

La transcription

Introduction :

La fonction capitale de l'ADN d'une cellule vivante est de contenir l'information génétique nécessaire au fonctionnement cellulaire. Ce fonctionnement est assuré, dans les cellules eucaryotes d'un organisme humain, par environ 30 000 protéines distinctes.

Le mécanisme général de formation des ARN qui seront nécessaires à la synthèse protéique nécessite :

- **ARN messenger (ARNm).**
- **ARN de transfert (ARNt) .**
- **ARN constitutifs des ribosomes (ARNr).**

Des études ont montré que le mécanisme de copie décrit chez les bactéries est valable aussi chez les eucaryotes à quelques nuances près. Cependant une différence fondamentale existe : l'absence de noyau chez les bactéries et sa présence dans la cellule eucaryote.

I) Définition de la transcription :

La transcription est le processus de copie du matériel génétique (ADN) en ARN. Chez les **procaryotes** une **seule ARN polymérase-ADN dépendante** effectue la transcription pour tous les types d'ARN, tandis que chez les **eucaryotes**, **trois** ARN polymérases (ARNpol) interviennent : l'**ARNpol I** pour les ARN ribosomiques transcrits dans le nucléole (28S, 18S et 5,8S), l'**ARNpol II** pour les ARNm, et l'**ARNpol III** pour les petits ARN (ARNt, ARNr 5S, ARNsn). Pour certains virus à ARN enfin, l'ARN est transcrit par une ARNpol-ARN dépendante appelée aussi réplécase.

Visualisation de la transcription

Le processus de la transcription peut être visualiser par microscopie électronique (première fois en 1970). **Les molécules d'ADN: filaments et les molécules d'ARN: Branches**



a) Les particularités des ARN polymérase sont :

- Qu'elles n'ont pas besoin d'amorces pour fonctionner.
- La polymérisation se fait dans le sens 5'-3', donc lecture du brin d'ADN dans le sens 3'-5'.
- L'ARN ne reste pas apparié au brin d'ADN pendant la synthèse.
- Les ARN polymérase n'ont pas d'activité de relecture (activité exonucléasique 3' → 5').
- La vitesse de synthèse est inférieure à celles des ADN polymérase (environ 50 nucléotides/seconde).

On peut d'emblée se poser une question : quel brin sera traduit ? Si l'on considère le brin d'ADN 3' → 5' antisens (transcrit), l'ARN lui sera complémentaire et antiparallèle.

Si l'on considère l'autre brin d'ADN en 5' → 3' (non transcrit, informatif), le brin d'ARN écrit en 5'3' lui sera identique, à l'exception toutefois de T remplacé par U.

Notons toutefois que lors d'une transcription, l'un des brins d'ADN est transcrit et l'autre pas, mais ce ne sera pas toujours le même brin. Dans une autre transcription, les rôles pourront fort bien être inversés (figure 1)

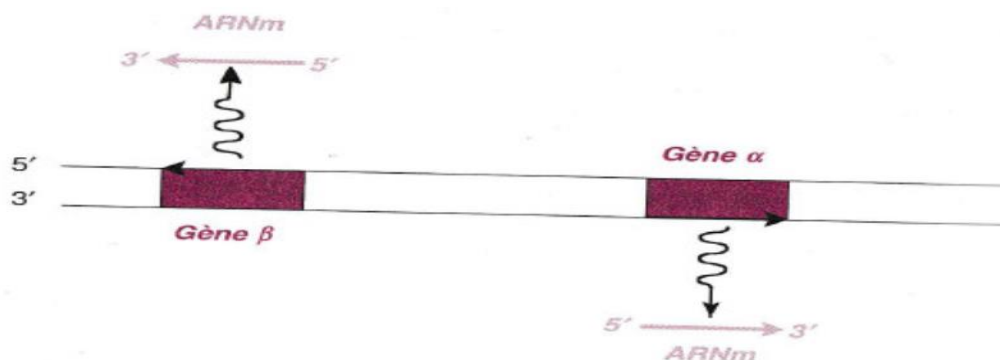


Figure 1 : Lecture des brins d'ADN lors de la transcription. Pour les gènes α et β , ce n'est pas le même brin d'ADN qui est lu, mais la lecture se fait toujours dans le sens 3' → 5', et la synthèse de l'ARNm de 5' vers 3'. (Beaumont S, 2007)

Gène :

Le gène est l'unité fonctionnelle de l'information génétique. Il est constitué d'un ensemble de nucléotides qui contient toutes les informations nécessaires pour transcrire un ARNm (figure 2)

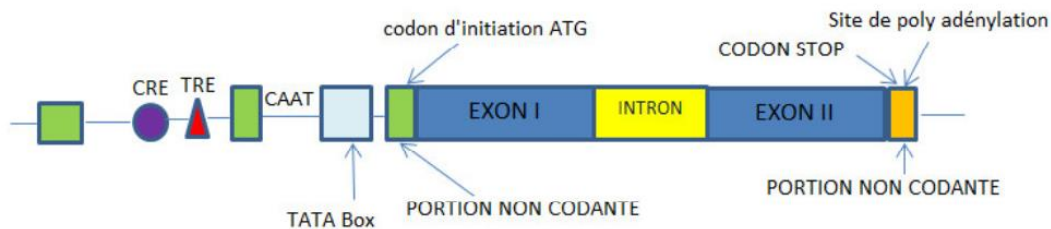


Figure 2: Structure général d'un gène chez les eucaryotes. (CRE [Cyclic AMP Responsive Element], TRE [12-O-Tetradacanolphorbol 13-acetate Responsive Element]: des éléments de réponse).

ARNm: Une succession de nucléotides sous forme de long filament a un seul brin. C'est un vecteur du message dictant la séquence des protéines entre chaque gène et la protéine correspondante.

Promoteur: Site sur l'ADN d'une centaine de bases auquel l'ARN polymérase (ainsi que les facteurs d'initiation requis) se fixe pour commencer la transcription. Chez E. coli il ya environ 2000 sites promoteurs (4.8×10^6 pb), elles sont qualifiés d'éléments cis

II) Transcription de l'ADN procaryote

La transcription s'effectue chez les bactéries grâce à une enzyme, l'ARN polymérase. C'est une enzyme de grande de taille (PM 450 000 Da) contenant 5 sous-unités ; chacune codée par un gène différent . Elle comprend la sous-unité σ et deux sous- unité α et un exemplaire de chacune des sous-unités β , β' (figure 3), et une sous-unité ω (n'est pas essentielle à la transcription mais elle a un rôle dans la stabilisation du complexe). La forme active de l'enzyme contient les sous-unités $\alpha 2 \beta \beta' \omega - \sigma$

L'ARN polymérase est une protéine ADN dépendante, **multimérique** possédant les sous unités α , β , β' et σ . Elle est présente sous deux formes **le core-enzyme** ($\alpha 2 \beta \beta'$) et l'**holoenzyme** ($\alpha 2 \beta \beta' \sigma$). Les ARN polymérases ne nécessitent pas d'amorce et ne possèdent pas d'activité exonucléasique et donc de correction d'erreur, le taux d'erreur est ainsi plus important que pour les ADN-polymérases, mais ce taux est supporté.

