

Université Frères Mentouri Constantine 1
Faculté Sciences de la Nature et de la Vie
Département de Microbiologie

Organisation des complexes de transcription
chez la levure

Support pédagogique destiné aux étudiants Master 1, spécialité Biologie Moléculaire des Microorganismes

Réalisée par : *P^R ALATOU R.*

1. Introduction générale

La cellule est l'unité fondamentale du vivant. En son sein se déroule un ensemble de réactions chimiques régulées appelé métabolisme qui lui permet de croître et de se reproduire. La cellule est composée de macromolécules très diverses en composition, en taille et en fonction. Soit elles sont nécessaires à la structure de la cellule, soit elles représentent un apport énergétique, soit elles confèrent une fonction spécifique à la cellule. La plupart de ces éléments constitutifs et fonctionnels sont synthétisés grâce à l'expression de l'information génétique.

Les réactions qui gouvernent cette expression de l'information génétique font partie du métabolisme de la cellule. Ces réactions sont, dans un premier temps, la transcription de l'acide désoxyribonucléique (ADN) en acide ribonucléique (ARN) et, dans un second temps, la traduction des ARN dits messagers (ARNm) en protéines. A ces réactions s'ajoute la réplication (ou duplication) de l'ADN, synthèse d'ADN à partir d'ADN, et la transcription inverse (ou reverse), synthèse d'ADN à partir d'ARN (figure 1).

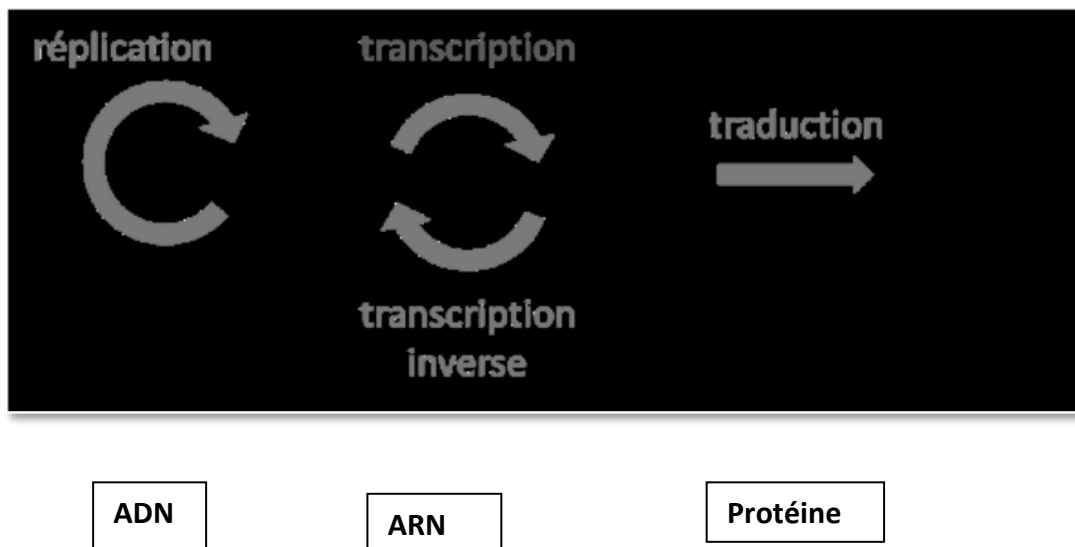


Figure 01 : Les différents types de transfert de l'information génétique dans une cellule

2. Quelques mots sur la levure *Saccharomyces cerevisiae*

Saccharomyces cerevisiae, appelée « levure de boulanger » ou aussi « levure de bière ». Cette levure est reconnue comme un organisme modèle d'études, représentant un eucaryote simple dont le génome est facilement manipulable. La levure partage avec la bactérie de nombreux avantages techniques qui ont permis une avancée rapide de leur étude génétique et moléculaire : une croissance rapide, des cellules dispersées, une facilité de réplication et d'isolation sur milieu solide, un système génétique bien défini et, surtout, un système de transformation de l'ADN très flexible. D'autant plus que ce microorganisme n'est pas pathogène. Elle devient alors un organisme modèle-clé dans les études génomiques. Ce

champignon (fungus), de la famille des ascomycètes. Son génome de 12 Mb contient au moins environ 6 000 gènes dont la plupart trouvent leurs orthologues chez l'homme (3 200 Mb, environ 30 000 gènes). On comprend alors que l'étude du génome de *S. cerevisiae* soit d'un aussi grand intérêt pour l'homme.

Le développement de la transformation de l'ADN a rendu la levure particulièrement abordable au clonage de ses gènes et aux techniques d'ingénierie génétique. De l'ADN peut y être introduit via des plasmides autonomes ou intégré dans le génome par recombinaison homologue. La levure a largement été utilisée en génétique classique et en génétique inverse, contribuant à l'étude de gènes jusque -là encore inaccessible.

Les techniques de biologie moléculaire ont largement été exploitées chez *S. cerevisiae* dans l'étude de la régulation de gènes, la relation structure-fonction des protéines, la structure des chromosomes et autres grandes questions de la biologie cellulaire.

