LA TRANSCRIPTION

Dans une cellule, eucaryote ou procaryote, l'ARN provient donc d'une polymérisation de ribonucléotide utilisant l'ADN comme matrice. En effet d'une manière générale l'ARN est incapable de se répliquer (il existe des exceptions dans certains virus qui possèdent une réplicase RNA polymérase RNA dépendante).

I) Les différentes sortes d'ARN

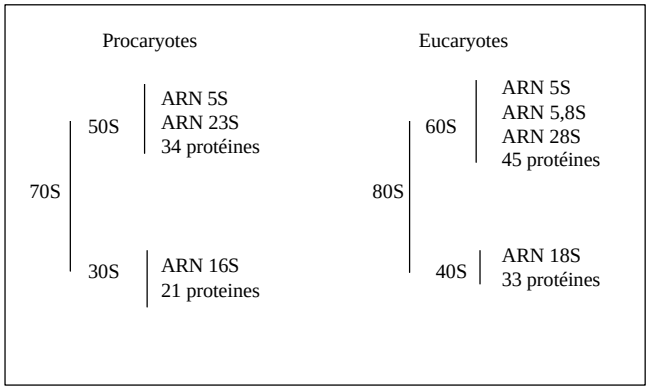
L'ARN diffère de l'ADN par une base (uracyle à la place de la tymine) et par le sucre (ribose à la place de désoxyribose). L'ARN est simple brin mais il existe des structures secondaires, alors que l'ADN est généralement sous forme double brin.
Ces différences entre l'ADN et l'ARN permettent une dégradation différentielle : l'ARN doit être dégradé (régulation de l'expression) alors que l'ADN doit rester stable.

1) ARN ribosomique

L'ARN qui compose des ribosomes représente environ 82% des ARNs totaux et 25 % des composants d'une bactérie. Les ribosomes sont composés d'ARN et de protéines répartis selon la figure.

L'ARN 5.8 S est présent uniquement chez les eucaryotes, la séquence correspond à l'extrémité 5' du 23S des procaryotes.

La différence entre les deux ribosomes, eucaryote et procaryote, peut être exploitée : certains antibiotiques agissent au niveau des ribosomes bactériens sans léser les ribosomes eucaryotes.



Les ribosomes des organelles sont différents des ribosomes du cytosol, ils ont diverses formes.
Les plus gros ont presque la taille des ribosomes bactériens et les plus petits sédimentent à 60S

Structure des ARN ribosomiques : Les ARNs ont de nombreuses structures secondaires, par exemple dans l'ARN 16S de *E. coli*, 50 % des bases forment des liaisons avec une autre base. Quoique on puisse théoriquement prédire la structure secondaire à partir de la séquence primaire, il y a de nombreuses possibilités pour de grandes molécules. Par exemple pour l'ARN 5S qui fait uniquement 120 b, 20 modèles ont été initialement proposés sur la base de la séquence primaire.

Il existe plusieurs méthodes pour choisir entre ces modèles parmi lesquelles on peut noter : 1) La comparaison de séquences entre des organismes proches, les régions qui sont importantes dans la structure secondaire sont conservées. 2) La mutagenése dirigée permet de voir si une liaison entre deux bases est importante, on mute une base et on regarde l'effet, puis on mute l'autre bases pour restaurer la complémentarité 3) La résolution de la structure trimentionelle.

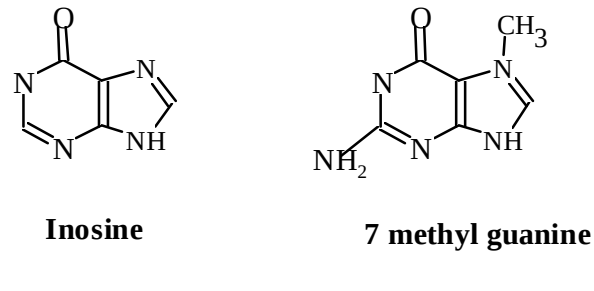
La conformation du RNA change avec le binding du tRNA, du mRNA ou avec l'association des deux sous unités.

2) Les ARN de transfert.

Ils représentent environ 15% des ARN d'une cellule

Ce sont de petits ARN servant à transporter les acides aminés du cytosol jusqu'au ribosome. Ils sont constitués d'une chaîne qui se replie en forme de feuille de trèfle si on considère uniquement la structure secondaire et en forme de L en réalité dans l'espace.

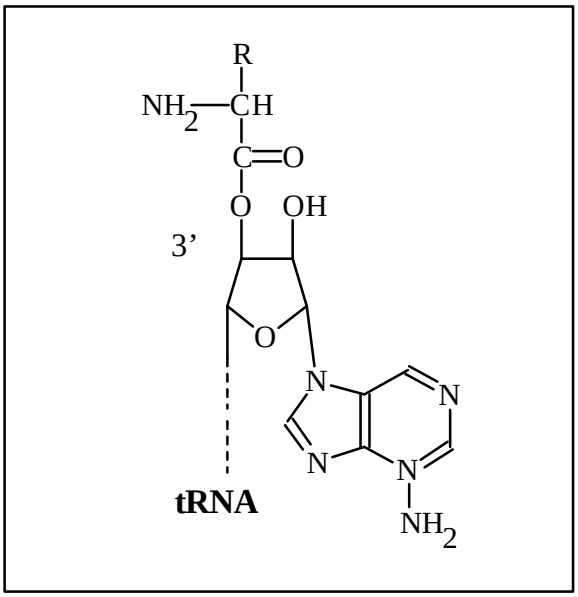
Ils sont constitués d'une chaîne unique et contiennent de nombreuses bases inhabituelles (7 à 15 par molécules). Ces modifications proviennent de modification enzymatique d'un ARNt précurseur. Par exemple l'inosine provient de la déamination oxydative de l'adénine.



Inosine **7 methyl guanine**

Site de fixation de l'acide aminé : Tous les ARN de transfert se terminent par CCA, c'est sur cette adenosine en 3' de l'ARN de transfert que sera fixé l'acide aminé à transporter par une liaison ester. Il est fixé soit sur le 3'OH soit sur le 2'OH de l'adenosine.

L'anticodon : La liaison au codon se fait avec une séquence anticodon c'est à dire qui porte la séquence complémentaire en formant un appariement de type Watson-Crick comme dans l'ADN.