Les nucléotides triphosphates sont additionnés à l'extrémité 3' de la chaîne en cours de synthèse par complémentarité de la matrice d'ADN. L'hydrolyse de la liaison anhydride fournit l'énergie pour la synthèse de la liaison phosphodiester.

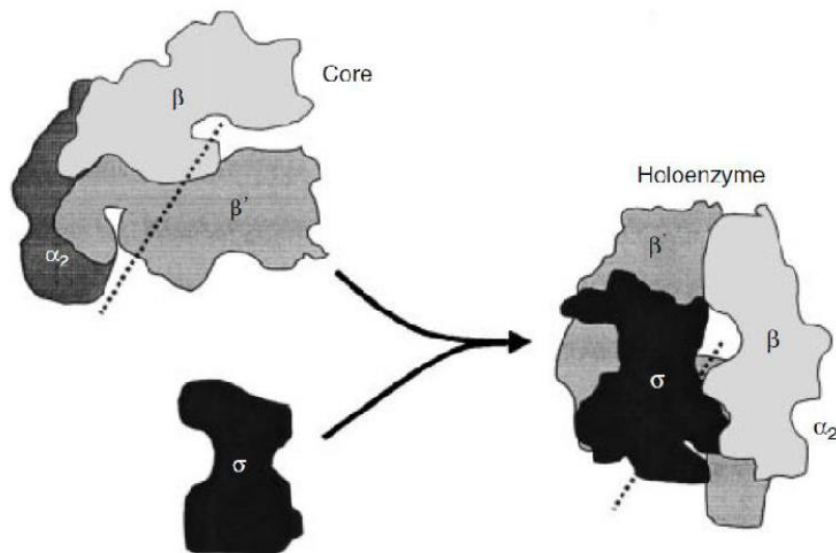
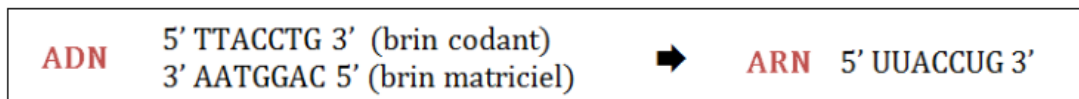


Figure 3 : Structure de l'ARN polymérase des bactéries.

La molécule d'ADN est composée d'un **brin matriciel** (ou **brin anti-codant**) dirigé par définition de **3' vers 5'** et servant comme son nom l'indique de matrice à l'ARN-polymérase, ainsi que d'un **brin codant** dirigé par définition de **5' vers 3'** et ayant une séquence identique à l'ARN transcrit mise à part le fait que la thymine est changée par l'uridine.

L'ARN messager est lui monocaténaire et dirigé tout comme le brin codant de 5' vers 3'



Chez E-coli, **une seule ARN-polymérase catalyse la synthèse de tous les ARN** de la cellule. La transcription chez les procaryotes ce fait au niveau du cytoplasme, elle est polycistronique car plusieurs gènes peuvent être transcrit par la même polymérase (exemple : L'opéron lactose). Par contre chez les eucaryotes, celle-ci se fait au niveau du noyau, et une polymérase transcrit un seul gène, on dit qu'elle est monocistronique (figure 4).

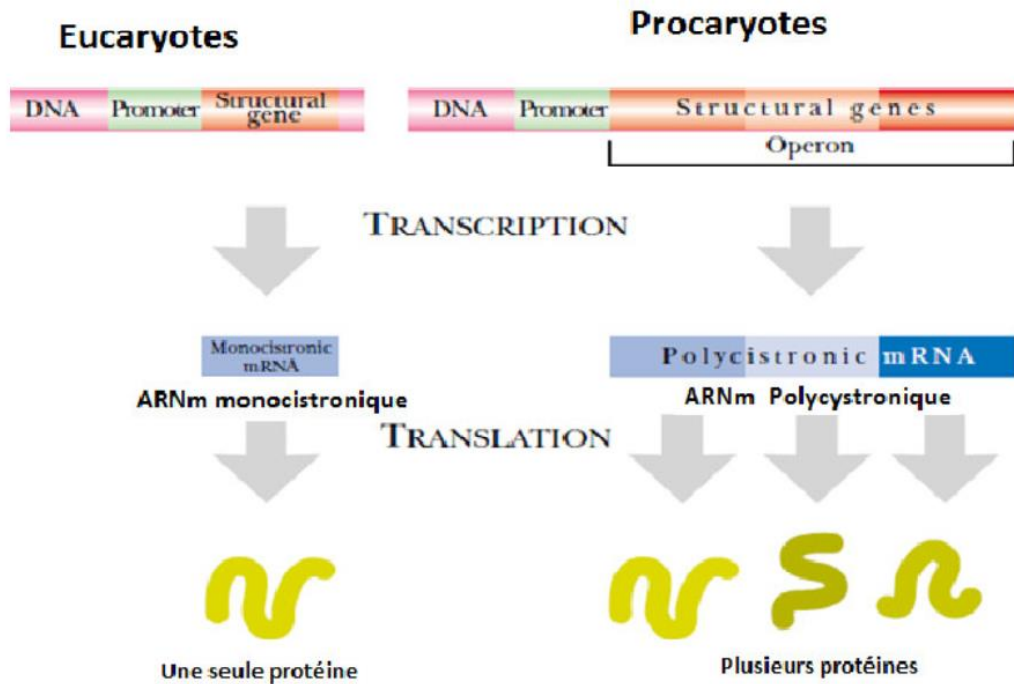


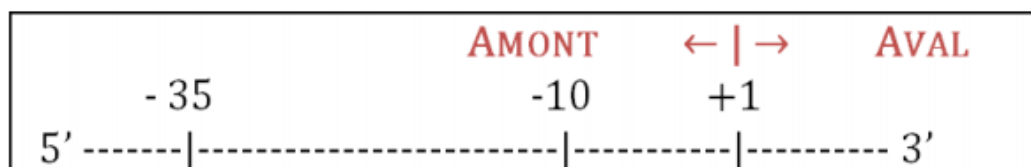
Figure 4: Synthèse monocistronique (eucaryotes) et polycistronique (procaryotes) de l'ARNm (Clark, 2005)

II.1 Étapes de la transcription chez les procaryotes

La transcription est divisée en 3 étapes : **l'initiation**, **l'élongation** et **la terminaison**. Le **promoteur** est une séquence d'une centaine de nucléotides située dans la région régulatrice et désignant le début de la transcription. Il est situé en **amont** du **site d'initiation** et porte des séquences reconnues par l'ARN-polymérase et déterminant le sens de la transcription.

Le promoteur est constitué de courtes séquences conservées d'une unité de transcription à l'autre et appelées **séquences consensus** :

- En **-10** du site d'initiation on trouve la **TATA box** ou **boîte de Pribnow** : «TATAAT»
- En **-35** du site d'initiation on trouve : « TTGACA »



L'affinité de l'ARN-polymérase pour l'ADN dépend de la forme de l'enzyme : le coreenzyme a une affinité faible et non spécifique, l'holoenzyme a une affinité très forte et spécifique pour le promoteur. On peut faire la remarque que la sous-unité sigma σ à l'état libre ne se fixe pas sur l'ADN. La sous-unité β' étant basique et l'ADN étant acide, ce sera elle qui facilitera l'interaction du complexe avec le promoteur.