3. La formation de l'ARN prémessager

La transcription des gènes codant pour les chaînes polypeptidiques chez les eucaryotes se déroule en deux étapes : formation d'un ARN prémessager puis maturation de cet ARN prémessager pour former un ou plusieurs ARN messagers.

3.1 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 etc... 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**.

➤ Le promoteur

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 sont surnommées 'boîtes', elles ont une importance particulière dans ce processus, particulièrement parce que ces séquences sont reconnues spécifiquement par différentes protéines appartenant au complexe d'initiation :

La boîte TATA qui est riche en T et A, la plus importante est située vers -25 à -30 nucléotides du site de démarrage de la transcription (noté +1)

Des éléments proximaux :

-La boîte CAAT qui est facultative, contenant de la C, est située vers -120 à -80 du site de démarrage du site de la transcription.

-La boîte GC qui est également facultative riche en G et C, peut être

présente entre la boîte CAAT et la boîte TATA.

Ces séquences sont souvent bien conservées, une variabilité non négligeable peut également être observée. Ainsi il existe des promoteurs sans la boîte TATA.

➤ Le complexe d'initiation

Contrairement à ce qui se produit chez procaryotes, l'ARNpol II des euracyotes ne reconnaît pas seule le promoteur proximal. Elle effectue ce travail en compagnie de nombreux cofacteurs protéiques qui se recrutent les uns les autres et qui forment avec elle un complexe d'initiation. Ces facteurs sont notés **TFIIA**, **TFIIB** etc...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'ARNpol II. La séquence d'assemblage du complexe d'initiation est décrite sur la figure 2.