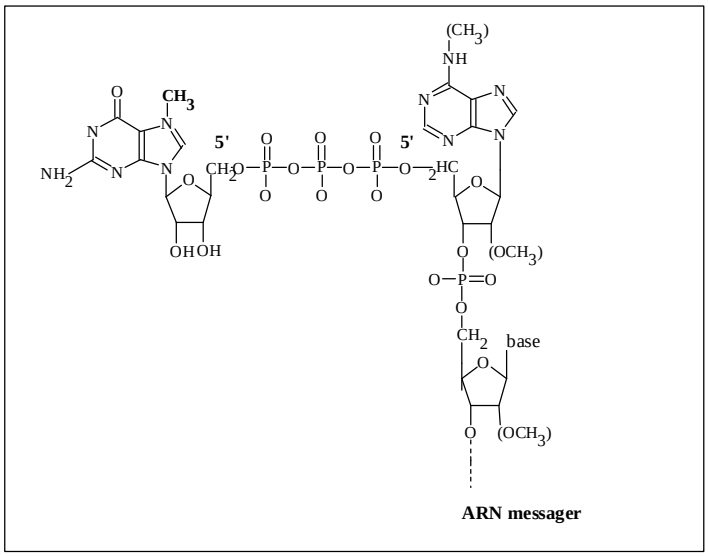
3) Les ARNS messagers. Leurs noms vient du fait qu'ils portent le message du gène. Ils sont de tailles très variables aussi sur un gel d'électrophorèse, ils forment une trainée contrairement aux ARN ribosomiques et de transfert. Dans une cellule, à un moment donné, il y en a environ 10 000 différents. Ils constituent une faible proportion des ARN d'une cellule (entre 1 et 5%). On distingue une partie codante (ORF) et une extrémité non codante en 5' et une extrèmité non codante en 3'.

Différence procaryotes / eucaryotes

Il y a plusieurs différences qui viennent du fait que chez la bactérie, transcription et traduction sont simultanées dans un seul compartiment alors que chez les eucaryotes, les deux processus sont spatialement et temporellement distincts.

* Nombre de parties codantes. Chez les procaryotes, l'ARN code pour plusieurs protéines différentes, il y a plusieurs phases de lecture ouverte. La reconnaissance par le ribosome s'effectue sur un site en amont de l'ATG d'initiation (Shine-Dalgarno box, site de liaison du ribosome, RBS). Chez les eucaryotes, l'ARN code pour une seule protéine (qui peut après être coupée lors de sa maturation). La reconnaissance de l'ARN par le ribosome est sur le Cap.

* Extrémité 5' : Chez les procaryotes l'extrémité 5' n'est pas modifiée, c'est donc un triphosphate. Chez les eucaryotes, après une synthèse d'environ 30 nucléotides, il y a addition de guanosine lié par une liaison 5'-5' au premier nuléotide. Cette guanosine est ensuite méthylée en position 7. De plus il y a d'autres méthylations sur les deux premiers nucléotides de l'ARN, ces dernières méthylations ne sont pas toujours présentes. Le cap a un rôle dans l'initiation de la synthèse des protéines et dans la stabilité du transcrit en cours d'élongation



* Extrémité 3' : procaryote : dernière base. Eucaryote : queue poly A.

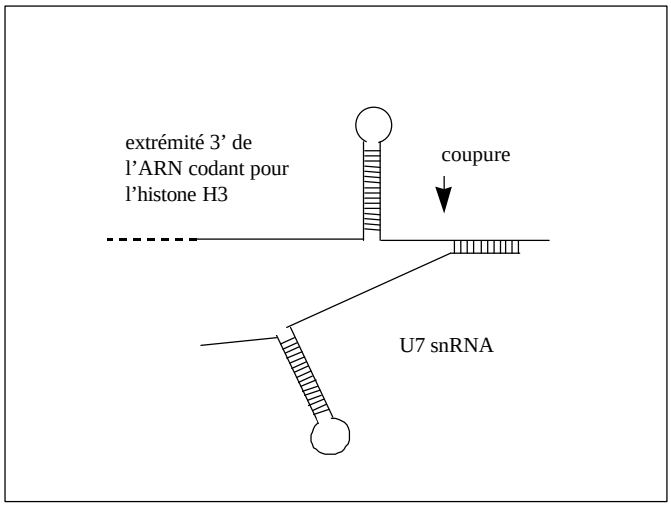
4) Les petits ARN régulateurs

Des interactions ARN-ARN sont impliquées dans la régulation de la réplication (voir l'exemple de la régulation du nombre de copies d'un plasmide) et de la traduction (voir chapitre suivant)

5) Les ribonucleoprotéines

a) Les petits ARN nucléaires

Le noyau cellulaire contient de nombreux complexes composés de protéines et de petits ARN (250 b au plus). Ils ont été dénommé U1, U2.. U12. Ils sont associés à des protéines pour former des petites particules nucléaires ribonucleoprotéiques (snRNP).



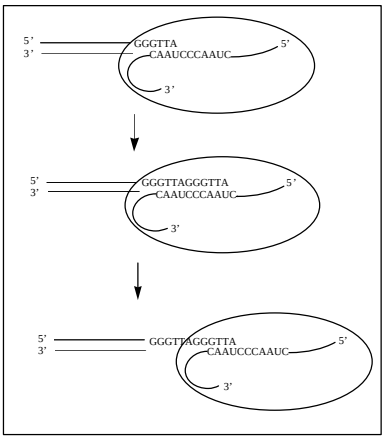
Leur rôles : ils agissent en s'hybridant avec d'autres ARNs : certains interviennent dans l'épissage des introns, d'autres dans le clivage des extrémités 3' de certains ARNs, d'autres dans la maturation des ribosomes et enfin certains ont une fonction encore inconnue. Ils ont donc pour rôle en s'hybridant de marquer une région de l'ARN où il y aura une maturation.

Exemple du rôle des petits ARN nucléaires : Clivage de l'extrémité 3' de certains ARNm : les ARN messagers codant pour les histones ont une particularité, ils ne sont pas polyadénylés, il y a une coupure en 3' qui nécessite un appariement avec une particule ribonucleoprotéique (snRNP U7).

b) les ARNs liés à des protéines

Il y a d'autres ARN qui sont liés à des protéines

C'est le cas de l'ARN 7SL. Cet ARN est une des composantes de la SRP (particule de reconnaissance du signal) avec 6 protéines (voir le cours de Biologie Cellulaire). Lorsque le ribosome fabrique une protéine destinée à être transloquée dans le réticulum, la SRP se fixe sur le ribosome et arrête la traduction. Le ribosome lié à la SRP se fixe au récepteur de la SRP sur la membrane de réticulum endoplasmique. La SRP est déplacée et la traduction continue.



Un deuxième exemple est la télomérase. C'est une ribonucléoprotéine qui catalysent l'addition de motifs répétés à l'extrémité des chromosomes de séqunece TTAGGG. L'ARN présent dans la télomérase sert de matrice pour la synthèse de l'ADN. La télomérase a donc une activité reverse transcriptase (la plupart des cellules somatiques n'ont pas d'activité télomérase et le chromosome se raccourci peu à peu alors que les cellules embryonnaires, les cellules souches ou les cellules cancereuses ont cette activité).

6) Les ribozymes

Jusqu'à récemment, on pensait que seules les protéines pouvaient catalyser des réactions. Cette idée venait du fait que la variabilité de la structures tridimensionnelle et la variabilité des résidus des

protéines qui pouvait générer un grand nombre de sites actifs. Mais quelques réactions catalytiques viennent de l'ARN, on a donc créé le terme de ribozyme.
On peut prendre comme exemple la ribonuclease P qui contient un ARN (375 bases) et une protéine (20 kDa), Cette RNAse est impliquée dans la maturation des ARN de transfert. Elle coupe le précurseur en 5' du tRNA mature. L'ARN est capable de couper le tRNA (mais moins vite), la protéine joue donc un rôle secondaire probablement en maintenant la structure de l'ARN.