La **sous-unité sigma σ** permet donc une reconnaissance spécifique du promoteur par l'ARNpolymérase et diminue l'affinité de l'enzyme pour les régions non promotrices. Il agit de manière cyclique, en effet après l'initiation faite, le facteur sigma se détache pour être recyclé et réutilisé pour d'autres initiations de gènes.

L'ARN-polymérase entraîne la **dénaturation des deux brins d'ADN** sur 14 paires de nucléotides, on parle de complexe ouvert qui augmente encore l'affinité de l'enzyme pour la double hélice.

II.1.A Initiation :

fixation de L'ARN polymérase sur les sites promoteurs.

- Chez toutes les bactéries, le démarrage de la transcription est conditionner par la reconnaissance de séquences spécifiques de l'ADN appelées sites promoteurs. L'analyse de nombreuses séquences d'ADN situées en amont du site de démarrage de la transcription, et d'autres expériences comportant l'étude de l'ADN après sa liaison à l'ARN polymérase (détermination des séquences d'ADN protégées par l'ARN polymérase) ont permis de définir des séquences « type » ou séquences consensus de liaison de l'ARN polymérase. Ces séquences constituent le site promoteur (Figure 5)..

Ces séquences sont appelées "éléments agissant en cis" (sur le même côté de la molécule d'ADN dans le gène lui-même). Contrairement aux "facteurs agissant en trans" qui sont des molécules qui se fixent à ces éléments d'ADN.



Figure 5 : Représentation du site promoteur des Procaryotes. Les chiffres indiquent la position des séquences par rapport à t'extrémité 3' de la TATA box prise comme référence. En gras apparaissent les nucléotides Les plus conservés. (Beaumont S, 2007).

L'initiation de la transcription se fait au niveau du promoteur. L'ARN polymérase réalise la polymérisation de ribonucléotides face à un bras d'ADN de 5 vers 3. Cette enzyme grâce à son facteur σ ; est capable d'assurer la reconnaissance de séquences spécifiques situées l'égerment en amant du point d'initiation de la transcription (+1).

La sous-unité σ se lie au site promoteur. Les séquences —10 et —35 sont reconnues par la sous-unité sigma σ de l'ARN polymérase. La sous-unité σ rentre en contact avec les deux séquences, mais semble également rentrer en contact avec une zone adjacente située sur l'autre brin.

Les points de contacts sont situés des mêmes côtés de la double hélice, ce sont essentiellement des groupements chimiques donneurs ou accepteurs de liaisons hydrogènes située dans le grand sillon. L'attachement se fait en deux temps :

- **D'abord à la région -35 ou une liaison relativement faible s'établit,**
- **puis à la région -10 ou une liaison très forte se produit. Cette deuxième liaison se traduit par une ouverture de la double hélice sur 15 paires de bases environ.**

L'initiation va durer le temps que la polymérase associe 7 à 8 ribonucleotides sous formes d'un polymère hybride au brin matrice.

Comme le core-enzyme présente une très forte affinité pour les hétéro duplex (brin d'ADN et brin d'ARN) le facteur σ se décroche et c'est le core-enzyme qui va seul continuer la transcription.

L'interaction de la **sous-unité β** est inhibée par la **rifampicine** et la **steptolydigne**.

Ces substances permettent donc d'inhiber le développement des bactéries en empêchant la synthèse des protéines qui leur sont nécessaires lors de leur croissance. Une mutation dans la sous unité β induit l'apparition de souches bactériennes résistantes à la rifampicine.

Le déroulement des premières étapes de la transcription est donc :

1. liaison non spécifique de l'holoenzyme.
2. formation d'un complexe fermé au niveau du promoteur
3. formation du complexe ouvert (déroulement sur 14 nucléotides)
4. Mise en place du premier nucléotide (très souvent A ou G)
5. Allongement de 7 à 8 nucléotides
6. Détachement du facteur sigma, après la transcription des premiers nucléotides.

Des points de contact supplémentaires entre l'ARN polymérase (les autres sous-unités) et l'ADN se forment, et la transcription va pouvoir commencer. La sous-unité σ se séparera alors de l'enzyme (dès que la chaîne d'ARN aura quelques nucléotides) pour aller se lier à une autre molécule d'ARN polymérase.

Sans (ou avec) la sous-unité σ , l'ARN polymérase a une affinité faible pour l'ADN ; elle « glisse » donc sur l'ADN jusqu'à ce que la sous-unité σ reconnaisse le site promoteur, ce qui facilite la reconnaissance du site promoteur.

La sous-unité σ se trouve d'un côté de la double hélice. C'est cette orientation qui va permettre la copie du brin matrice orienté 3' --> 5'. C'est ce qui explique que l'ARN polymérase ne copie qu'un seul brin. Il y a donc reconnaissance de la TATA box dans le sens 3' --> 5', pour que l'ARN polymérase polymérise l'ARN dans le sens 5' --> 3'.

C'est aussi l'orientation du site promoteur qui explique que dans une molécule d'ADN bactérien, l'une ou l'autre des chaînes sera copiée.

Des séquences d'ADN situées plus loin, entre —50 et —150 nucléotides, sont aussi très importantes pour que l'activité du site promoteur soit maximale (elles influencent la vitesse de l'initiation).

II.1.B Elongation

Le core-enzyme continue la polymérisation toute en déroulant l'ADN en aval et en le ré-enroulant une fois la séquence copiée. En amont du site de polymérisation on trouve un hybride de 17 paires de bases (figure 6).

Les conditions de la formation de l'ARNm par l'ARN polymérase sont : (Simon Beaumont, biologie moléculaire, Dunod, Paris 2007)

- présence d'une matrice d'ADN
- présence des 4 ribonucléotides (UTP, ATP, CTP, GTP)
- présence d'ions métallique Mg^{2+} , ou Mn^{2+}

L'élongation est inhibée par des aminosides ou amino-glucosides.

Parmi les sous-unités de l'ARN polymérase, c'est la sous-unité β qui joue le rôle catalytique, alors que β' servirait essentiellement à maintenir la liaison entre l'ADN et l'ARN polymérase.

Remarque :

1. Les premiers nucléotides de la chaîne d'ARN formée, ne sont pratiquement jamais ceux que l'on retrouve dans l'ARN traduit.
2. Des nucléases enlèvent quelques nucléotides ; parfois un clivage plus éloigné se produit. Ce qui sera important, en fait, c'est le codon d'initiation pour la traduction, et les séquences d'ARN permettant l'interaction avec le ribosome.