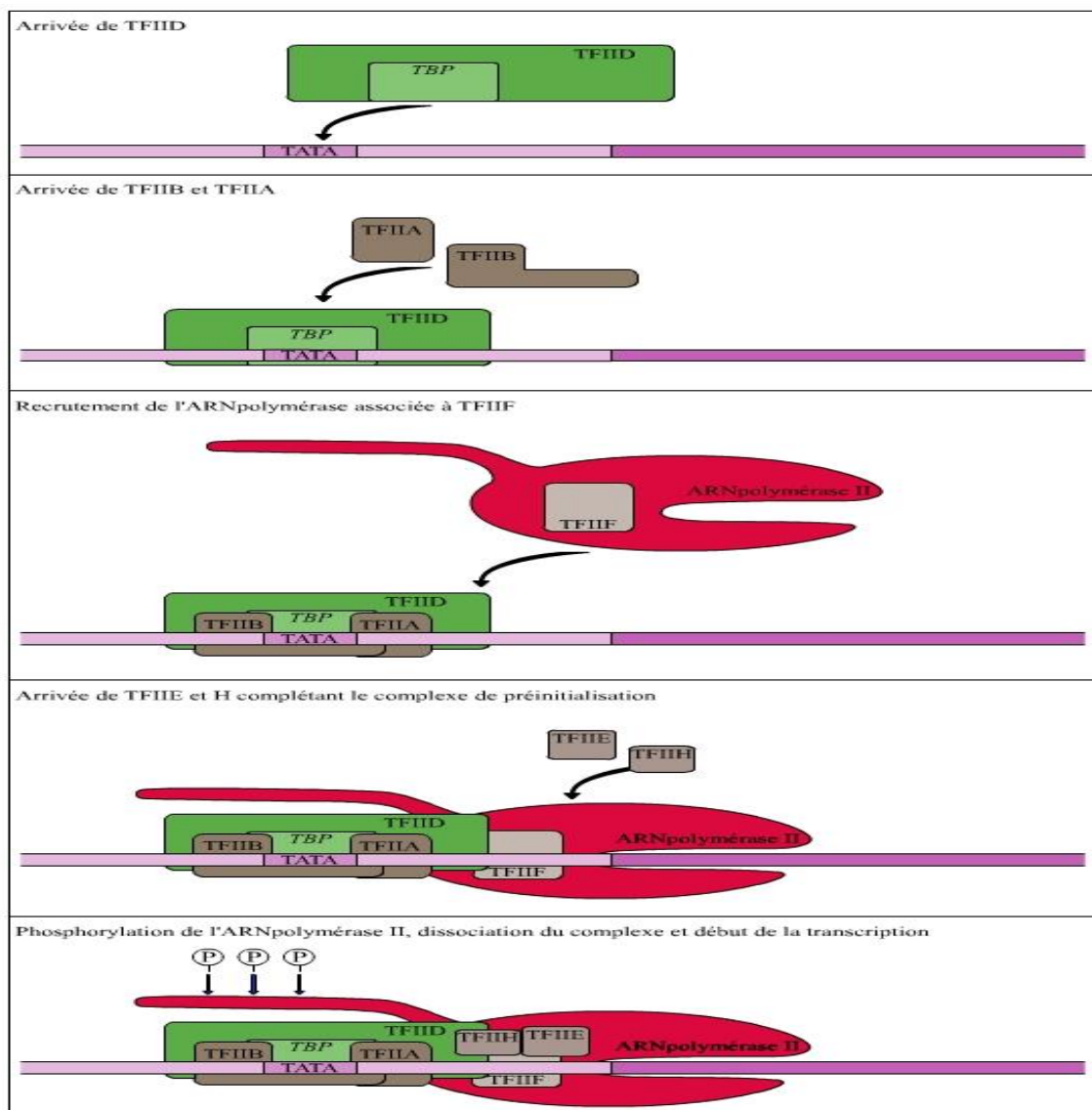


Figure 02 : Initiation de la transcription

La TBT pour TATA box-Binding protéine (protéine de liaison à la boîte TATA) 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 TFIID (Transcription Factor for ARN polymerase II ; facteurs généraux de la transcription de l'ARN polII) comporte plusieurs activités enzymatiques dont une activité hélicase permettant l'ouverture de la double hélice de l'ADN au niveau du promoteur, et une activité kinase responsable de la phosphorylation de la queue C-terminale de l'ARN polymérase II. Cette phosphorylation provoque une modification de la structure tridimensionnelle de l'ARN polymérase qui entraîne la dissociation du complexe d'initiation et de la transcription.

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.

➤ Intervention de facteurs spécifiques de la transcription

Le complexe d'initiation composé de l'ARNpol II et des différents TFII est suffisant pour obtenir une activité transcriptionnelle. L'augmentation de cette activité (ou sa repression) est sous la dépendance de facteurs spécifiques qui vont interagir avec le complexe d'initiation. Ces protéines activatrices ou inhibitrices (éléments trans-regulateurs) se lient à des promoteurs distaux spécifiques de l'ADN, appelées **amplificateurs (enhancer)** lorsqu'ils recrutent des cofacteurs activateurs, ou **silenceurs (silencers)** lorsqu'ils recrutent des cofacteurs inhibiteurs. Ces promoteurs distaux peuvent être 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 par le jeu de courbure de l'ADN, des facteurs de transcription et du médiateur qui maintient liés tous ces facteurs (figure 03)

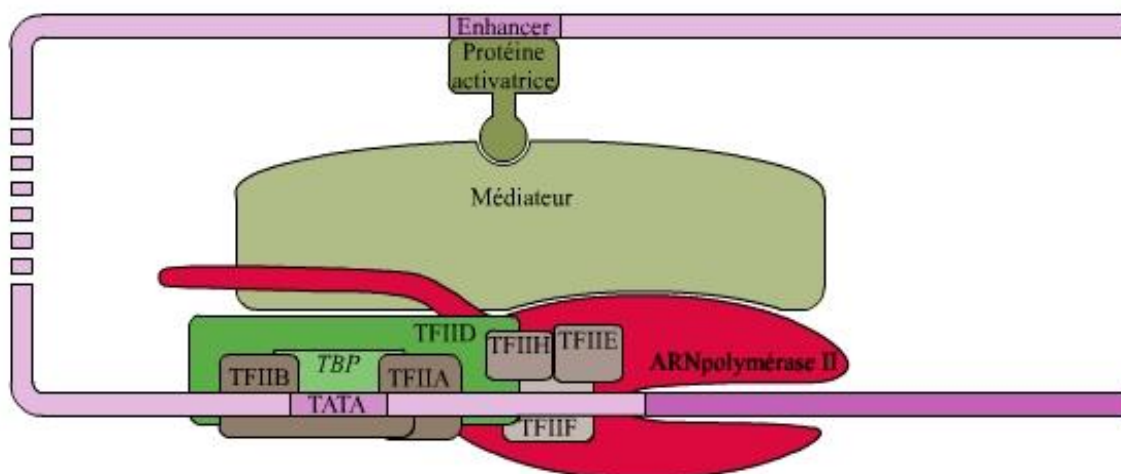


Figure 03 : Facteurs spécifiques de la transcription

Les amplificateurs (enhancers) sont des séquences situées parfois à plusieurs milliers de nucléotides d'une promoteur et qui, par un jeu d'interactions protéiques, stabilisent le

complexe d'initiation, favorisant ainsi la transcription. Ces interactions sont rendues possibles par la courbure de l'ADN qui rapproche des éléments situés à de grandes distances les uns des autres. Cette activation est spécifique du gène et utilise une multitude de facteurs de transcription spécifiques (les protéines activatrices) qui agissent généralement sous forme dimérique. La répression spécifique peut agir selon le même type de modalité.

3.2 L'élongation

L'ARNpol II est équipée de facteurs protéiques d'élongation qui facilitent sa progression à travers d'une chromatine dont ils relâchent la structure. Un ARN pré-messager complémentaire d'un brin matrice de l'ADN (brin antisens), donc identique au brin codant de l'ADN, aux riboses et uraciles près, commence à être synthétisé selon la direction 5'-3' (figure 04). Le brin d'ARN néosynthétisé est donc identique au brin codant d'ADN, aux uraciles et riboses près.

