

# La transcription

---

## I. Introduction

Un **transcriptome** eucaryote est beaucoup plus complexe qu'un transcriptome bactérien et il varie selon les fonctions.

Lors de la transcription il y a tout d'abord synthèse d'un **transcrit primaire** qui peut soit donner un ARN codant soit un ARN non codant.

L'**ARN codant** donne un messenger fonctionnel chez les bactéries et un prémessenger qui subira des modifications post-transcriptionnelles chez les eucaryotes pour devenir mature.

L'**ARN non codant** donne des ARNtm et des ARN 6s de petite taille chez les bactéries. Chez les eucaryotes il donne de petits ARN nucléaires (sn), nucléolaires (sno), cytoplasmiques (sc) et des ARNpremi qui donneront des micro-ARN ; mais aussi des préARNt et préARNr.

L'**ARN polymérase ADN dépendante** utilise un brin d'ADN comme matrice, c'est le brin transcrit. Le brin non transcrit est donné par défaut car il possède la même séquence que le messenger avec des « T » à la place des « U ».

Il n'y a pas d'amorce pour la synthèse de l'ARN de 5' vers 3', la sélection des NTP (nucléotides triphosphates) se fait par complémentarité des bases. La transcription est un processus asymétrique, c'est-à-dire que pour un gène ou un groupe de gènes donné, seul l'un des deux brins de l'ADN est transcrit.

Dans la **cellule bactérienne**, une seule ARN polymérase permet la synthèse des ARNr, ARNt et ARNm.

Dans le noyau des **cellules eucaryotes**, on retrouve trois ARN polymérases ADN dépendantes :

- ARN polymérase I qui transcrit les gènes des ARNr (28s, 18s et 5,8s) ;
- ARN polymérase III catalyse la synthèse des ARNt, de l'ARNr 5s et est aussi responsable de la transcription des petits ARN nucléaires ;
- ARN polymérase II responsable de la transcription des gènes codant les protéines traduites par les ARNm.

Ces ARN doivent reconnaître des signaux constitués d'éléments/motifs de séquence au niveau du promoteur (= région d'ADN qui possède des éléments de séquence nécessaires à la mise en place du complexe d'initiation à la transcription.)

**Chez les bactéries**, le **promoteur** se situe en 5' du site d'initiation de la transcription c'est-à-dire en amont de la région +1 qui est le premier nucléotide incorporé dans l'ARN. L'ARN polymérase se fixe directement sur l'ADN.

**Chez les eucaryotes**, le **promoteur** peut se trouver en amont ou en aval du site d'initiation, des facteurs de transcription généraux se fixent sur le promoteur et sont spécifiques à une ARN polymérase donnée. La plateforme est donc constituée par les facteurs de transcription fixés sur l'ADN et reconnu par l'ARN polymérase qui induira le début de la transcription.

Mise en évidence de la fixation d'un facteur de la transcription sur sa cible : expérience de retard sur gel

Le fragment d'ADN étudié, rendu radioactif et constituant la sonde, est incubé en présence des protéines susceptibles de s'y associer, dans les conditions où seules les protéines recherchées se complexent spécifiquement à la sonde. L'ADN étudié est en excès par rapport à la quantité de protéines. Si la protéine reconnaît sa cible, elle se fixe et on obtient alors un complexe ADN-protéine. Après formation des complexes spécifiques, le mélange est soumis à électrophorèse sur gel de polyacrylamide en conditions non dénaturantes. On effectue aussi un dépôt témoin constitué uniquement d'ADN marqué. Le complexe ADN-protéine migrera plus lentement que le fragment d'ADN libre et sa position pourra être visualisée sur le gel par autoradiographie. Cette méthode permet donc de comparer les concentrations relatives de la protéine spécifique dans différents tissus.