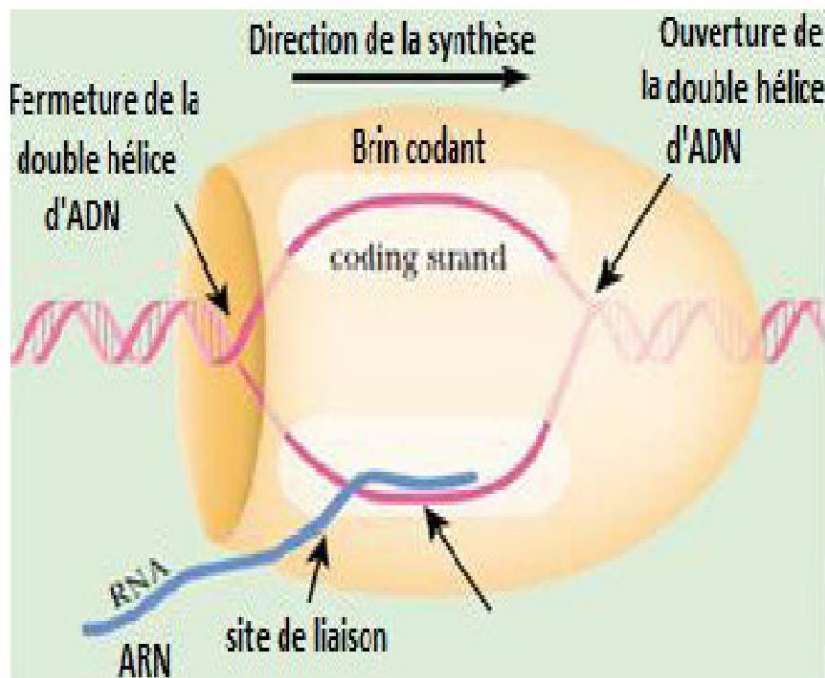


Figure 6: Elongation et direction de la synthèse

II.1.C Terminaison

La terminaison de la chaîne est assurée par des séquences particulières appelées **signaux de terminaison**. Un signal de terminaison est précédé par des sites dits de pause où la polymérase sera ralentie. Ces sites sont, soit des séquences **riches en G-C**, difficiles à ouvrir, soit des structures à caractère **palindromique**, c'est-à-dire présentant une symétrie.

On en distingue deux types :

- Signaux ayant une structure d'ADN particulière comportant un centre de symétrie d'ordre 2.

Ces séquences particulières, lorsqu'elles seront transcrites en ARNm, permettront à celui-ci de se boucler sur lui-même puisque ses séquences seront complémentaires et pourront donc s'apparier grâce à des liaisons hydrogènes.

Quand cette structure en épingle à cheveux se forme (Figure 7), elle correspond à une dissociation du complexe ADN-ARN sur environ 20 nucléotides. Il reste alors une séquence poly U liée à l'ADN. C'est un hybride très instable qui pourra se dissocier immédiatement de l'ADN (donc tout l'ARN est dissocié de l'ADN).

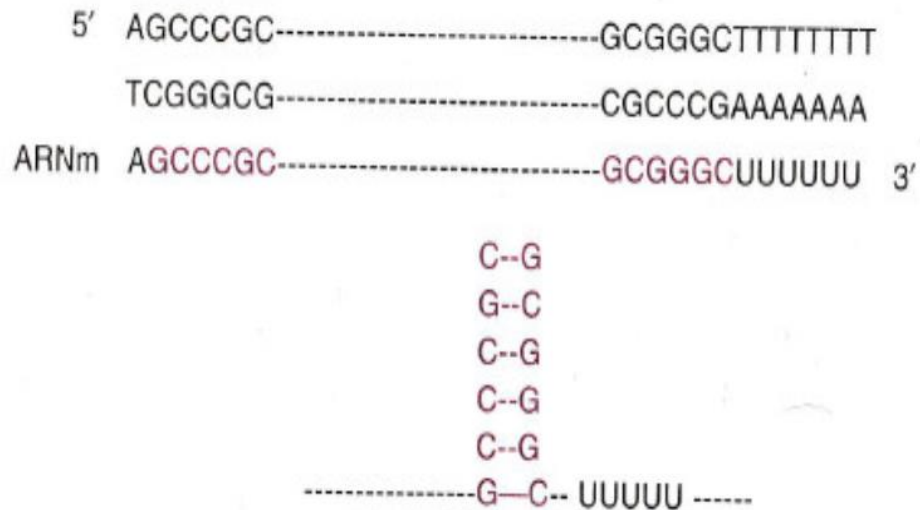


Figure 7 : Terminaison de transcription chez les Procaryotes

En haut les deux brins d'ADN, dont l'un est transcrit en ARNm qui contient des nucléotides capables de former la structure en épingle à cheveu. (Beaumont S, 2007)

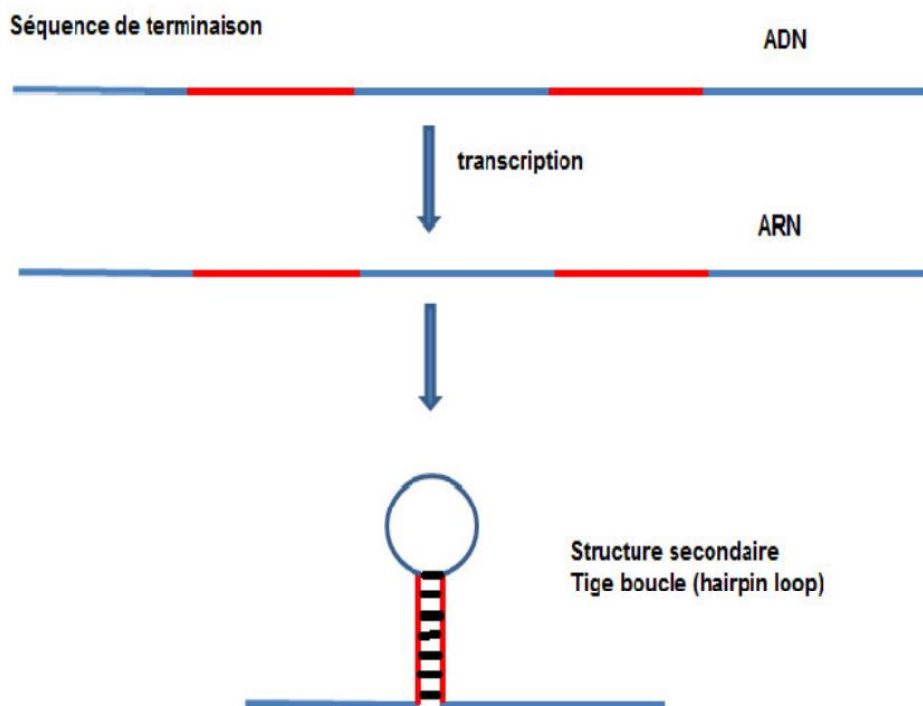


figure 8: Formation du signal de terminaison de la transcription chez les bactéries (Aouf, 2016)

Ces structures secondaires suivies d'une série d'uridines sont des terminateurs efficaces de la transcription (figure 8). Des séquences riches en GC suivie par autres riche en AT sont aussi des sites spécifiques de terminaison. Ces régions qui ne nécessitent pas de facteurs additionnels sont nommées "**terminateurs intrinsèques**".