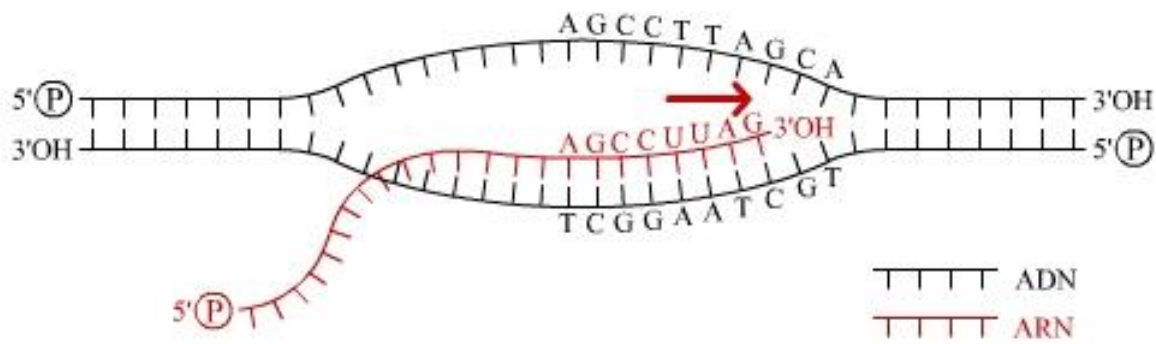


Figure 04 : L'élongation

3.3 La terminaison

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

3.4 La formation d'un ou plusieurs ARNmessenger (s)

Le transcrit primaire n'est pas utilisé tel quel pour la synthèse protéique (la traduction). Il doit subir des modifications qui répondent à plusieurs impératifs (augmentation de la demi-vie, modification de la séquence). Toutes ces modifications sont réalisées au fur et à mesure de la progression de la synthèse du préARNm dans le nucléoplasme. Il existe trois grands types de modification, catalysées chacune par des enzymes de nature protéique ou ribonucléique.

3.5 L'addition d'une coiffe en 5'

Elle a lieu dès le début de la transcription avant que la chaîne ne compte plus de 30 nucléotides. Elle consiste en l'ajout d'un nucléotide à la G sur l'extrémité 5' de l'ARN suivi de sa méthylation sur l'azote 7 de la base, ainsi que la méthylation en 2' du ribose du premier

ou des deux premiers nucléotides du transcrit primaire. La particularité de cet ajout consiste dans le type de liaison mis en jeu (figure 05). Il en résulte que l'extrémité 5' de l'ARNm n'est pas porteuse des trois acides phosphoriques libres habituel, mais d'un GMP, ce qui limite la réactivité de cette extrémité et sa reconnaissance par les exonucléases (protection contre la dégradation). Cet ensemble sert de coiffe protectrice à l'extrémité 5' de l'ARNm. Elle est également nécessaire à l'exportation de l'ARNm vers le cytoplasme et à la liaison de ce dernier avec la petite sous-unité du ribosome, lors de l'étape d'initiation de la traduction.

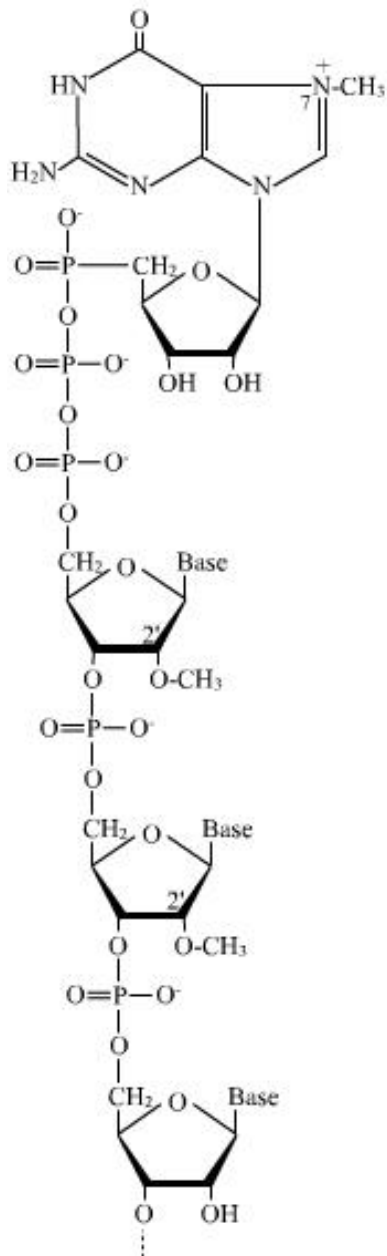


Figure 05 : Formation de la coiffe

3.6 L'excision –épissage

Chez les eucaryotes, les archéobactéries et les cyanobactéries, les gènes sont morcelés : constitués d'une alternance d'exons (parties codantes du gène) et d'introns (parties non codantes, bornées par des séquences de bases spécifiques : 5'GU et 3'AG), ils sont d'abord intégralement recopiés dans l'ARNpm, puis subissent une opération d'excision des introns, suivie d'un épissage (splicing), qui veut dire la réunion bout à bout des exons restants qui constituent l'ARNm. Ce remaniement se déroule au fur et à mesure de la progression de la transcription. L'excision des introns s'opère par l'entremise d'une formation dite 'en lasso' (figure06).