Empreinte à la DNase I :

Le site de fixation de l'ARN polymérase peut être étudié par la méthode dite de *foot-printing*, qui consiste à soumettre le fragment d'ADN protégé par l'enzyme (protéine) à une digestion endonucléasique ménagée qui va pouvoir couper toutes les liaisons phosphodiester accessibles, mais en provoquant statistiquement une coupure par molécule. Comme dans la méthode des gels de séquence, les fragments engendrés sont séparés en fonction de leur taille par électrophorèse sur gel. Après autoradiographie, on obtient pour chaque liaison rompue, une bande sur le gel correspondant au fragment allant de l'extrémité marquée jusqu'au site de coupure. La comparaison des bandes obtenues sur le gel pour le fragment d'ADN non protégé et le fragment protégé fait apparaître que des bandes manquent sur le gel correspondant au fragment protégé ; il s'agit de la région où la polymérase fixée sur l'ADN a empêché l'action endonucléasique. En combinant cette méthode avec le séquençage, on pourra déterminer la position du site de fixation et sa séquence.

Expérience d'interférence :

Cette expérience consiste à modifier chimiquement l'ADN, par exemple la DMS méthyle spécifiquement le N7 de la guanine. Si le complexe ADN-protéine ne se forme pas lorsqu'on ajoute la protéine cela signifie que la base modifiée était importante dans la formation du complexe.

Pour identifier les bases modifiées, on réalise des traitements chimiques qui suppriment d'abord la base modifiée. Ensuite on supprime le ribose qui ne possède plus sa base et on dépose l'ADN sur le gel dénaturant.

Isolation des facteurs protéiques de la transcription :

- Par *chromatographie d'affinité* : un fragment d'ADN ou un oligonucléotide synthétique, contenant le site de liaison de la protéine, est immobilisé sur une colonne d'affinité. L'extrait contenant la protéine est ensuite passé sur la colonne à basse force ionique ce qui permet la liaison des protéines à leurs sites de liaison. La protéine spécifique de la séquence immobilisée est retenue sur la colonne alors que toutes les autres protéines passent à travers. Une fois que ces protéines indésirables ont été complètement éliminées par lavage, la colonne est éluée à haute force ionique, ce qui détruit le complexe ADN-protéine. On recueille la protéine de liaison purifiée.
- Par *criblage d'une banque d'expression* : ADNc est de l'ADN synthétisé à partir des messagers qui servent de matrice. Le clonage est l'insertion de ces ADN dans un vecteur qui constitue la banque d'ADNc. On transforme des bactéries en boîte de culture, l'ADNc code pour une protéine synthétisée par ces bactéries. On transfère les bactéries sur un filtre (membrane) et on ajoute de l'ADN marqué renfermant la cible du facteur de transcription qu'on veut cloner = formation d'un complexe ADN-protéine si la protéine reconnaît le site de fixation. On lave dans des conditions douces pour ne garder que le complexe d'intérêt et pour enlever toutes les liaisons aspécifiques. On fait une autoradiographie et on observe des spots noirs qui correspondent aux complexes ADN<sub>radioactif</sub>-protéine. On recommence l'expérience plusieurs fois pour obtenir sur l'autoradiographie uniquement la culture qui possède la protéine d'intérêt.

## II. Transcription chez les bactéries

### A. L'ARN polymérase

Core enzyme +  $\sigma^{70}$  = holoenzyme qui est la seule qui est capable d'initier la transcription.

L'holoenzyme est donc constituée de 6 sous-unités :  $\alpha_2\beta\beta'\sigma^{70}$ . L'élément  $\omega$  compte pour moins d'un exemplaire par molécule d'holoenzyme et il n'est pas certain qu'il fasse parti de l'holoenzyme fonctionnel.

$\alpha_2$  : assemblage du core enzyme qui interagit avec des protéines régulatrices et parfois avec l'ADN ;

$\beta\beta'$  : impliqués dans la catalyse ;

$\sigma^{70}$  : spécificité de reconnaissance du promoteur.

### B. Fixation de l'ARN polymérase sur l'ADN

Le core enzyme est basique et  $\sigma$  permet la fixation spécifique de l'holoenzyme en présence du promoteur. Une fois que l'ARN polymérase reconnaît un promoteur il y a formation d'un complexe binaire (holoenzyme/promoteur) fermé. Dans ce complexe, l'ADN est bicaténaire. Le complexe s'ouvre rapidement et forme un complexe ternaire (ADN/enzyme/chaîne naissante d'ARN) après synthèse de 9-10 nucléotides d'ARN sans que l'ARN polymérase

bouge. Dans ce cas 18 pbs sont dénaturées pour former la bulle de transcription. Il peut y avoir arrêt de la transcription si l'amorce ARN n'est pas assez longue, sinon l'ARN polymérase entre en phase d'élongation lorsque le facteur  $\sigma$  se détache du complexe. C'est l'enzyme minimum (core) qui poursuit la polymérisation. La dissociation du facteur  $\sigma$  diminue la stabilité du complexe d'initiation et permet la progression de l'enzyme minimum qui se déplace le long du brin matrice de l'ADN, et la région d'ouverture de la double hélice accompagne ce déplacement.