Site catalytique: dans les protéines, le site catalytique est formé de chaînes latérales d'acide aminé, chargé positivement ou négativement, hydrophile ou hydrophobe. Dans l'ARN, la variabilité des résidus est plus restreinte et est en fait limitée aux 4 bases. Les ribozyme utilisent l'ARN comme substrat, il y a dans ce cas utilisation de l'appariement entre les bases homologues.

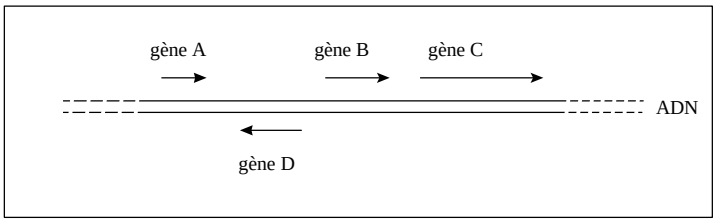
La découverte de ces ribozymes a fait penser que l'ARN était la molécule primitive à l'origine des êtres vivants. Ainsi l'ARN aurait précédé l'ADN et les protéines. Cette hypothèse est étayée par le rôle central de l'ARN.

- Il peut se fabriquer par assemblage spontané de nucléotides.
- Il est capable de se répliquer par appariement préférentiel: la séquence d'un petit oligonucléotide peut servir de matrice pour la polymérisation d'un autre polymère et donc reproduction de l'original. (phénomène très important car il est alors possible de conserver un avantage)
- Il a un pouvoir catalytique sur lui même (auto-excision, auto-epissage, auto élongation....)
- Il peut servir de matrice à l'ADN

II) Mécanisme de la transcription

Structure d'un gène

On peut définir un gène comme une région de l'ADN qui est transcrite en ARN. A cette région on rajoute les parties en cis qui interviennent dans la transcription. Cette définition met en fait en valeur une définition fonctionelle, des mutations dans cette région, sur la partie transcrite comme sur les parties régulatrices affectent l'expression de la protéine codée par le gène.



Une seule des 2 chaînes de l'ADN sert de matrice pour générer un transcrit mais la transcription peut avoir lieu sur des brins différents pour générer plusieurs protéines. Comme la polymérase avance de 5' vers 3', un changement de brin implique un changement de sens. Chez *E. coli*, le sens est pour la plupart du temps celui de la fourche de réplication.

La transcription permet la conversion en ARN de l'information contenue dans l'ADN. Elle fait intervenir l'ARN polymérase, un promoteur, des facteurs d'initiation et d'élongation et des signaux de terminaison.

ARN polymérase

1) Se lie à un site spécifique de l'ADN (promoteur) sur une zône d'environ 60 nucléotides.

2) Ouvre la double hélice
3) Initie la polymérisation par association des deux premiers ribonucleotides triphosphates.
4) Se déplace sur le brin matrice. L'allongement d'un nucléotide dans le sens 5' vers 3' à une vitesse moyenne de 30 nucléotides/seconde.
5) Referme la double hélice à l'arrière

Il y a donc deux sites dans l'ARN polymérase, un site de déroulement et un site de réenroulement. L'ARN polymérase bactérienne est une enzyme de grande taille (> 500 kDa) à sous unités multiples, associées à plusieurs sous-unités protéiques supplémentaires qui se lient temporairement au complexe polymérase ADN selon les étapes de la transcription.

La polymérase de *E. coli* a 4 sous-unités (2 α, β, β'et σ). On ne sait pas encore pourquoi l'ARN polymérase est si complexe, certaines ARN polymérases de virus bactérien sont plus simples avec une seule chaine polypeptidique avec une masse 5 fois plus faible, et elle marchent aussi bien. Vraissemblablement certaines sous unités sont impliquées dans des processus de régulations qui ne sont pas encore définis

Chez les eucaryotes, il y a trois ARN polymérases qui peuvent être différentiées par leur fonctions et par leurs sensibilités à l'amanitine (de l'amanite phalloide notamment).

ARN pol I fabrique les gros ARN ribosomaux, insensibles à l'amanitine, son activité transcriptionelle est responsable du nucléole .

ARN pol II fabrique les ARNm et la majorité des petits ARN qui forment les snRNP (impliqués dans l'épissage des ARN), elle est sensible à l'amanitine

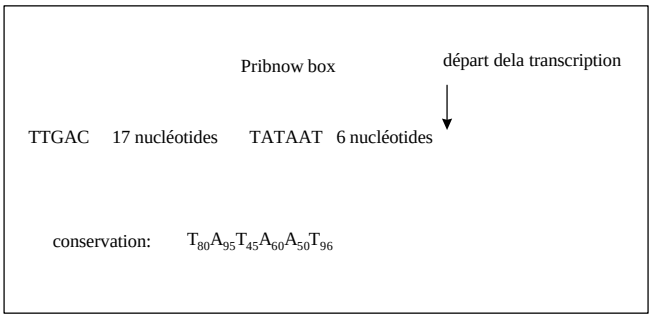
ARN pol III fabrique les petit ARN ribosomaux (5S) et les ARN de tranfert.

Le promoteur

- Signale à la polymérase où elle doit se lier (et de ce fait, détermine le sens de la transcription).
- Contient le site de départ (ou d'initiation) de la transcription

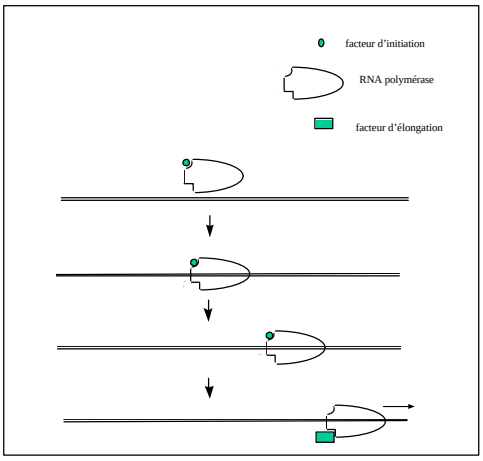
Chez *E. Coli*, il y a deux séquences consensus hexanucléotidiques

(notion de séquence consensus : ces séquences sont les plus fréquentes, 100 signifie que le nucléotide est toujours rencontré alors que 25 signifie qu'il est présent une fois sur 4, sa fréquence est alors due au hasard).