En résumé le phénomène d'excision-épissage entraîne la séparation de l'intron d'avec l'exon 1 (situé en amont) et la formation d'une structure en lasso interne à l'intron. Ensuite, l'extrémité 3' de l'exon 1 réagit avec l'extrémité 5' de l'exon 2 permettant l'épissage des deux exons et la libération du lasso qui sera dégradé par des ribonucléases.

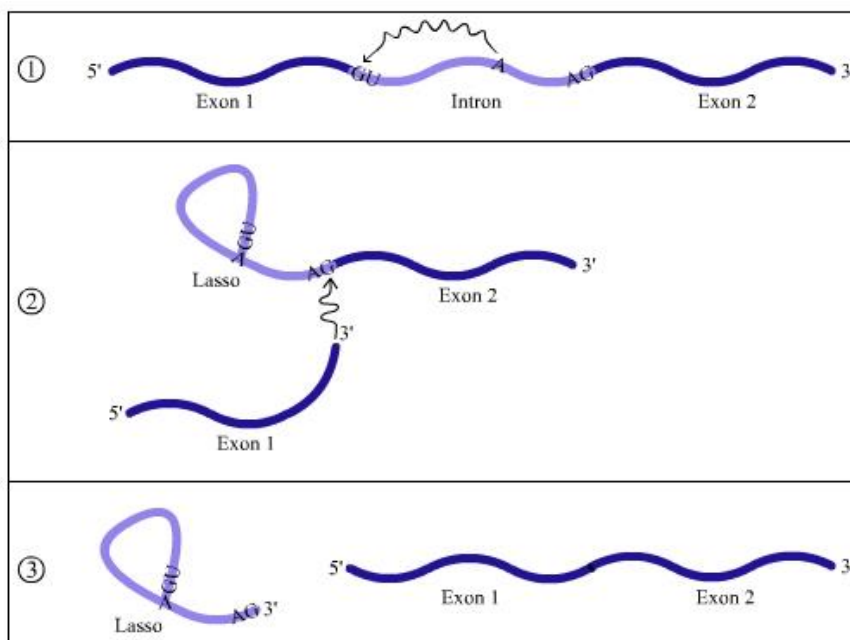


Figure 06 : Mécanisme de la formation du lasso

La formation du lasso (Splicésome) fait appel à des facteurs spécifiques : ribonucléoprotéines contenant des petits ARN (snRNP), ribozymes (ARN ayant des propriétés catalysantes), des enzymes spécifiques (maturases). La structure secondaire du transcrit met en contact trois séquences de l'intron : la séquence du début de l'intron (extrémité 5'-GU ou site donneur), la séquence de la fin de l'intron (extrémité 3'-AG ou site accepteur) et un nucléotide à A et un nucléotide à A (A du branchement) situé à 40 nucléotides avant le site receveur (figure 07).

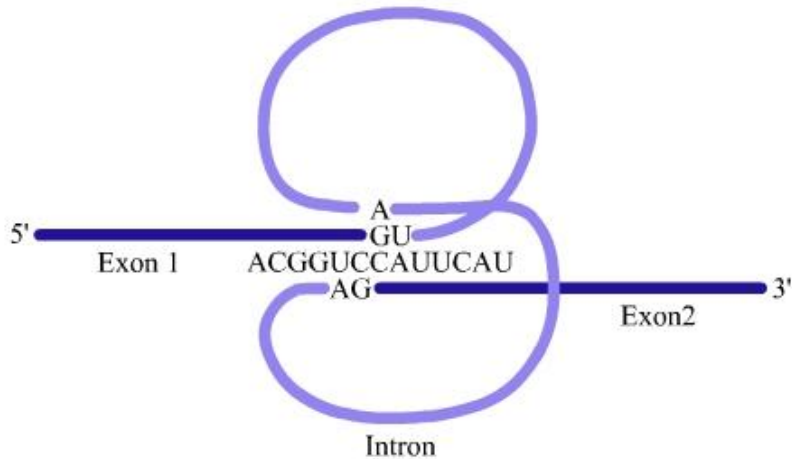


Figure 07 : Splicéosome

3.7 Addition d'une queue polyA en 3'

Le site de polyadénylation est codé au niveau du gène. Le site de clivage déterminé par le dinucléotide CA est entouré par une séquence AAUAAA très conservée, située entre 10 à 30 nucléotides en amont du site de clivage, et par séquences DES (Down Stream Element) riche en U ou en GU, situé à une trentaine de nucléotides en aval du site de clivage. La séquence AAUAAA est reconnue spécifiquement par un complexe protéique appelé CPSF (Cleavage and Polyadenylation Specific Factor), et la séquence DES par un complexe CstF (Cleavage Stimulation Factor). Ces deux complexes et d'autres composants, comme l'ARN polymérase II et la poly (A) polymérase (PAP) vont interagir en formant le complexe de clivage qui va cliver la molécule du pré-ARN au niveau du site de clivage (Figure 08).