□ Certains signaux de terminaison ont une structure en épingle à cheveu insuffisante pour entraîner une dissociation immédiate de l'ARN. Ils ont besoin de la liaison d'une **protéine p (Rho)**. La protéine p se fixe, dans une réaction impliquant l'hydrolyse d'ATP, sur l'ARN, alors que l'ARN polymérase est dans un site de pause ; l'ARN entoure la protéine p sur 80 nucléotides ; elle interagit aussi avec les protéines associées et détache l'ARN de sa liaison à l'enzyme (l'ATP sert de donneur d'énergie dans cette réaction) (figure 9)

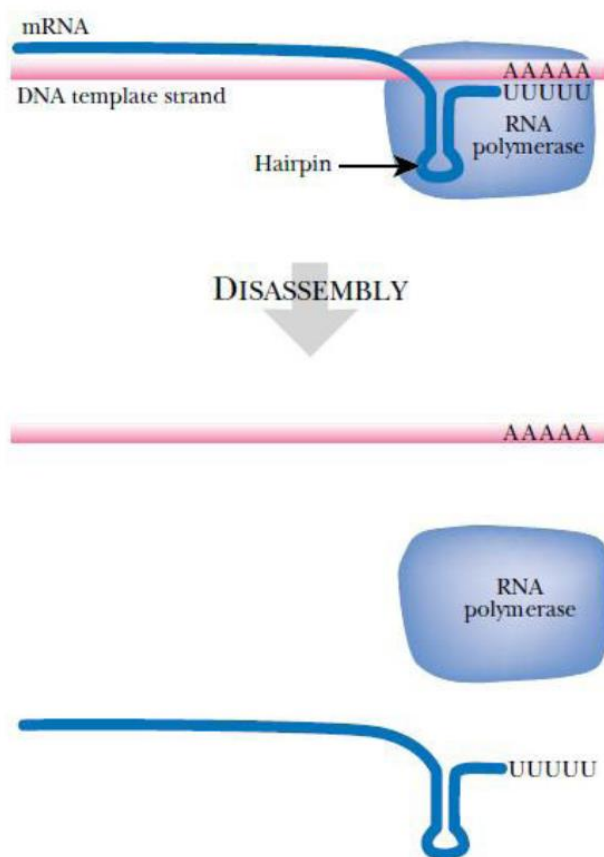


Figure 9: Terminaison de la transcription par des terminateurs intrinsèques (Clark, 2005)

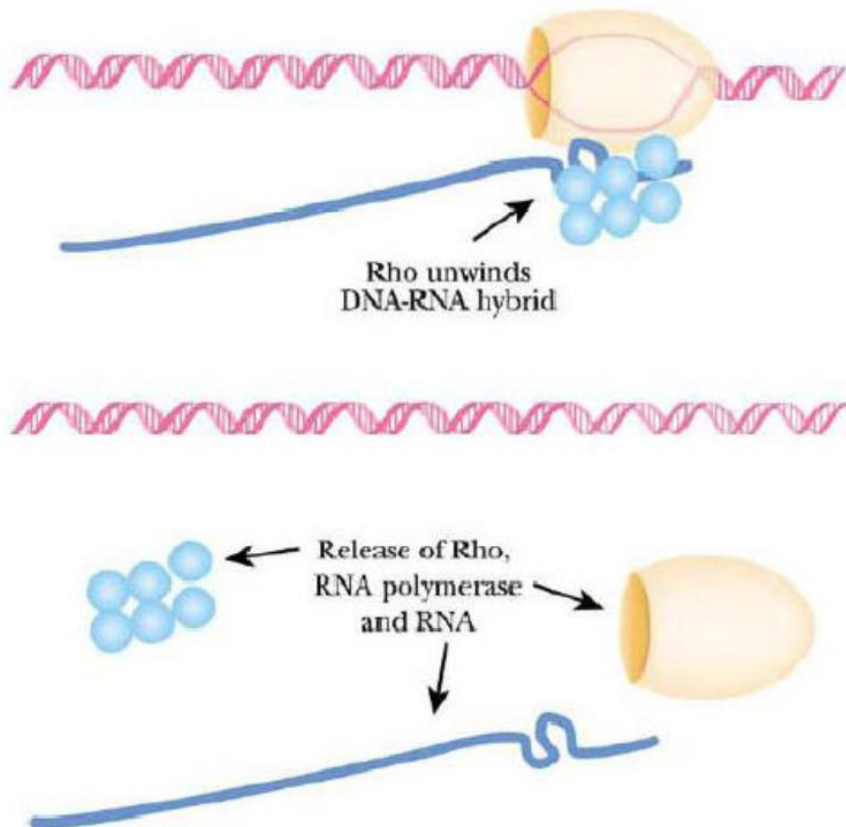


Figure 10: Terminaison de la transcription chez les bactéries par le facteur *Rho* (Clark, 2005).

III) Transcription de l'ADN eucaryote

Chez les eucaryotes le mécanisme de base de la transcription est identique à ce qui a été décrit pour les procaryotes ; cependant **la structure du promoteur** est différente et les transcrits primaire obtenus sont toujours **monocistronique**.

Trois points essentiels distinguent la transcription des eucaryotes et des procaryotes :

- chez les eucaryotes, l'ADN est contenu dans le noyau et la transcription s'effectue dans celui-ci alors que la traduction en protéines s'effectue dans le cytoplasme. Chez les procaryotes, ces deux phénomènes sont associés et la traduction commence avant même que la transcription soit fini. De plus, l'ADN des Eucaryotes est présent sous forme d'une structure appelée la chromatine. La transcription de l'ADN en ARN implique une structure décondensée de la chromatine, c'est-à-dire une chromatine où il existe peu d'histone H1, où la structure est en collier de perles avec des zones sans nucléosomes.

- Chez les Eucaryotes, il existe trois ARN polymérase distinctes (alors qu'une seule fait «tout le travail» chez les bactéries)

- ARN polymérase I : ARNr (sauf l'ARN 5S)
- ARN polymérase III : ARNt, ARN 5S et petits ARN nucléaires
- ARN polymérase II : ARNm.

Ces polymérase ne fonctionnent correctement (en particulier pour la fixation sur des sites promoteurs) qu'en présence de très nombreuses protéines supplémentaires appelées **facteurs de transcription**.

- chez les eucaryotes, la transcription d'un gène en ARN m, r, t ne produit pas immédiatement une molécule fonctionnelle. Dans le noyau c'est un précurseur (ou transcrit primaire) qui est produit dans la grande majorité des cas, il subira plusieurs modifications avant son entre dans le cytoplasme.

III.1 Transcription par l'ARN polymérase I :

La synthèse de l'ARN ribosomal s'effectue dans un compartiment spécialisé du noyau, le nucléole. Cette synthèse représente plus de la moitié de l'ARN total cellulaire avec près de 80 % de tous les ARN.

Elle s'effectue sous forme d'un précurseur de grande taille (en moyenne 45S) qui sera converti en ARNr 18S, 5,8S et 28S. Ces ARN, avec l'ARN 5S transcrit sous contrôle d'une autre ARN polymérase (ARN polymérase III), vont former les constituants non protéiques du ribosome.

Cette synthèse est sous contrôle d'une ARN polymérase spécialisée, l'ARN polymérase I.