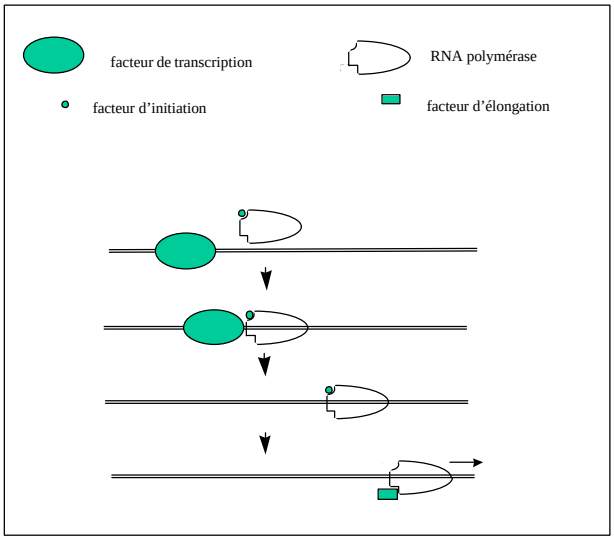


Par convention, on représente les séquences sur la chaîne non transcrite, identique à la chaîne d'ARN.

Initiation et élongation



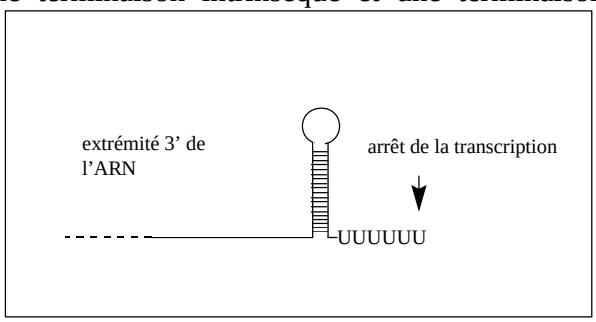
Chez les procaryotes, il existe des facteurs d'initiation de la transcription (Chez *E. coli*, la sous-unité σ a ce rôle). Ils se fixent sur la RNA polymérase et permettent à l'enzyme de trouver les séquences consensus des promoteurs. Puis il se détache (après la synthèse d'environ 8 nucléotides).D'autres protéines, les facteurs d'élongation (ex: nusA chez coli), se lient au complexe et sont nécessaires pour l'élongation de la chaîne.



Chez les eucaryotes, les ARN polymérases ne reconnaissent pas les promoteurs sans la présence de facteurs de transcription, TF I, TF II ou TF III selon qu'ils servent les polymérases I, II ou III. Il en existe plusieurs appelés II A, IIB etc... selon l'ordre de leurs découvertes. Ils se fixent sur l'ADN indépendemment de l'ARN polymérase (ils sont différents des facteurs d'initiation qui se fixent sur l'enzyme). Par exemple le TFIID (facteur TATA) se fixe sur la boite TATA (séquence conservée riche en A T) qui se situe 25 bp en amont du site de la transcription.

Terminaison

Au niveau d'un signal de terminaison, la polymérase libère la matrice et l'ARN monocaténaire néosynthétisé.



Chez *E. coli*, ily a deux sortes de terminaison, une terminaison intrinséque et une terminaison dépendante d'une protéine, la protéine ρ. Dans la terminaison intrinséque, la polymérase s'arrête quand elle synthétise une série de résidus U dans la mesure où elle a d'abord synthétisée une séquence autocomplémentaire capable de se replier. C'est le repliement, l'hélice en épingle à cheveux, qui se forme rapidement dans cette région, qui est cruciale pour l'arrêt de la transcription. La séquence de la région autocomplémentaire peut varier.

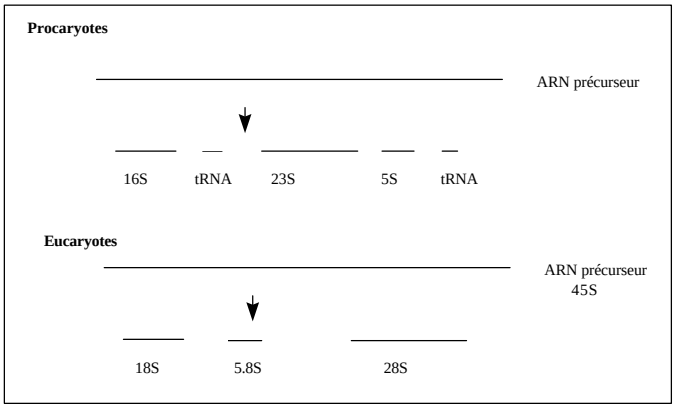
Lors de la terminaison ρ dépendante, la protéine ρ reconnaît une séquence sur l'ARN néosynthétisé et catalyse l'arrêt de la transcription.

Chez les eucaryotes, il y a arrêt a la fin du premier gène (il n'y a en général pas de gene dit polycistronique comme chez les procaryotes)
- Pour la RNA polymérase I (ribosome), la terminaison se situe à un site 1000 b après l'extrémité 3'
- La RNA polymérase II n'a pas d'arrêt de transcription défini. Il y a transcription plus d'un kb après un site dit de polyadénylation (AAUAA). Ce signal est nécessaire à la coupure (10-30b après) puis à la polyadénylation par un endonucléase puis par un poly(A) polymérase.
- La RNA polymérase III (tRNA) a un arrêt de la transcription à un site qui correspond a celui de l'ARN mâture (2ème U de 4 U qui se suivent dans une région riche en GC)

III Maturation des ARNs

ARN ribosomiques

Chez les bactéries,
Il y a 7 opérons qui ne sont pas en cluster, entre le 16S et 23S il y a un ou plusieurs tRNA (variable selon les opérons)



Chez les eucaryotes
La biogenèse des ribosomes a lieu dans le nucléole. Pour synthétiser les ribosomes, il y a généralement plusieurs gènes en effet il n'y a pas d'étape d'amplification. La plupart des protéines les plus abondantes (hémoglobine, myoglobine ..) sont synthétisées à partir d'un seul gène. La grande quantité de protéine est due au fait que les ARNm transcrits à partir du gène peuvent être traduits en plus de 10 protéines / min. plus l'ARN est stable, plus il y aura de protéine.
Les gènes codant pour les ARN ribosomiques sont regroupés et répétés en tandem et forment de grandes boucles d'ADN d'ARNr (10 chez les cellules humaines c'est à dire 5 chromosomes différents).
Les gènes d'ARNr sont transcrits (ARN polymérase I) sous la forme d'un d'un ARN 45 S précurseur. Ce précurseur est intégré à un complexe de 70 chaînes polypeptitiques (provenant du cytoplasme), Il est ensuite coupé pour donner 3 ARN (18S, 5.8S et 28S)
l'ARN 5S est lui aussi répété en tandem mais n'est pas associé aux 3 autres, il n'est pas fabriqué dans le nuclèole et il y est importé pour s'associer aux ribosomes.