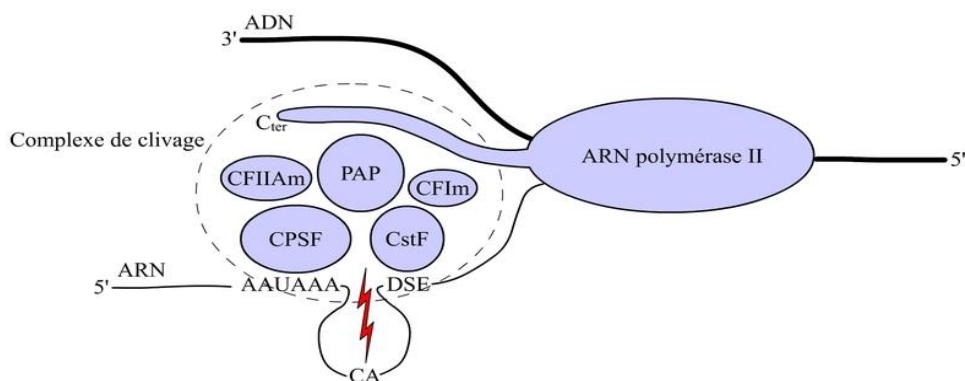


Figure 08 : Formation de la queue poly A

La PAP va alors ajouter environ 200 nucléotides, une poly (A) binding protéine II (PABP2) qui interagit avec la PAP se chargeant d'activer l'enzyme et de contrôler la longueur de cette queue poly (A). Il est intéressant de noter que cette polymérase n'utilise pas de matrice ADN pour créer cette séquence poly (A). La queue poly (A) n'est donc pas codée par un e génome. Notons que ce n'est pas un exemple isolé, d'autres mécanismes comme l'édition ou la modification enzymatique de certaines bases (C, U par exemple) étant encore beaucoup plus intrigants. Cette queue poly(A) confère de la stabilité au futur ARNm et se perd au fur et à mesure qu'il traduit.

Université Frères Mentouri Constantine 1
Faculté Sciences de la Nature et de la Vie
Département de Microbiologie

Remodelage de la chromatine

Support pédagogique destiné aux étudiants Master 1, spécialité Biologie Moléculaire des Microorganismes

Réalisée par P^R ALATOU R.

Introduction

Le génome de la levure, comme celui des autres eucaryotes, a une structure plus complexe que celui des bactéries. Tout d'abord, l'ADN est condensé dans l'espace nucléaire sous forme de chromatine. Tout en structurant le génome, la chromatine doit laisser un accès à l'ADN. On trouve chez les eucaryotes de nombreux complexes protéiques qui ont pour rôle de lever la barrière chromatiniennemais aussi de participer au maintien de cette structure. La modulation de cette structure fait partie intégrante des mécanismes de régulation de l'expression des gènes.

Dans les cellules eucaryotes, le matériel génétique est organisé en une structure complexe constituée d'ADN et de protéines qui est localisée dans un compartiment spécialisé, le noyau. Cette structure a été baptisée '**Chromatine**'. Environ deux mètres d'ADN dans chaque cellule doit être contenu dans un noyau. En plus de cet énorme degré de compaction, l'ADN doit être rapidement accessible afin de permettre son interaction avec les machineries protéiques régulant les fonctions de la chromatine : Réplication, réparation et recombinaison (figure 01). Donc toute cette organisation dynamique de la structure de la chromatine influence potentiellement toutes les fonctions du génome.