III.1.A Fonctionnement de la transcription

Le site promoteur comporte deux séquences situées vers -140 (site distal) et vers -35 (site proximal). Le premier site est le plus important pour la transcription, mais une transcription très active ne peut avoir lieu sans le deuxième site (figure 11)



Figure 11 : Structure du site d'initiation de transcription pour l'ARN pol I. (Beaumont , 2007)

La séquence de ces deux sites est spécifique de l'espèce, ce qui correspond à une fixation de facteurs de transcription qui ne reconnaissent que le site promoteur de l'espèce dont ils proviennent.

Deux protéines se lient au promoteur :

- SL1 ou TFI D ou TFI B (fixation sur les deux sites)
- UBF (fixation entre les deux sites).

Seule la protéine SL1 a la spécificité d'espèce.

Remarque :

Un autre facteur protéique TFI C semble nécessaire à ce démarrage de la transcription. La structure ressemble à la sous-unité σ de la polymérase *d'E. coli*.

Enfin l'ARN polymérase I est nécessaire. Elle est composée de 65 sous-unités chez les vertébrés. La plus grande ressemble à β *d'E. coli*.

La terminaison est assurée par une structure composée de deux parties :

- une première où se fixe une protéine qui a été identifiée aussi bien chez un eucaryote inférieur, la Levure *S. cerevisiae* (Reb-1p), que chez les Mammifères (TTF-1)
- une seconde qui est indispensable aussi à la libération du transcrit d'ARN et qui est une séquence riche en paire A-T.

La protéine va bloquer la polymérase en situation de « pause » et cette dernière va reculer ce qui va créer un hybride ARN-ADN instable et dissocier l'ARN de sa liaison avec l'ADN et avec l'ARN polymérase.

III.1.B Maturation du précurseur 45 S et assemblage :

Quand la copie de l'ADN en pré-ARNr est faite, le pré-ARNr est incorporé dans un grand complexe ribonucléoprotéique qui clive ce pré-ARNr. Ce complexe contient la ribonucléoprotéine U3 : découpage d'abord dans la partie 5' puis du côté 3', puis entre les

deux. Ce n'est pas une élimination d'introns sous forme d'épissage, mais une action de type nucléasique. Après transcription, le pré-ARNr ou ARN 45 S (environ 14 000 bases) subit aussi des modifications post-transcriptionnelles : méthylation, réduction, formation de pseudo-uridine, etc. comme chez les Procaryotes.

L'assemblage des ARNr obtenus après maturation avec les protéines constitutives des deux sous-unités ribosomales 40S et 60S (85 protéines différentes) et avec l'ARN 5S (qui s'associe à la grande sous-unité 60S) se fait dans le nucléole. Ce n'est que lorsque l'assemblage est terminé que les sous-unités vont quitter le nucléole pour aller vers le cytoplasme.

III.2 Transcription par l'ARN polymérase III :

L'ARN polymérase III copie :

- l'ARN 5S constitutif des ribosomes
- les ARN de transfert
- certains ARN nucléaires appelés petits ARN nucléaires ou ARNsn.

III.2.A Le site promoteur pour l'ARN 5S et la fixation de l'ARN polymérase III :

On sait qu'il est situé à l'intérieur du gène et qu'il se lie à un facteur de transcription TFIII A. La liaison de ce facteur à l'ADN est une structure dite doigt de zinc Cys 2-His 2 (deux Cys et deux His en interaction avec un ion zinc II) qui se lie au grand sillon de l'ADN.

La partie C-terminale de TFIII A se lie à TFIII B et C. L'ensemble va ensuite interagir avec l'ARN polymérase III. Le complexe recouvre alors la totalité du gène 5S.

C'est en définitive l'interaction ARN pol III-TFIII B qui permet à celle-ci de commencer la transcription. Il est à noter que cette transcription se fait sans dissociation du complexe TFIII A-ADN.

III.2.B La transcription des ARNt :

Le site promoteur pour l'ARNt est constitué de deux séquences (+ 10 à + 20) et (+ 50 à + 60) qui vont fixer le facteur de transcription TFIII C.

L'interaction se fait en plusieurs étapes :

- TFIII C se lie à l'ADN
- après fixation, TFIII B se fixe à TFIII C
- le complexe stable formé permet alors la liaison de la polymérase III.

III.2.C La transcription et la maturation du transcrit primaire :

- 5S, ARNsn : pas de maturation.
- L'ARN de transfert subit un processus d'épissage : élimination d'une séquence « inutile » (intron) et « recollage » des séquences « utiles », avant de passer dans le cytoplasme. Ce

processus d'élimination de l'intron est assez différent de celui utilisé pour l'ARNm. Ce n'est qu'après l'élimination de l'intron que l'ARNt passe dans le cytoplasme.

III.3 Transcription par l'ARN polymérase II :

III.3.1 Transcription d'un ARN messager :

III.3.1.1. L'initiation de la transcription :

Tout l'ADN n'est pas transcrit, seules les régions correspondant à des gènes le sont. Et encore, cette expression peut être régulée selon le stade de développement, le type cellulaire, l'environnement,... Dès lors, un acteur doit intervenir pour déterminer à quel endroit une région d'ADN doit commencer à être transcrite : c'est le rôle du promoteur.

III.3.1.1.a. Site promoteur :

Il faut bien distinguer, étant donné le niveau d'expression variable des gènes, deux zones distinctes :

- le site promoteur proximal
- les sites de régulation.

Le promoteur correspond à une région non transcrite de l'ADN, généralement juste en amont du début de la région transcrite, dont la séquence permet le recrutement de l'ARNpol II. Certaines séquences du promoteur (surnommées "boîtes") ont une importance particulière dans ce processus, essentiellement parce que ces séquences sont reconnues spécifiquement par différentes protéines appartenant au complexe d'initiation:

□ la "**boîte TATA**" riche en thymine et adénine, la plus importante, est située vers **-25** à **-30** nucléotide du site de démarrage de la transcription (noté +1)

□ des éléments proximaux :

o la "**boîte CAAT**" (facultative), contenant de la cytosine, est située vers **-120** à **-80** nucléotides du site de démarrage de la transcription.

o la "**boîte GC**" (facultative également), riche en guanine et cytosine, peut être présente entre la boîte CAAT et la boîte TATA.

Signalons que si ces séquences sont souvent bien conservées, une variabilité non négligeable peut également être observée. Ainsi, il existe des promoteurs sans "boîte TATA".

III.3.1.1.b. Le complexe d'initiation :

Contrairement à ce qui se passe chez les procaryotes, l'ARN pol II des eucaryotes ne reconnaît pas seul le promoteur proximal. Elle effectue ce travail en compagnie de nombreux co-facteurs protéiques qui se recrutent les uns les autres et qui forment avec elle un complexe d'initiation. Ces facteurs sont notés TFIIA, TFIIB,... pour **Transcription Factor for RNA**

polymerase II. Ils correspondent aux facteurs généraux de la transcription car ils s'assemblent sur tous les promoteurs utilisés par l'ARN pol II. La séquence d'assemblage du complexe d'initiation se fait dans un ordre précis : TFIID puis TFIIA, B, ARN pol II se fixe ensuite suivie par TFIIF, E, H, pour produire finalement un complexe fonctionnel capable d'initier la transcription (figure 12).