Les deux sous unités du ribosome se séparent et passent dans le cytoplasme, là il y a réassociation entre les deux sous unités et des étapes de maturation du ribosome. Cette dernière étape de maturation dans le cytoplasme empêche d'avoir des ribosomes fonctionnels dans le noyau qui pourraient traduire des ARNhn (prémRNA)

Les ARN des transfert

Les gènes codant pour les tRNAs se trouvent le plus souvent en cluster et répétés à l'exception de tRNA du tryptophane (cet opéron est donc indispensable). Chez les procaryotes, ils sont transcrits en tant que

précurseur qui contiennent plusieurs tRNAs. Les tRNAs prennent leur structure secondaires et sont séparés par la RNase P qui coupe à l'extrémité 5' (cette enzyme est composée d'ARN et de protéine, les deux sont important pour la catalyse) une autre RNase (sans doute RNase III) permet de raccourcir l'extrémité 3'
Il existe 2 types de tRNA: Type 1 ont le CCA dans la séquence. la RNase coupe juste après le A, type II le CCA n'existe pas et est rajouté par une enzyme , la tRNA nucleotidyl transférase.

Bases modifiées dans les ARN ribosomiques et les ARN de transfert.

Il y a un très grand nombre de bases modifiées dans les ARNs de transfert, il y a des modifications dans tout les ARN mais d'habitude c'est uniquement de la méthylation, dans les ARN de transfert comme dans les ARN ribosomiques, il y a des modifications importantes:

Exemple : l'inosine se trouve comme intermédiaire normal dans la synthèse des purines mais n'est pas incorporé directement, elle résulte de la déamination oxydative de l'adénine

Les modifications habituellement impliquent l'altération de la base incorporée sauf dans le cas de la queunosine (Q) ou il y a remplacement du nucléoside complet en coupant de chaque coté et en refaisant les ponts. Ces modifications ont un rôle dans l'appariement, plus particulièrement dans l'anticodon, ils augmentent ou diminuent le wobble (voir plus loin)
exemple : l'inosine s'apparie avec U, C et A et la thiouridine s'apparie uniquement avec A

Maturation de l'ARNm

Chez les eucaryotes comme chez les archéobacteries, les ARN immédiatement transcrits ne sont que des précurseurs. Ils sont appelés ARNnh (nucléaires hétérogènes). Certaines parties sont retirées (épissage), elles sont appelées introns. Les parties restantes sont appelées exons.

a)Ils subissent le capping

Après une synthèse d'environ 30 nucléotides, il y a addition et méthylation d'une structure composée de guanosine liée par une liaison 5' 5' au premier nuléotide. Cette modification est donc cotranscriptionelle.
Ce cap a deux rôles
- dans l'initiation de la synthèse des protéines, c'est le cap qui est reconnu par le ribosome.
- dans la stabilité du transcrit en cours d'élongation.

b) La queue poly A

La RNA polymérase II n'a pas d'arrêt de la transcription bien défini. Il y a transcription plus d'un kb aprés un site de polyadénylation AAUAAA, puis coupure 20-30 bases après ce site par une endonucléase puis ajout de la queue poly A par une polyA polymérase.

La plupart des ARNm eucaryotes possèdent donc une longue série de résidus adénosine (env. 200). L'utilité de la queue polyA n'est pas encore bien connue, on suppose qu'elle protège de la dégradation et qu'elle permet le transport vers le cytoplasme.

Utilisation technique : purification des RNA messagers sur colonne d'oligo dT.

c) Epissage

Le transcrit primaire est appelé ARN hétérogène nucléaire ou hnRNA. C'est une copie fidèle de l'ADN (à la coiffe et la queue poly A près). Le transcrit mâture est plus petit (500-3000 bases) et généralement code pour une protéine.

Donc, chez les eucaryotes, une grande partie du transcrit initiale (ou du gène) ne se retrouve pas dans l'ARN messager et donc n'est pas codante. Les parties qui se retrouvent dans l'ARN méssager sont les exons, les parties qui seront excisées et détruites sont les introns. Les exons seront assemblés : c'est l'épissage. Il existe une très grande variabilité dans le nombre d'exons selon les gènes. Certains en ont une cinquantaine alors que d'autre n'en ont pas.

Cet ensemble d'événements se déroule dans le noyau de la cellule : l'ARN ne passe dans le cytoplasme que sous la forme mâture.

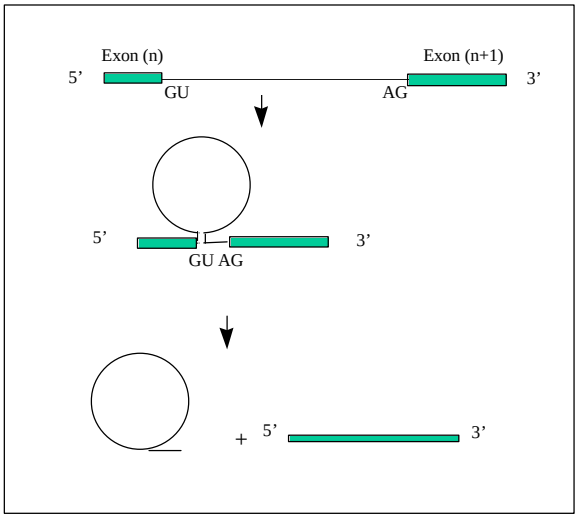
Mécanisme de l'épissage:

Les introns commencent et finissent par des séquences consensus.

La séquence consensus à l'extrémité 5' de l'intron commence toujours par **GU**, celle à l'extrémité 3' finit toujours par **AG.**

$A_{64}G_{73}$ / $G_{100}U_{100}A_{62}A_{68}G_{84}U_{69}$ .....12 pyC_{65} $A_{100}G_{100}$ /

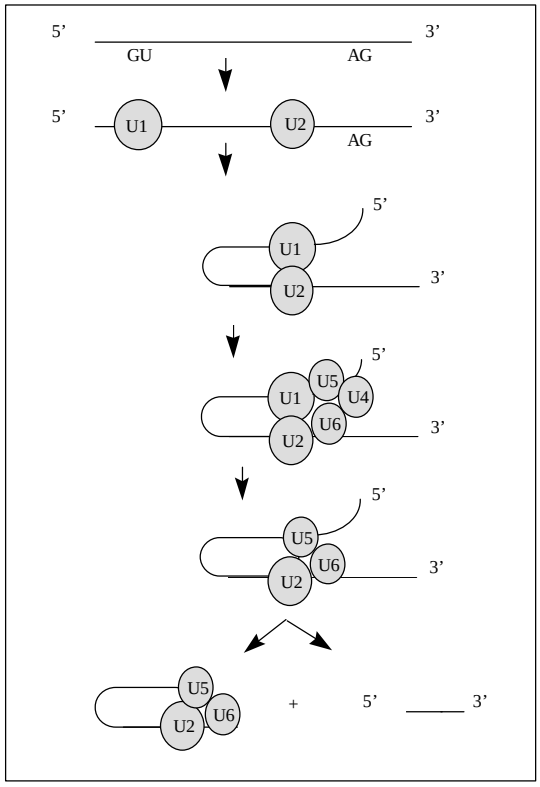
L'intron est donc orienté: le site en 5' est appelé «site donneur » et le site en 3' est appelé «site accepteur ». Le site est consensus pour tous les introns, il n'y a pas de spécificité due à l'exon, on peut mettre un intron entre deux exons de gènes différent, il sera excisé. Ainsi tout les sites donneurs pourrait interagir avec tout les sites accepteurs, mais en fait ça ne se passe pas.