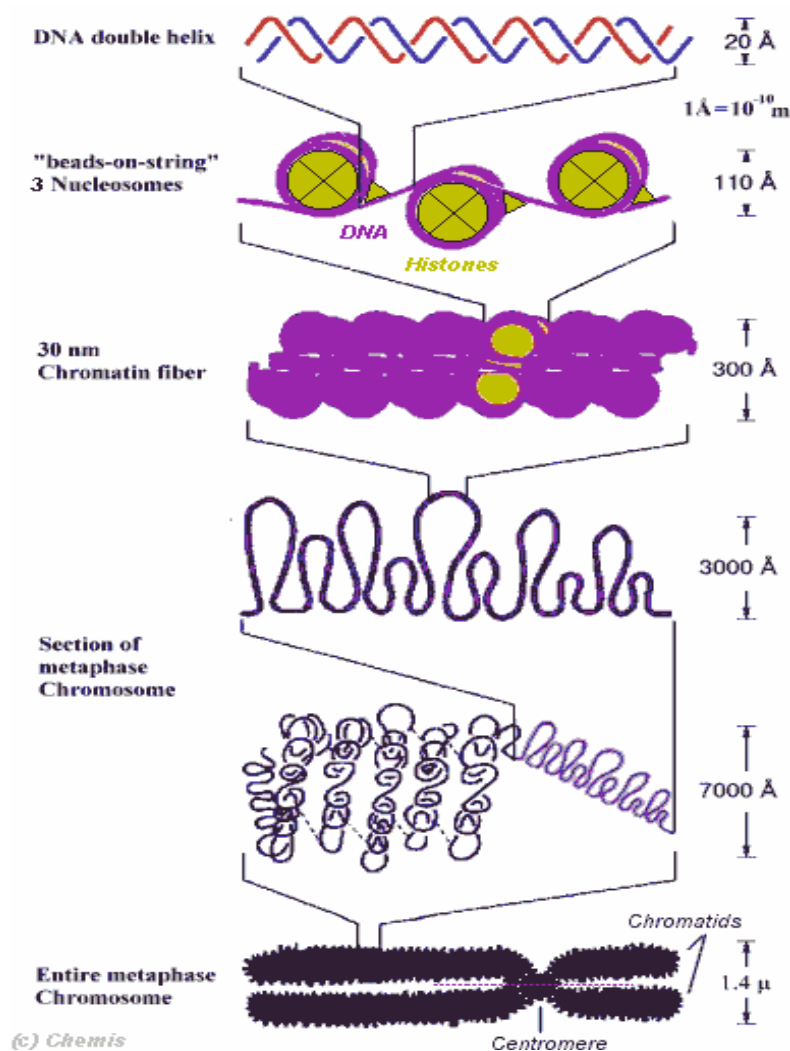


Figure 01 : Les différents niveaux de compaction de la chromatine

1. Définition de la chromatine

Chez les eucaryotes, le phénomène de condensation de la molécule d'ADN à des niveaux très élevés en donne la **Chromatine**. Il existe des protéines associées à cette structure d'ADN nommées '**Histones**'. Les histones sont définies comme étant des protéines de charge positive ce qui favorise leur liaison avec les molécules d'ADN qui sont chargées négativement.

Deux types de protéines sont associés à l'ADN :

- ✓ Les histones: H2A, H2B, H3, H4, H1. $6 \cdot 10^7$ molécules/cellule. Ils jouent un rôle dans la compaction de l'ADN et la régulation de la transcription
- ✓ les protéines non-histones: 10^4 molécules/cellule, ont un rôle dans la réplication, la transcription et la régulation de l'expression génique

2. Nucléosome

Figure 03 : Octamère d'histone

L'unité fondamentale de la chromatine est le **nucléosome** qui est composé d'ADN et d'histones (figure 02). Il constitue le premier niveau de compaction de l'ADN dans le noyau. Cette structure est par la suite régulièrement répétée pour former le nucleofilament qui peut, lui-même adopter des niveaux d'organisation plus compacts.

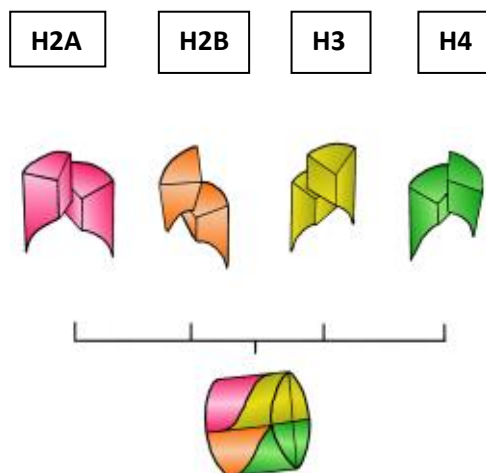


Figure 02 : Structure du nucléosome

Un nucléosome est constitué de 2 dimères H3H4 (un tétramère) et de 2 dimères H2A H2B (un tétramère) (**figure 03**)

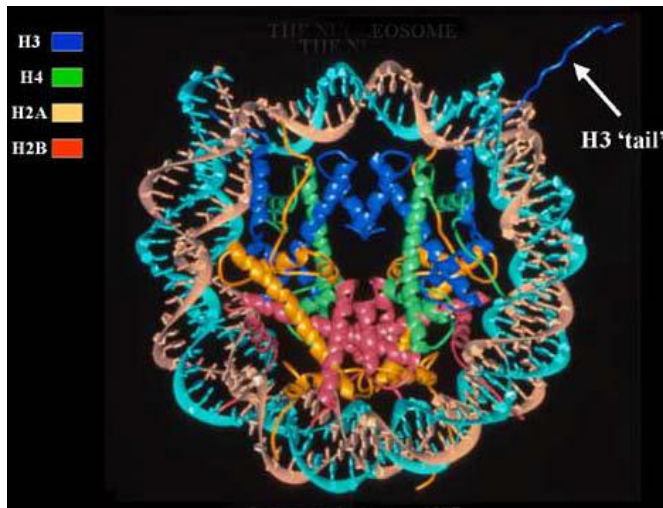


Figure 3 : Un tétramère

La chromatine a été divisée en : **Euchromatine** et **hétérochromatine**.