Les topoisomérases sont aussi impliquées dans la transcription puisqu'il y a aussi formation de super-tours.

#### a. Promoteur

Chez les bactéries, il est constitué des régions consensus -35 (TTGACA) et -10 (TATAAT). Les expériences à la DNase I ont montré que le site de fixation apparaît localisé entre -50 et +20. Les expériences d'interférence ont montré des points de contact en amont du site d'initiation, ce sont des motifs communs (consensus) comportant six paires de bases (hexamères), appelés respectivement séquence -10 (TATA) et -35. Au niveau de ces consensus il y a des éléments invariants mais on constate que si on modifie par mutations les séquences consensus il y a altération de la transcription.

L'initiateur se trouve dans une région très conservée qui possède le site +1.

La région entre les 2 boîtes -10 et -35 est importante car si on modifie cette distance en réalisant des délétions il n'y a plus de transcription. Cette distance permet de positionner correctement les boîtes sur la même phase de l'hélice.

Certains promoteurs sont dits forts/faibles : les *promoteurs forts* sont les séquences -10 et -35 qui sont proches de séquences consensus et qui assurent l'initiation très efficace de la transcription ; les *promoteurs faibles* ont une efficacité moins importante (ex : opéron lactose).

#### b. Facteur $\sigma$

Les facteurs ou éléments  $\sigma$  influencent les propriétés de liaison du noyau de la polymérase.  $\sigma$  accroît l'affinité du noyau de la polymérase envers les séquences propres aux promoteurs.

Les bactéries ont plusieurs  $\sigma$  :  $\sigma_{70}$  qui est le facteur général,  $\sigma_{32}$  qui est un facteur spécial induit par l'augmentation de la température et le  $\sigma_{34}$  induit lors de la diminution d'azote.

Mutations compensatoires : elles permettent au facteur  $\sigma$  muté de reconnaître la séquence mutée. Le facteur  $\sigma$  seul ne peut pas se fixer à l'ADN car son domaine de fixation est masqué par la partie N-term. Il y a un changement de conformation lorsqu'il est fixé au core enzyme, ce qui permet de démasquer le DBD (DNA Binding Domain).

Il y a des mutations qui ont montré que c'est au niveau de la boîte -10 que se fait l'ouverture de l'ADN et que la boîte -35 est importante dans la spécificité de la reconnaissance.

Le facteur  $\sigma$  est responsable de la transcription de gènes spécifiques (ex : infection de *B.subtilis* par le phage SPO1) : les gènes précoces utilisent l'ARN polymérase bactérienne, parmi ces gènes précoces il y en a qui codent pour le facteur phagique  $\sigma_{p1}$ .  $\sigma_{p1}$  remplace le facteur  $\sigma$  bactérien au niveau de l'holoenzyme qui forme alors une ARN polymérase bactérienne au profit du phage. La transcription des gènes intermédiaires code pour le

facteur  $\sigma_{p2}$  qui prend à son tour la place de  $\sigma_{p1}$  au niveau de l'holoenzyme. L'ARN polymérase ainsi formé catalyse la transcription de gènes tardifs.

### C. Elongation

Il y a formation d'un hybride. Le facteur  $\sigma$  ne participe pas à cette étape. L'ARN polymérase transcrit le gène sur lequel il y a un signal de terminaison.

### D. Terminaison

#### a. Mécanisme intrinsèque

Au cours de l'allongement de la chaîne, la vitesse de synthèse d'ARN est influencée par les séquences transcrites. Aussi quand le complexe transcripteur siège sur une séquence riche en GC, la matrice d'ADN s'entrouvre plus difficilement pour former la bulle de transcription. Ces séquences, appelées sites de pause, sont capables de ralentir la transcription d'un facteur 10 à 100.