Les introns sont éliminés sous l'aspect de molécules en forme de lasso. Il y a tout d'abord coupure au site donneur en 5'. Le bout 5 ' forme une sorte de lasso et se branche sur le 2' d'une adenosine environ 30 bases avant la fin de l'intron au niveau d'un site dit « site de branchement. Le rôle du site de branchement est d'identifier le site accepteur le plus proche. Il y a ensuite coupure au niveau du site accepteur puis ligation. Il y a donc chimiquement une transesterification, un pont ester est transféré d'un endroit à un autre.

Ce mécanisme est du à des snRNP (small nuclear ribonucleoproteins, complexe de protéines et de petits ARNs. Ils sont composés de petits ARN (de 250

nucléotides au plus) appelés ARN U1..U12 et forment avec les protéines un complexe qui ressemble au ribosome mais en plus petit. Ils s'associent à des séquences spécifiques de l'ARN pour former le complexe d'épissage ou "spliceosome". L'association est due en partie à des complémentartités entre sequences, si on mute l'ARNm il n'y a plus d'epissage, si on mute l'ARN U au site complémentaire on restaure l'epissage.



- U1 snRNP se lie au site 5', l'examen d'un des snRNA (l'ARN U1) a montré qu'il existait dans sa partie 5' une séquence complémentaire à la séquence consensus définie par l'extrémité 5' du site d'épissage . U1 snRNP s'hybride au site donneur, mais pas U1 seul, il faut les 8 protéines.
- U2 snRNP se lie au site de branchement (U2 a des séquences complémentaires).
- U1 snRNP et U2 snRNP réagissent ensemble, ce qui forme le lasso. Il y a ici une reconnaissance protéine/protéine.
- Un complexe comprenant U5snRNP et U4/U6 snRNP viennent se lier, tout est en place sans qu'il y ait la réaction
- U1snRNP se dissocie, U5 snRNP migre vers le site 3'
- U4 snRNP s'en va, U4 était partiellement hybridé à U6 (il le séquestrait) U6 maintenant libre interagit avec U2

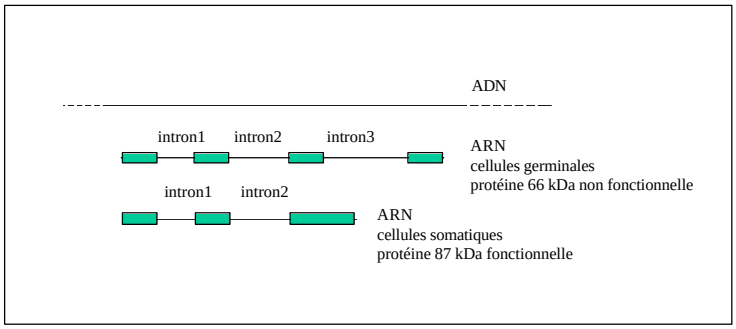
On pense que c'est cette dernière structure qui est responsable de la transestérification. Les snRNP interagissent entre eux et avec l' ARN m par formation de duplex, ces intéraction permettent des changements de structure ce qui créent des centres catalytiques

Ordre de l'épissage: lorsqu'il y a plusieurs introns dans un gène, il y a un ordre préférentiel de l'épissage, il semble que la composition de l'ARN influence l'accessibilité des sites de l'épissage. Lorsque l'intron est retiré, il y a un changement de conformation, un autre site devient accessible

Epissage alternatif:

D'une manière générale, l'appariement a lieu avec le site en 3' le plus proche sur la séquence linéaire, toutefois ce n'est pas toujours le cas.

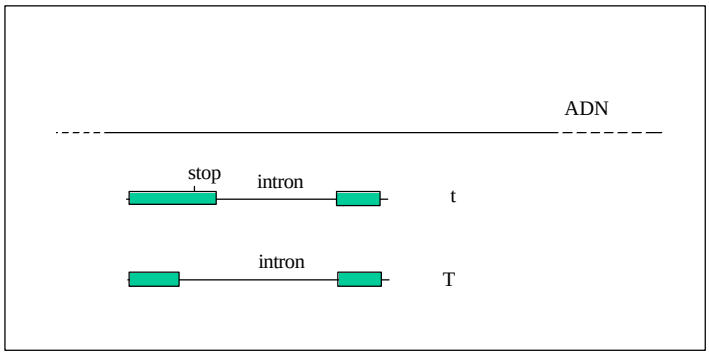
L'épissage alternatif sert souvent de regulation de l'expression au niveau postranscriptionel, en effet il permet de transcrire plusieurs protéines différentes à partir du même gène. Ces protéines sont différentes par exemple selon le tissus ou selon les stades du développement.



Exemple de la transposase de l'élément P de la drosophile. Dans les cellules somatiques, il y a une protéine qui inhibe

l'épissage de l'intron 3 en se liant une séquence de l'intron ce qui empêche l'association du spliceosome.

Souvent un seul des deux sites varie, c'estle cas par exemple de l'antigène T/t de SV40, là le site 3' ne change pas, seul le site en 5' est alternatif



la taille de la protéine ne dépends pas de la longeur de l'ARN, elle dépend de la longueur de la phase de lecture ouverte, si bien qu'un ARN plus grand peut donner une protéine plus petite. Dans le cas de l'antigène t, la protéine t est plus petite alors que son ARNm est plus grand que celui codant pour T. L'épissage dépend des cellules, ce qui donne plus de t par rapport à T dans certaines cellules. Dans ces cellules on a isolé une protéine SF2 qui se lie à l'ARN et qui promouvoir l'utilisation du site 5' le plus proche du site 3'.

Dans ces deux exemples on voit que l'épissage alternatif est due à des protéines qui se lient à l'ARN.

Autoépissage:

L'existence du lasso dans le mécanisme d'épissage a été déroutant lors de sa découverte. A priori, ce n'est pas le mécanisme le plus simple. Une explication est que ce mécanisme permet l'autoépissage ou plus exactement provient de l'autoepisage. L'autoépissage des introns de type II qui se trouvent dans les mitochondries de champignon ressemblent à l'épissage des gènes nucléaires dans le sens ou il suivent la séquence consensus GT....AG et il y a un lasso. Il y a dans ce cas une forte structure secondaire dans l'intron, avec 9 boucles peuvent être comparée à la combinaison entre U6 et U2 ou avec la combinaison de U2 avec le site de branchement. Il semble donc que le spliceosome du gène nucléaire ait évolué en compensant la perte progressive de la capacité d'autocatalyse.

Le trans-épissage
Dans ce cas le mRNA mature provient de la ligation de deux ARN, ce phénomène a été démontré chez *Trypanosoma brucei* et chez *Caenorhabditis elegans.* L'extrémité 5' de l'ARN est dans ce cas toujours la même pour tous les ARNm.

