des facteurs spécifiques.

Les messagers d'histones :

- *Les messagers synthétisés constitutivement* proviennent de gènes avec des introns, ils possèdent une queue polyadénylée. Ces histones sont utilisées pour réparer la chromatine.
- *Les messagers synthétisés en fonction de la réplication* proviennent de gènes ne possédant pas d'introns et **ne possèdent pas de queue polyadénylée**. Ces histones sont utilisées pour compacter l'ADN néosynthétisé.

Sur ces messagers, il y a des signaux importants pour la mise en place des partenaires responsables de la coupure du prémessager : on peut observer une structure en tige-boucle en amont du site de coupure du prémessager ; et un motif HDE conservé en 3' du site de coupure et qui présente une séquence complémentaire et antiparallèle au snRNP-U7.

Le snRNA-U7 est présent sous la forme de particule ribonucléoprotéique et porte une coiffe triméthylée. Une fois qu'il est fixé en 3' du site de coupure, une autre protéine vient se fixer sur la structure en tige-boucle : ce complexe forme le signal du site de coupure. On peut donc dire que le snRNP-U7 joue le rôle de guide en s'hybridant au pré-ARN.

4. Epissage en trans

Ces mécanismes existent au niveau de certains prémessagers, mais ils sont rares. Le trans-splicing est une forme d'épissage qui joint 2 exons qui ne proviennent pas du même transcrit primaire : ce mécanisme est aussi une raison de l'augmentation de la capacité codante du génome.

Chez le Trypanosome, tous les messagers matures possèdent la même extrémité 5' constituée de 35 nucléotides et apportée lors d'un épissage en trans (trans-splicing). Ces 35 nucléotides sont apportés par le petit ARN SL.

5. Epissage des introns du groupe I et II

Ces introns s'éliminent par un processus catalytique, ils diffèrent par leur structure secondaires et présentent des régions conservées en séquence.

Leur auto-épissage se fait de la même façon que dans le spliceosome mais uniquement à partir de l'ARN : les introns du groupe I et II réalisent le même épissage que le spliceosome mais sans avoir besoin de protéines. Ce sont des ribozymes, c'est-à-dire des ribonucléotides à activité enzymatique.

- *Epissage des introns du groupe I*

Il est caractérisé par deux trans-estérifications. Le 3'-OH d'une guanosine libre ou d'un cofacteur (GTP, GMP, GDP) attaque le 5'-P du site donneur de l'intron ; ensuite l'hydroxyle de l'extrémité 3' de l'exon 1 attaque le P de l'extrémité 5' de l'exon 2 par réaction nucléophile et la seconde trans-estérification résulte de l'assemblage des deux exons.

- *Epissage des introns du groupe II*

Au niveau de la structure secondaire, on observe une ressemblance avec les mécanismes de trans-estérifications du spliceosome.

Le 2'-OH (OH en position 2' du ribose) du résidu adénylique présent au sein de l'intron attaque l'extrémité 5' du site donneur de l'intron ce qui forme le « lasso » (*en* : lariat) ; ensuite l'hydroxyle libéré à l'extrémité 3' de l'exon 1 attaque le 5'P de l'exon 2 ce qui provoque la deuxième réaction de trans-estérification qui permet de joindre les 2 exons et libérer l'intron.

6. Edition (editing)

Sous la pression de sélection de l'évolution, les cellules ont inventé constamment de nouvelles « astuces » pour produire un protéome plus important à partir d'un génome donné. La modification post-transcriptionnelle de la séquence nucléotidique d'un ARNm est appelée **édition de l'ARN**. Les protozoaires comme les Trypanosomes, agents de la maladie du sommeil et de la maladie de Chagas, utilisent intensivement l'édition de l'ARN, afin d'**inclure** ou de **supprimer des nucléotides** uracylés dans l'ARNm. L'édition peut modifier les instructions contenues dans un ARN de manière ciblée et conférer de nouvelles propriétés à une protéine.

L'insertion ou la suppression d'uridine peut être soit due à la présence d'enzymes seulement, soit à la présence de petits **ARN guides** (ARNg) qui s'hybrident en 3' de la région de l'ARN à modifier.

L'édition de l'ARNm du gène humain de l'apolipoprotéine B en fournit un bon exemple de l'édition : le gène de cette protéine code pour un polypeptide de 456 acides aminés appelés apolipoprotéine B100, synthétisé dans les cellules hépatiques et à travers le corps. Une protéine proche, l'apolipoprotéine B48 est synthétisée par les cellules intestinales. Cette protéine n'a qu'une longueur de 2153 acides aminés et est synthétisée à partir d'une version « éditée » le l'ARNm de la protéine complète. Dans les cellules intestinales, cet ARNm est modifié par la désamination d'une cytosine, la transformant en uracile.

Chez certains mammifères, une enzyme **modifie l'ARN** et entraîne l'apparition d'un codon stop dans le messager et induit la synthèse de deux protéines différentes.

L'épissage alternatif, la terminaison variable de l'ARN par la polyadénylation alternative et l'édition de l'ARN sont donc des variations sur un seul et même thème : la diversification de la palette des produits par un traitement différentiel des précurseurs.