L'ampleur de la pause s'accroît dans les régions d'ADN présentant une symétrie binaire, régions appelées palindromes. Pendant la transcription d'une séquence palindromique, l'ARN néoformé peut se replier en épingle à cheveux, ce qui extirpe le brin d'ARN de la bulle de transcription ; en dissociant ainsi prématurément une partie de l'ARN naissant, la formation de cette épingle à cheveux déstabilise l'hybride ARN-ADN attaché au complexe transcripteur. Il y a une région riche en U (8U) située en 3' de cette structure secondaire qui est facile à rompre ce qui entraîne la déstabilisation de l'hybride et le relargage de l'ARN polymérase.

#### b. Mécanisme Rho dépendant

Rho est une protéine hexamérique présentant un domaine de fixation à l'ARN et reconnaît une région riche en C. Elle présente une fonction hélicase qui fixe 6 ATP.

Le facteur rho se fixe en 3' de la phase de lecture ouverte (entre AUG et « stop »), casse l'hybride ARN-ADN et entraîne un changement de conformation de l'enzyme et donc un relargage de l'ARN polymérase.

#### c. Opéron lactose

Un **opéron** est un ensemble de gènes sous le contrôle d'un promoteur unique. Ces gènes codent pour une protéine impliquée dans une voie métabolique identique.

Le **terminateur** se situe en aval des gènes. Au niveau du promoteur il y a un **opérateur** qui est une séquence d'ADN reconnue par le facteur de régulation (répresseur ou activateur). L'opérateur peut chevaucher ou intégrer le promoteur.

Le gène régulateur possède son propre promoteur/terminateur et il code pour une protéine qui agit en trans.

- Régulation négative : le gène régulateur code pour un répresseur (protéine) qui se fixe spécifiquement au niveau de l'opérateur pour inhiber la transcription de celui-ci. Mais il y a toujours une transcription de base.

- Induction : le lactose pénètre dans la bactérie grâce à la perméase qui a été transcrite, il est ensuite transformé en allolactose par la  $\beta$ -galactosidase. L'allactose est un inducteur qui se fixe sur le répresseur et induit son changement de conformation. Le répresseur ne peut plus se fixer sur l'opérateur.

#### d. Opéron tryptophane

Un **atténuateur** est un terminateur de la transcription interne localisé au début d'une unité de transcription.

L'opéron Trp renferme des gènes qui sont impliqués dans la biosynthèse du Trp. En 3' de ces gènes il y a un terminateur et en 5' il y a un ARN leader de 140 nucléotides qui présente différents motifs complémentaires (séquences complémentaires et antiparallèles) : le motif 1 s'associe avec le motif 2 et le 3 avec le 4. Mais si le motif 2 ne se fixe pas au motif 1 il peut se fixer au motif 3. A la fin du motif 4 il y a une région de 8U qui peut former une structure en « épingle à cheveux » qui peut entraîner l'arrêt de la transcription.

L'ARN leader peut adopter 2 types de structures secondaires. Il possède une petite phase de lecture ouverte avec 2 Trp pour 14 codons qui est importante pour la régulation de l'opéron.

Chez les bactéries, la transcription et la traduction sont liées très étroitement. Il y a 2 structures alternatives :

- Cas où il y a du Trp : la traduction démarre en présence d'un codon AUG, le ribosome masque environ 60 nucléotides et fonctionne jusqu'au codon « stop ». Une partie de la région 2 s'apparie avec 3 et la partie libre de la région 3 s'apparie avec la région 4. L'appariement de 3 avec 4 forme l'atténuateur qui entraîne un arrêt de la transcription de l'opéron Trp.
- Cas où il n'y a pas de Trp : il y a peu de Trp dans la bactérie donc peu d'ARNt sont chargés avec du Trp. Le ribosome traduit l'ARN leader puis attend les W-tRNA au niveau des codons Trp de l'ARN leader. Pendant ce temps l'ARN polymérase synthétise la région 2 qui s'apparie avec la région 3 car la région 1 est masquée par le ribosome. La région 4 synthétisée reste alors simple brin et il n'y a pas formation de l'atténuateur.