L'édition

Un dogme de la biologie moléculaire est que la séquence du mRNA correspond à la séquence de l'ADN (et que la séquence de la protéine correspond à la séquence du mRNA). Il exise toutefois une exeption, l'information change au niveau du mRNA
Exemple : chez la mitochondrie du trypanosome, on observe plusieurs bases qui sont systématiquement ajoutées ou délétées. La cytochrome oxydase II, il y a adjunction de plusieurs uridines. Il y a ici hybridation avec un ARN (RNA guide) et l'ARN messager qui va être édité, cet ARN guide est codé ailleurs.
Il existe aussi deux exemples chez les mammifères où on observe une substitution de bases. Dans l'apolipoprotéine-B, 4563 codons donnent une protéine de 512 kDA, mais dans l'intestin, on trouve une

protéine plus petite de 250 kDa, La séquence est identique, mais il y a un changement de C en U à la position 2153, CAA (glutamine) donne UAA (stop). Cette modification pourrait venir d'un épissage alternatif, mais il n'a pas été détecté, de plus cette édition a pu s'obtenir *in vitro*.
L'autre cas connu est celui du récepteur glutamate. L'édition change ici un CAG (Glutamine) en CGC (arginine). Cette modification change les propriétés électrophysiologiques du récepteur, change le flux ionique. Elle est donc importante au niveau physiologique. L'adénosine est déaminé en inosine. L'édition est donc due à une adénosine désaminase et une activité de reconnaissance de la structure.

V) Régulation transcriptionnelle chez les procaryotes

Intérêt de la régulation génétique : chez *E. coli* la transcription et la traduction consomment 50% de l'ATP produit. Il y a donc un énorme avantage d'avoir un système régulé pour gérer les ressources du milieu extérieur sans dépense d'énergie inutile.

Le contrôle de cette expression génétique peut se faire à plusieurs niveaux. Au niveau:
- de la transcription
- de la stabilité des ARN
- de la traduction (voir chapitre suivant)
- de l'activité enzymatique

1) Régulation au niveau de la transcription

Chez les procaryotes, la régulation de l'expression génétique se fait principalement au niveau de la transcription. On considère qu'il y a des promoteurs faibles, les gènes en aval de ces promoteurs sont peu transcrit, les protéines sont donc peu traduites et sont donc en faible quantité. En contrepartie, il y a des promoteurs forts.
Lorsqu'il y a régulation transcroptionelle, le taux de transcription peut être augmenté ou diminué en fonction de facteurs indépendants.

Chez les porocaryotes, l'expression de plusieurs gènes est gouvernée par le même système de régulation. Cet ensemble est appelé l'opéron, c'est l'unité d'expression génétique.

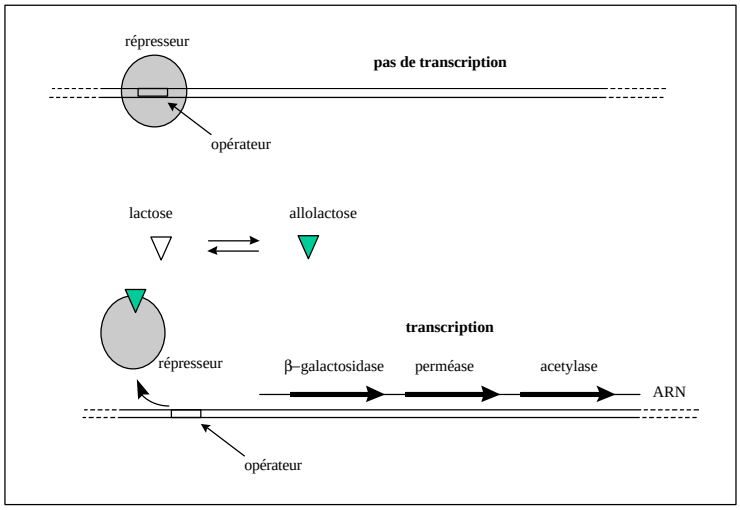
La régulation peut être due à des modifications de l'ADN

La transcription est souvent lié au taux de méthylation de l'ADN. La Dna A est codée par un gène qui n'est transcrit que si le gène est méthylé sur les deux brins par la Dam méthylase. Lorsque l'ADN vient d'être répliqué au niveau du gène codant pour la Dna A, l'ADN n'est pas encore méthylé, le gène codant pour la Dna A n'est pas transcrit et il n'y a donc pas de protéine. Comme la Dna A intervient dans l'initiation de la réplication en se fixant sur ori C, cette absence de Dna A participe à éviter qu'il y ait plusieurs réplications trop rapprochées, avant la séparation des deux cellules filles.

La régulation peut être due à des protéines qui se fixe en cis au niveau du promoteur

Exemple de régulation négative : l'opéron lactose

En l'absence de lactose, la bactérie ne fabrique pas les enzymes nécessaires à sa métabolisation (β-galactosidase, galactoside perméase et galactoside acétylase).



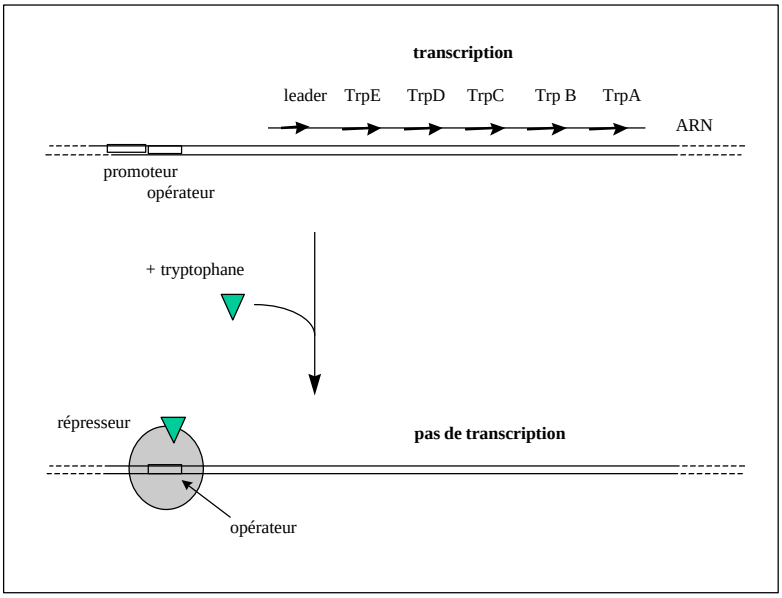
L'expression de ces gènes est gouvernée par le même sytème de régulation, ils sont sous la dépendance du même promoteur.

En présence de lactose, l'allolactose se forme dans la cellule grâce au peu de β-galactosidase présente dans la cellule, l'allolactose se fixe à un répresseur, induit un changement allostérique de la protéine qui se détache de l'opérateur, l'ARN polymérase peut se fixer sur le promoteur qui synthétise un ARN polycistronique codant pour les trois protéines

C'est une régulation négative, le répresseur fixé empêche la transcription.. Le répresseur est codé par le gène i qui est transcrit sans arrêt à un taux faible.