Le recrutement de la polymérase II est très compliqué. Jusqu'à récemment, on considérait l'accrochage séquentiel suivant pour les gènes : TFII D, puis TFII A qui augmente la liaison de TFII D à TATA, puis TFII B qui reconnaît TBP (TATA box Binding Protein) et permet le recrutement de l'ARN pol II avec II F, II E, faiblement associé, II H et II B.

Il est capital de comprendre que la transcription effectuée en présence de ces différents facteurs est très faible. Le rôle de différentes protéines régulatrices qui s'associent à des séquences situées en amont ou en aval, voire à l'intérieur du gène, sera essentiellement d'augmenter la stabilité du complexe d'initiation de la transcription : liaison à TBP mais surtout aux TAF (TBP Associated Factors), liaison aux SRB. La fixation de différents facteurs de régulation va augmenter de 100 à 1 000 le taux de transcription

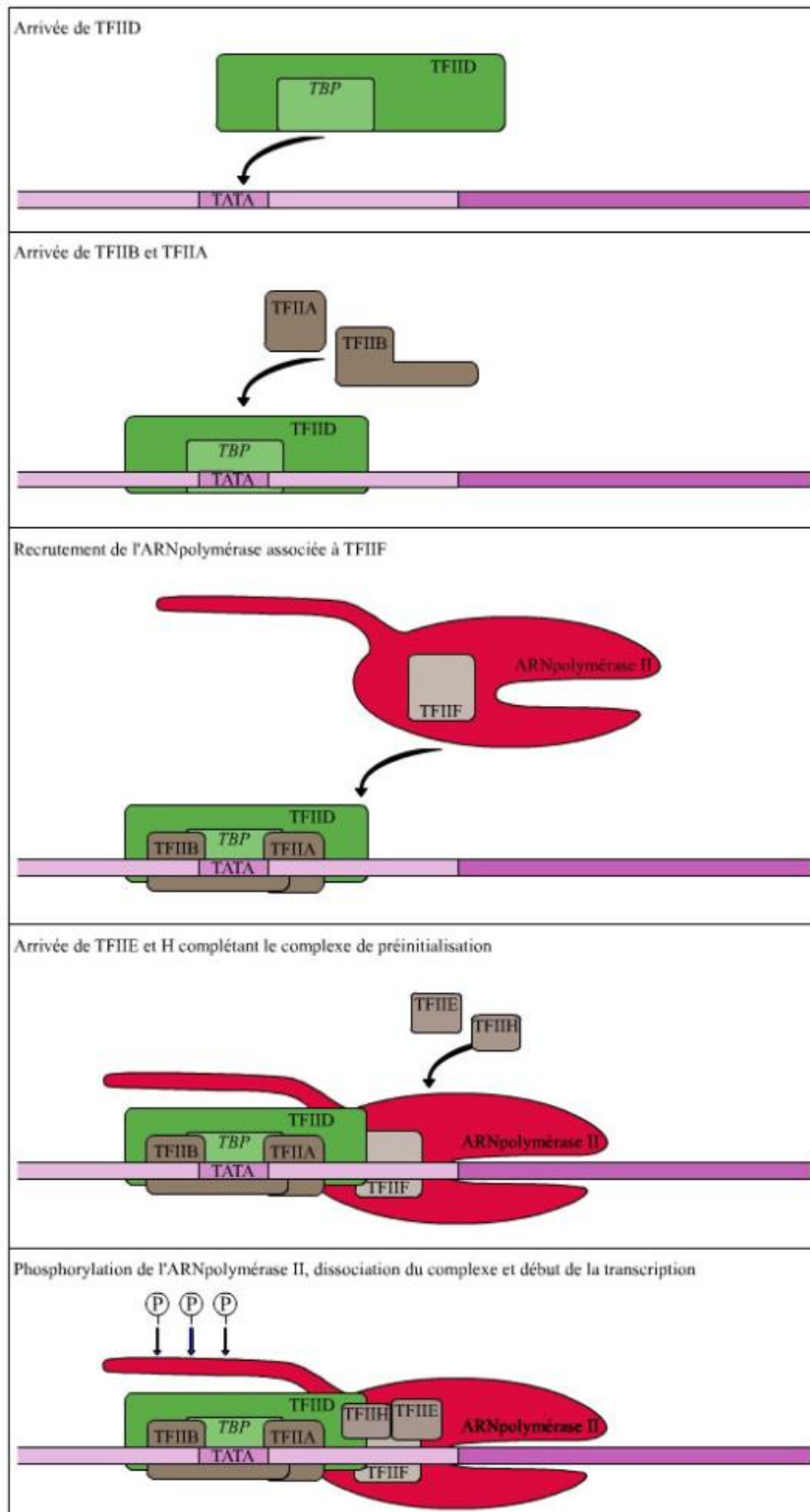


Figure 12: Mise en place du complexe d'initiation

La liaison du complexe de transcription au promoteur proximal provoque l'ouverture et le déroulement des deux brins de son ADN, tout en indiquant le brin qui va être transcrit.

III.3.1.1.c. Intervention de facteurs spécifiques de la transcription :

Le complexe d'initiation composé de l'ARN pol II et des différents TFII est suffisant pour obtenir une activité transcriptionnelle *in vitro* mais à très faible taux. L'augmentation de cette activité basale (ou sa répression) est sous la dépendance de facteurs spécifiques qui vont interagir avec le complexe d'initiation. Ces protéines activatrices ou inhibitrices se lient à des promoteurs distaux, séquences spécifiques de l'ADN, appelées **enhancer** lorsqu'il recrute des cofacteurs **activateurs**, ou **silencer** lorsqu'ils recrutent des cofacteurs **inhibiteurs**. Ces promoteurs distaux peuvent être situés à des milliers de nucléotides du promoteur proximal et agissent sur le promoteur proximal par le jeu de courbures de l'ADN (voir figure 13). Mais d'autres participants existent encore. On trouve ainsi des promoteurs proximaux alternatifs et des promoteurs distaux, situés à des milliers de nucléotides du promoteur proximal. Malgré la distance qui sépare les promoteurs proximaux des promoteurs distaux, ces derniers agissent sur le promoteur proximal comme activateurs (enhancers; voir figure 13) ou répresseurs (silencers) par le jeu de courbures de l'ADN, des facteurs de transcription et du médiateur qui maintient liés tous ces acteurs