Le Trp est un **corépresseur** qui interagit avec un répresseur. Le complexe se fixe au niveau de l'opérateur et induit l'inhibition de la transcription.

En fonction de la teneur en Trp-ARNt<sup>Trp</sup>, les ribosomes masquent soit partiellement soit totalement la région 1 : forte [Trp-ARNt<sup>Trp</sup>] induit la formation d'un atténuateur et l'arrêt de la transcription alors qu'une faible [Trp-ARNt<sup>Trp</sup>] induit une transcription complète de l'opéron.

**Chez les eucaryotes**, les facteurs généraux de la transcription se fixent sur le promoteur situé en amont ou aval de la région transcrite. Le complexe de pré-initiation (PIC) est constitué des facteurs généraux sur lesquels se met en place l'ARN polymérase.

Il y aura des contacts protéine/protéine et protéine/ADN. Les facteurs de transcription sont des protéines modulaires avec un domaine d'activation DA et un domaine de fixation à l'ADN DBD.

### III. Transcription chez les eucaryotes

Alors que chez les procaryotes, transcription et traduction sont couplées, c'est-à-dire que la traduction d'un ARNm peut débuter avant que la synthèse ne soit achevée, chez les eucaryotes, la transcription a lieu dans le noyau tandis que la traduction s'effectue dans le cytoplasme de sorte que les ARN devront migrer du noyau vers le cytoplasme.

#### A. Transcription par l'ARN polymérase II

L'ARN polymérase II est constituée de 12 grandes sous-unités dont 2 sont analogues à celles des bactéries et 3 sont communes aux 3 ARN polymérases.

La particularité est la présence de la grande sous-unité : c'est un motif (CTD) de 7 acides aminés répété YSPTSPS (26 fois chez la levure et 52 fois chez l'homme). Au niveau de l'initiation, le CTD n'est pas phosphorylé mais quand l'ARN polymérase II rentre en phase d'élongation il y a phosphorylation du CTD selon des combinaisons différentes : quand S5 est phosphorylé il y a une interaction avec l'enzyme permettant l'addition de la coiffe du messenger par la guanylyltransférase (étape de maturation des pré-messagers).

Le core promoteur est situé entre -40 et +40 et renferme des éléments reconnus par les facteurs de transcription. Ils sont très diversifiés et modulables. La TATA-box n'est présente que dans 10% des gènes transcrits et elle est située à environ 20 pbs du site +1, elle peut suffire à transcrire un gène en association avec l'initiateur (InR). Le site +1 correspond d'ailleurs au « A » de l'InR. Le motif BRE en amont est reconnu par le facteur général TFIIIB de la transcription. Parfois on peut trouver DPE, et il est toujours associé à l'InR et situé en aval du +1 (+28 et +32).

Dans tous les gènes sans TATA-box, leur transcription se fait à différents endroits dans une fenêtre de 100 pbs.

#### Approche expérimentale :

Pour étudier les éléments promoteurs, il faut les caractériser. On effectue une comparaison des régions promotrices pour voir si des séquences sont conservées et sont alors impliquées dans la transcription :

- on isole le fragment d'ADN qui renferme les motifs ;
- on l'insère dans un vecteur qui est capable de se répliquer dans les cellules eucaryotes avec un gène rapporteur qui permet d'observer s'il y a transcription ou pas. Le gène rapporteur code pour la luciférase qui a une activité enzymatique avec production de lumière ;
- on fait un extrait cellulaire dans lequel on teste l'activité de la luciférase qui induit la présence d'éléments promoteurs ;
- on réalise des expériences de délétions/mutations pour savoir quel motif est important : si on perd de la lumière, on perd un motif.

Attention : lorsqu'on réalise des délétions on raccourcit les distances avec le site d'initiation ! Il vaut mieux muter chaque motif.



Etude du promoteur d'un gène codant pour un messager :

La région promotrice d'un gène est définie par rapport au site +1. Si on compare la séquence d'ADNc à la séquence du génome on peut restituer la position du site +1 dans le génome car elle correspondra à la région 5' de l'ADNc. Mais il faut que l'ADNc utilisé soit **complet**.