Exemple de régulation négative : l'opéron tryptophane

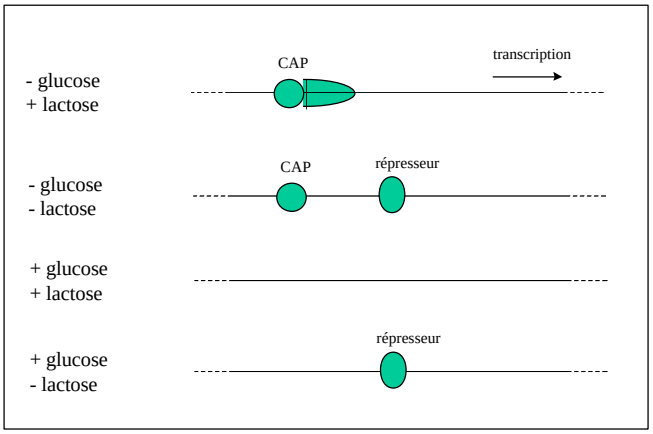
La synthése du tryptophane nécessite plusieurs étapes qui sont catalysées par 5 enzymes codées par les gènes TrpA, TrpB, TrpC, TrpD, et TrpE. Cet opéron contient en outre un promoteur, un opérateur et une séquence leader qui joue un rôle de régulation au niveau traductionel (voir plus loin). En l'absence de Trp, les 5 gènes sont transcrits, traduits et le Tryptophane est synthétisé. En présence de Trp, l'acide aminé se fixe sur le répresseur qui lui même se fixe sur l'opérateur et l'opéron n'est plus transcrit.



Il s'agit encore ici d'un régulation négative, le répresseur fixé empêche la transcription mais ici la fixation du ligand permet à la protéine de se fixer à l'ADN.

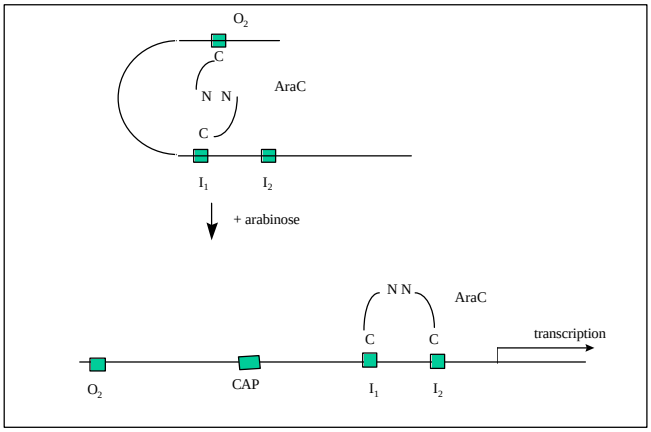
Exemple de régulation positive, la protéine d'activation catabolique (CAP).

A coté de la régulation négative, il existe aussi des régulations positives, dues à des activateurs : l'activateur favorise localement l'action de l'ARN polymérase. La CAP (protéine d'activation catabolique) est un activateur qui permet de n'utiliser d'autres sources de carbone qu'en l'absence de glucose chez *E. coli*. En l'absence de glucose, l'AMP cyclique induit un changement de conformation, la CAP se fixe sur l'ADN, l'opéron lactose est transcrit. Ce phénomène est connu sous le nom de répression catabolique



Exemple de régulation positive : l'operon arabinose

AraC est une protéine dimérique qui s'accroche sur deux sites distants de 210 bp en l'absence d'arabinose. Les deux protéines interagissent si bien que l'ADN forme une boucle qui réprime la transcription. Lorsqu'il y a de l'arabinose, il se fixe sur AraC. Il s'ensuit un changement conformationel. Le site distal n'est plus utilisé et AraC se fixe sur un autre site proximal. Dans cette configuration, AraC agit en tant qu'activateur transcriptionel.



2) Régulation par la stabilité de l'ARN messager

Il y a de grandes différences de stabilité des ARNm chez les eucaryotes

La stabilité peut dépendre :

de l'ARN synthétisée

> Les ARNs instables codent pour des protéines régulatrices dont le taux varie rapidement dans la cellule. Par exemple l'ARN de la β-globine a une durée de demie vie de plus de 10 heures, alors que l'ARN d'un facteur de croissance a une durée de demie vie d'environ 30 min.

du cycle cellulaire

.

> Par exemple les ARN codant pour les histones qui ne sont pas adénylés, mais présentent en 3' une structure en "tige et boucle". Leur instabilité est modulée par la synthèse d'ADN par la cellule. Leur demie vie est d'une heure lorsque la cellule synthétise de l'ADN (phase S) et de 10 min. en l'absence de synthèse.

de la localisation tissulaire

La stabilité différentielle peut être cellule spécifique, être responsable de la localisation tissulaire de certaines protéines

Généralement cette stabilité dépend des séquences non traduites en 3'. Pour analyser ce phénomène on transfère les séquences du gène analysé sur un gène rapporteur tel que celui codant pour la β-globine. Si on transfère l'extrémité d'un gène codant pour une histone, l'ARN devient plus stable. Par contre si on ajoute une séquence riche en A et U présente dans les ARN instables comme ceux codant pour un facteur de croissance, l'ARN devient instable.

V Transcription *in vitro*

On utilise la transcription in vitro dans les techniques de biologie moléculaire pour obtenir un ARN.

L'ADN d'intérêt, généralement un cDNA est cloné dans un plasmide. Un cDNA est une copie ADN d'un ARN messager. De chaque coté du site de clonage il y a un promoteur spécifique d'une RNA polymérase. On utilise généralement les promoteurs des ARN polymérase des phages T3, T7 ou SP6.
Dans la bactérie on peut propager ces plasmides, l'ADN inséré n'est pas transcrit en effet les promoteurs des phages ne sont pas reconnus par les RNA polymérase d'*E. coli*.

Si à cet ADN on rajoute des ribonucléotides (UTP, ATP, CTP et GTP, ou NTP), et de la RNA polymérase correspondant à un promoteur, on obtient un ARN. L'ARN a une taille indéfinie car l'ARN polymérase fait le tour du plasmide et n'a pas de raison de s'arrêter à un ppoint bien précis.
pour avoir un arrêt on peut soit introduire un site de terminaison de *E. coli* (une série de U suivant une structure secondaire) soit couper au préalable le plasmide juste après l'insert. C'est généralement cette dernière solution qui est envisagée.

Suivant l'orientation de l'insert et le promoteur utilisé, on peut obtenir un ARN sens, c'est à dire dans la même orientation que l'ARN normalement transcrit dans les cellules par le gène cloné. Mais, on peut aussi obtenir un ARN antisens, qui habituellement n'exite pas dans les cellules.

les ARN sens et antisens s'hybrideront avec l'ADN mais seul l'ARN antisens s'hybridera avec l'ARN.