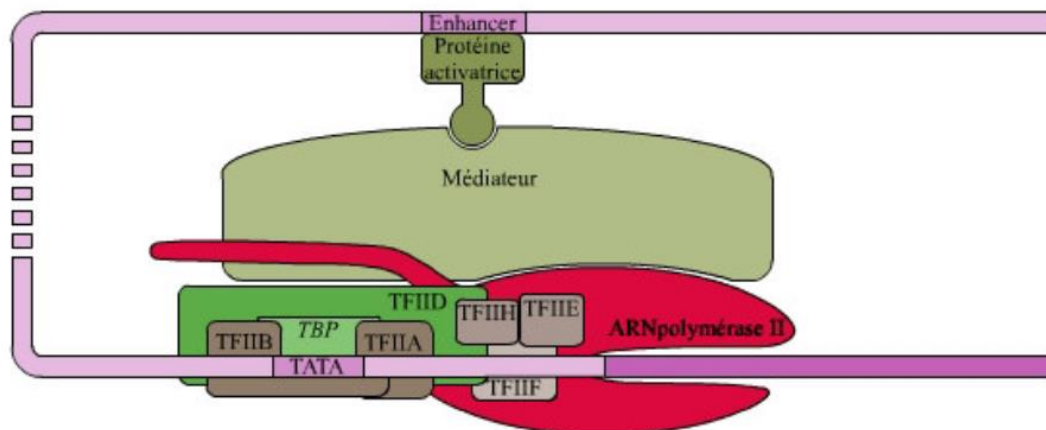


Figure 13 : La régulation spécifique de la transcription

Les enhancers (activateurs) sont des séquences situées parfois à plusieurs milliers de nucléotides d'un promoteur qui activent ce dernier par un jeu d'interactions protéiques qui stabilisent le complexe d'initiation. Ces interactions sont rendues possibles par la courbure de l'ADN qui rapproche des éléments situés à de grandes distances les uns de

autres. Cette activation est spécifique du gène et utilise une multitude de facteurs de transcriptions spécifiques (les protéines activatrices) qui agissent généralement sous forme dimérique). Une répression spécifique peut agir selon le même type de modalité.

III.3.1.2. L'élongation :

La TBP est la première protéine qui reconnaît une séquence spécifique de l'ADN initiatrice de la transcription (la boîte TATA). Le facteur TFIIB semble impliqué dans la sélection précise du site d'initiation (nucléotide à partir duquel se déroule la transcription). Le facteur TFIIF comporte plusieurs activités enzymatiques dont une activité **hélicase** permettant l'ouverture de la double hélice d'ADN au niveau du promoteur, et une activité **kinase** responsable de la phosphorylation de la queue C-terminale de l'ARN polymérase II (domaine CTD, pour C-Terminal Domain) est un des événements « clef » du passage à l'élongation. Cette phosphorylation entraîne une modification de la structure tridimensionnelle de l'ARN polymérase entraînant la dissociation du complexe d'initiation et le début de la transcription.

La grande sous unité de pol II possède sur l'extrémité C-terminale une série de séquence répétée de 7 acides aminés (heptapeptide: Tyr-Ser-Pro-Thr-Ser-Pro-Ser) appelée domaine carboxy terminal (CTD) ou la queue. Le nombre de répétitions dépend des espèces: 27 fois chez les levures, 45 fois chez la mouche *Drosophila*, et 52 fois chez l'homme. Chaque motif répété contient des sites de phosphorylation par des kinases spécifiques, notamment une qui est une sous unité de TFIIF. La régulation du niveau de phosphorylation du CTD de Pol II contrôle aussi les étapes ultérieures (élongation, maturation).

TFII = Transcription Factor for RNA polymerase II (Facteurs généraux de la transcription pour l'ARN polymérase II).

TBP = TATA box-Binding Protein (Protéine de liaison à la boîte TATA).

L'ARN pol II est équipé de facteurs protéiques d'élongation qui facilitent sa progression au travers d'une chromatine dont ils relâchent la structure. Un ARN pré-messager complémentaire du brin matrice de l'ADN (brin antisens), donc identique au brin codant de l'ADN (brin sens) aux riboses et uraciles près, commence à être synthétisé selon la direction 5'-3' (voir figure 14).

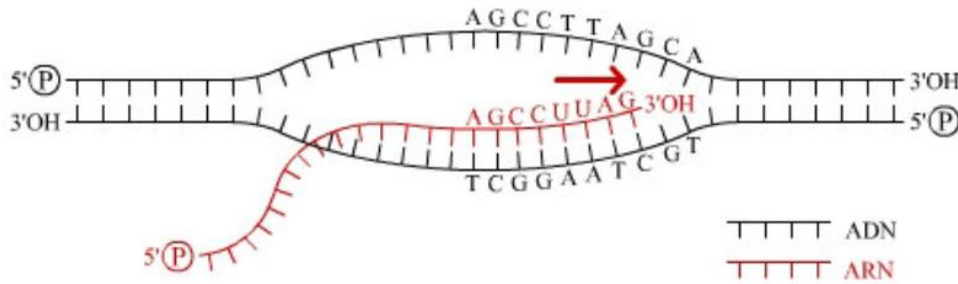


Figure 14 : La phase d'élongation de la transcription

L'ARN polymérase lit le brin antisens qui est complémentaire du brin codant (ou brin sens) depuis son extrémité 3' vers son extrémité 5'. Le brin d'ARN néosynthétisé est donc identique au brin codant d'ADN, aux uraciles et riboses près.

III.3.1.3. La terminaison :

L'ARN pol II est également équipée de facteurs protéiques de terminaison, et reconnaît ainsi un ou plusieurs signaux de terminaison portés par le brin progressivement parcouru et qui annoncent la fin de la transcription sur le brin d'ADN matrice (TTATTT). Elle arrête bientôt son travail de transcription et libère l'ARN prém qu'elle vient d'assembler.

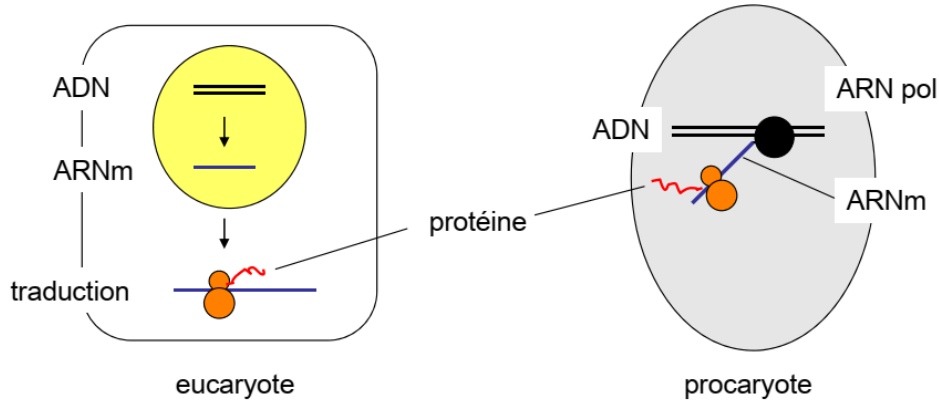
III.4 Inhibiteurs de transcription :

Il existe certains antibiotiques qui se comportent comme des inhibiteurs de la transcription, en agissant à des niveaux différents :

- *La rifampicine*, qui inhibe l'étape d'initiation de la transcription, en se plaçant dans le site actif de l'ARN polymérase.
- *L'actinomycine D*, s'intercale entre les plans des bases de l'ADN, et l'empêche alors de servir de matrice à la fabrication de l'ARN.
- La toxine de l'amanite phalloïde (champignon vénéneux), *l'alpha amanitine* inhibe l'ARN polymérase II des Eucaryotes en s'y fixant avec une très grande affinité.

Transcription des procaryotes vs eucaryotes

Chez les procaryotes, la transcription est couplée à la traduction alors que chez les eucaryotes les ARNm seront exportés hors du noyau avant d’être traduits.



Chez les procaryotes, plusieurs gènes peuvent être transcrits à partir d’un promoteur unique