Pour préparer l'ADNc, on réalise tout d'abord une chromatographie avec support d'oligo-T sur lequel seuls les ARNm qui présentent un poly-A sont retenus. On hybride ensuite l'ARNm à un oligo-T complémentaire du poly-A. La transcriptase reverse (= ADN polymérase ARN dépendante) ajoutée en même temps qu'un mélange de nucléotides T, A, G et C va utiliser le poly-oligo-T comme amorce pour la synthèse d'un ADN complémentaire de l'ARNm. L'ARNm est ensuite éliminé par hydrolyse avec NaOH. L'ADNc est ensuite dupliqué sous l'influence d'une ADN polymérase I ADN dépendante qui utilise comme amorce une courte séquence auto-complémentaire toujours présente à l'extrémité 3', repliée en épingle à cheveux. La boucle est coupée par la nucléase S1, spécifique de l'ADN monocaténaire aboutissant alors à un ADNc bicaténaire.

Mais il arrive qu'il y ait un arrêt prématuré de la transcriptase inverse, l'ADNc synthétisé ne sera donc pas complet et sa région 5' ne correspondra pas au site +1 de l'ADN.

Il existe des ARNm qui ne sont pas complets et qui présentent des extrémités 5'P ou 5'OH et qui ne possèdent donc pas de « cap ». Il y a donc une méthode qui permet de ne sélectionner que les brins d'ARNm complets. (cf. document à la fin du cours)

Dans le mélange, on a une population hétérogène avec des ARNm « capés » et donc complets, des ARNm avec une extrémité 5'P et des ARNm avec une extrémité 5'OH. Tout d'abord on réalise un traitement à la *phosphatase alcaline* qui éliminera le 5'P des ARNm incomplets, on obtiendra donc des ARNm 5'OH. Ensuite on élimine la coiffe des ARNm complets à l'aide d'une *pyrophosphatase* qui coupera entre les deux premiers phosphates et qui libèrera du Gpp. Pour ajouter l'étiquette de séquence connue à l'ARNm 5'P formé, on réalisera une *ligation* entre le 5'P de l'ARNm et le 3'OH de l'étiquette. Grâce à une *transcriptase inverse*, il y aura synthèse d'un simple brin d'ADNc qui possèdera le complémentaire de l'étiquette à son extrémité 3' et un oligo-dT à son extrémité 5'. Afin de pouvoir cloner l'ADNc synthétisé un peu plus tard, on ajoute un extrinsèque au niveau du 5' qui permettra de l'insérer dans un vecteur de clonage. Après *dégradation alcaline* de l'ARNm, on *amplifie* le simple brin d'ADNc *par PCR* en utilisant une amorce complémentaire à l'étiquette 3' de l'ADNc, on obtiendra alors un ADNc double brin complet. On digère les extrémités par une enzyme de restriction Sfil qui permettra d'insérer l'ADNc dans un vecteur de clonage.

Modèle séquentiel de formation du complexe de transcription de base :

L'initiation de la transcription par l'ARN polymérase II requiert la formation d'un complexe multiprotéique sur l'ADN. Ce complexe comprend, outre la polymérase II, les facteurs généraux de transcription TFIIA, -B, -D, -E, -F et -H ainsi que le médiateur. TFIIID est constitué par la TATA Binding Protéine (TBP) et 13 à 14 facteurs associés (TAFs). TBP participant à la transcription par les trois ARN polymérases nucléaires, il est souvent considéré comme le facteur « universel ». Le facteur TBP se fixe dans le petit sillon de l'ADN et induit une courbure de celui-ci.



Il y a tout d'abord *formation du complexe de pré-initiation* (PIC) au niveau de la TATA-box en association avec l'initiateur Inr. Le CTD (domaine carboxy-terminal de la grande sous-unité de l'ARN polymérase II) n'est pas phosphorylé et le *complexe PIC est fermé*. Ensuite il y a *déroulement de l'ADN* qui induit l'*ouverture du complexe d'initiation*. Finalement il y a *phosphorylation du CTD de l'ARN polymérase II* par un complexe enzymatique contenu dans les TFIIA. Le CTD phosphorylé n'interagissant plus avec le TBP, le complexe *ARN polymérase II-TFIIIF-TFIIIE-TFIIH initie la transcription*. Les autres facteurs restent fixés au niveau de la TATA-box.

L'**élongation** est réalisée à une vitesse de 20 à 40 nucléotides/min et les gènes sont de grande taille (ex : Dytrophine 2400kb transcrit en 20h). Les *facteurs d'élongation* (13) suppriment les pauses de l'ARN polymérase II et modifient la chromatine.

La **terminaison** est encore un mécanisme inconnu. La transcription du signal de polyadénylation est couplée à l'arrêt de la transcription et les facteurs de la polyadénylation sont apportés par le CTD.

En conclusion, le core promoteur permet l'assemblage d'un complexe de transcription qui permet une transcription de base (in vitro).

Une transcription plus efficace est assurée par la présence d'un activateur (= éléments de séquence donnée pouvant être localisé à 1kb du site initiateur de la transcription).

Le site de fixation d'un facteur protéique peut être ubiquitaire ou tissu spécifique.

Le **médiateur** de l'activation de la transcription, un complexe de 25 protéines, fait le lien entre les activateurs se liant à l'ADN et la Pol II. Il établit le contact entre tous les facteurs de la transcription.

#### Promoteurs des gènes ARNs transcrits par l'ARN polymérase II :

Deux éléments sont localisés en amont du +1 :

- PSE (-58pb) : c'est le site de reconnaissance d'un facteur de transcription spécifique SNAPc ; cette séquence assure la transcription de base et est toujours associée à un élément distal DSE ;
- DSE (-100 à -200pb) : c'est un site sur lequel se fixe un activateur donné qui stabilise SNAPc au niveau du promoteur de base assurant une transcription normale des gènes

En 3' de la région transcrite, il y a un élément de séquence qui assure la coupure du transcrit primaire.

## B. Transcription des gènes par l'ARN polymérase I

L'ARN polymérase I est responsable de la transcription des ARNr 18S, 5,8S et 28S. Ces gènes sont répétés dans le génome :



Le +1 est un initiateur de la transcription au niveau du core promoteur, SL1 se met en place au niveau de ce core promoteur. La fixation d'UBF sur UPE permet une transcription optimale. Ce motif UPE n'est pas indispensable à la transcription mais la distance entre celui-ci et le core promoteur est importante.

Dans la région en aval du +1 il y a un élément qui assure la terminaison : c'est un terminateur qui constitue le site de fixation d'une protéine de la terminaison.

Cette synthèse est effectuée au niveau du nucléole dans lequel se concentrent tous les gènes qui codent pour l'ARNr. C'est une région dans laquelle la chromatine est décondensée.

Lorsque l'ARN polymérase I entre en phase d'élongation, les facteurs généraux de la transcription restent en place au niveau du promoteur.

### C. Promoteurs des gènes transcrits par l'ARN polymérase III

Il existe des promoteurs internes (à l'intérieur de la région transcrite) et externes de 3 types :

- **Type I** *ARNr5s* (interne) : possède 3 séquences conservées pour un organisme donné A, IE et C. Ces séquences constituent le site de fixation des facteurs généraux de la polymérase III :
  - **TFIIIA** sur A et C puis
  - **TFIIIC** se met en place puis
  - **TFIIIB** se fixe en amont du +1, il est incapable de se fixer seul mais une fois fixé il ne se détache plus.

Cette plateforme protéique sera reconnue par l'ARN polymérase III.

Il peut y avoir des contacts entre les différents facteurs de transcription. Ceux-ci restent aussi en place lorsque la polymérase III entre en phase d'élongation.

La courbure de l'ADN induit une transcription rapide car lorsque l'ARN polymérase III se détache de l'ADN elle est tout de suite reprise et réutilisée au site +1. Le signal de terminaison de l'ARN polymérase III est une succession de T.

- **Type II** *ARNt* (interne) : possède 2 boîtes entre lesquelles il y a des introns ce qui signifie que la distance entre les séquences A et B peut être variable. La fixation de **TFIIIC** au niveau du promoteur permet la fixation de **TFIIIB**. Ce complexe sera reconnu par l'ARN polymérase III. Le complexe pic est mis en place et il y a initiation de la transcription.
- **Type III** *hARNsnU6* (externe) : il est situé en amont du site +1 et possède 3 éléments : **TATA** et **PSE** qui constituent le promoteur de base et **DSE** en amont. PSE seul constitue le signal de reconnaissance de la polymérase II, TATA seul constitue aussi le signal de reconnaissance de la polymérase II mais l'ensemble des 2 éléments constitue le signal de reconnaissance pour la polymérase III. On retrouve SNAPc (qui chasse TFIIID) et TFIIIB' qui entraînent la fixation de la polymérase III.

La mise en évidence des interactions protéines/protéines au niveau des complexes de transcription peut se faire par deux méthodes in vitro :

- **Immuno-précipitation**
- **Gst pull down**

Immuno-précipitation :

Pour savoir s'il y a des interactions avec les facteurs 1 et 2 on réalise une chromatographie avec comme support un anticorps spécifique (IgG) du facteur 1.

La protéine 1 sera donc reconnue et fixée lors de son passage dans la colonne.

La 2<sup>ème</sup> protéine ajoutée sera soit marquée radioactivement, soit incluse dans un extrait nucléaire.

Témoin : on fait passer le facteur 2 uniquement sur un support IgG sans facteur 1.

- Dans le cas où le facteur 2 est radioactif, s'il interagit avec le facteur 1 il restera fixé sur la colonne, sinon il sera éliminé. On lave et après ajout d'une solution dénaturante on dépose l'extrait sur un gel de polyacrylamide SDS pour effectuer une électrophorèse en conditions dénaturantes. On observera un signal uniquement si la 2<sup>ème</sup> protéine est fixée au facteur 1.

- Dans le cas où le facteur 2 n'est pas radioactif, on réalise un Western Blot (=transfert sur un champ électrique) et on révèle le facteur avec un anticorps radioactif contre lui.

Gst pull down :

Le facteur 1 est inséré dans un vecteur de clonage en 3' de l'ADNc codant pour la Gst. Le vecteur est ensuite introduit dans des bactéries qui synthétisent une protéine chimère qui porte le Gst dans la région N-term et le facteur d'intérêt dans la région C-term. On fait une chromatographie d'affinité avec comme support du glutathion sépharose. La protéine chimère peut donc se fixer spécifiquement sur la colonne grâce à la forte affinité de Gst pour son substrat. On ajoute ensuite le facteur 2 qui ne se fixera que s'il interagit avec le facteur 1. On obtiendra alors un marquage sur gel dénaturant.

Témoin : on fait passer le facteur 2 uniquement sur un support contenant du glutathion sépharose.

## D. Effets de la chromatine sur la transcription

Il existe différents domaines de chromatine :

- Domaines actifs : euchromatine avec possibilité de transcription ;
- Domaines inactifs : hétérochromatine sans transcription car c'est un domaine de forte condensation.

Les histones peuvent être modifiées : on parle de « code des histones »

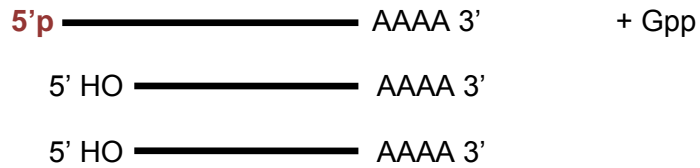
Lorsqu'une « queue » sort (ex : H3 sur le doc.) elle est accessible dans le nucléosome et peut subir des modifications : les lysines de cette queue peuvent être acétylées et dans ce cas on a une transcription positive, les lysines peuvent être méthylées et dans ce cas il n'y a pas de transcription car une autre protéine qui reconnaît cette méthylation va condenser l'ADN.



Traitement à la phosphatase alcaline



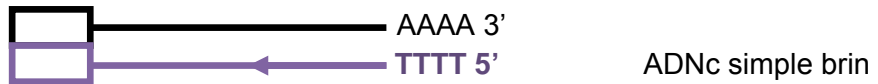
Traitement à la pyrophosphatase



Ajout de l'étiquette et ligation  
(uniquement au niveau des 5'P)



Transcriptase inverse



Dégradation alcaline & Amplification par PCR

Amorce complémentaire de l'étiquette de l'ADNc synthétisé



Comment sélectionner des brins d'ARNm complets avant la synthèse de l'ADNc ?