L'hétérochromatine a été définie comme une structure stable qui ne se change pas d'état de condensation où est localisée la région silencieuse de l'ADN, par contre l'euchromatine apparaît décondensée indiquant des fragments d'ADN en transcription ou en réplication.

3. Les facteurs d'assemblage de la chromatine

3.1. Les facteurs interagissant avec les histones

Des facteurs de nature acide peuvent former des complexes avec les histones et ainsi favoriser leur mise en place. Ces facteurs sont considérés comme des chaperon d'histones qui facilitent la formation de la partie nucléosomale sans faire partie intégrante de cette partie. Ces facteurs sont nommés également '**facteurs d'assemblage de la chromatine**'. Nous pouvons citer le facteur CAF-1 (Chromatin Assembly Factor 1) qui interagit avec les histones H3 et H4 acétylées nouvellement synthétisées, et participe à l'assemblage de la chromatine couplé à la réplication de l'ADN.

Les différentes modifications des histones pourraient donc constituer un signal de ciblage et d'ancrage d'activité de remodelage vers certaines régions chromatiniennes.

a. L'acétylation des histones

La réaction d'acétylation est un transfert d'un groupement acétyl depuis l'acétylcoenzyme A sur un résidu lysine de la protéine cible. Dans le cas des histones, elle est catalysée par les Histone Acétyl-Transférases (**HAT**). Ces enzymes ont pour substrats privilégiés les histones H3 et H4.

b. La méthylation

La méthylation est un transfert d'un groupement méthyle depuis la S-adenosylméthionine vers la protéine cible. Les histones peuvent être méthylées par des histone méthyltransférases (**HMT**) au niveau de résidus lysine ou arginine. Les premières histone méthyltransférases agissant sur les résidus lysine à avoir été isolées sont les protéines Suv39H1 humaine (Aagaard et al., 1999). Ces protéines méthylent *in vitro* la lysine 9 de l'histone H3 (Rea et al., 2000). L'histone H3 peut également être méthylée sur la lysine 4. Il existe d'autres HMT nommées **PRMT1** (Protein Arginine Methyltransferase 1) possédant une activité ciblée spécifiquement sur les résidus arginine.

Comme l'acétylation, la méthylation des histones a été corrélée à la régulation de la transcription. Cependant, la méthylation des histones n'a pas comme conséquence systématique une activation transcriptionnelle. Le résidu, son emplacement, le nombre de méthylation ainsi que la combinaison avec d'autres modifications covalentes des histones déterminent l'effet transcriptionnel.

C. la phosphorylation

La phosphorylation fait intervenir des kinases spécifiques aux histones. La phosphorylation touche la sérine 10 de l'histone H3. La phosphorylation des histones a aussi été impliquée dans l'activation transcriptionnelle (revues dans Spencer and Davie, 1999 ; Berger, 2001).

4. Remodelage de la chromatine par des complexes ATP-dépendants

Des facteurs interviennent également pendant l'étape de maturation de la chromatine en organisant et en maintenant un état bien défini de cette dernière. Ces facteurs peuvent affecter la chromatine en induisant des changements conformationnels au niveau du nucléosome mais également au niveau de larges domaines de la chromatine. Ces facteurs existent en deux catégories : les premiers nécessitent un apport énergétique sous forme d'ATP et sont généralement appelés **facteurs de remodelage de la chromatine**, et les seconds sont des enzymes qui modifient post-traductionnellement les histones.

Le complexe SWI/SNF

➤ Isolement du complexe

Le complexe SWI/SNF (pour mating type SWItching/Sucrose Non Fermenting) est le premier complexe de remodelage de la chromatine à avoir été isolé chez la levure. Plusieurs composants de ce complexe ont été originellement identifiés dans des cribles génétiques comme des régulateurs positifs de la transcription du gène de l'endonucléase HO (gènes de la famille SWI) ou de l'invertase SUC2, requise pour la croissance sur saccharose (gènes de la famille SNF ; pour revue, voir Winston and Carlson, 1992). Des études génétiques et biochimiques ont par la suite montré que ces différentes protéines forment un même complexe, composé d'au moins 11 sous-unités et d'une masse moléculaire d'environ 2 MDa (Côté et al., 1994). Son mécanisme d'action repose sur la mise en place d'un activateur transcriptionnel ce qui permet l'accessibilité de l'ADN qui permet le changement du positionnement et de la structure des nucléosomes. L'étude de mutants du complexe SWI/SNF a permis d'établir un lien entre ce complexe et le remodelage de la chromatine. La perte de fonction de SWI/SNF se traduit par des phénotypes très variés. Des études récentes par puces à ADN ont montré que SWI/SNF n'a pas un rôle général dans la transcription. Il n'est requis (directement ou indirectement) pour la transcription que de 3 à 6 % des gènes selon les études (Holstege et al., 1998 ; Sudarsanam *et al.*, 2000). De plus, SWI/SNF a un rôle positif et négatif sur la transcription. Ceci indique que ce complexe contribue à rendre l'ADN plus ou moins accessible selon les gènes.

Le fait que SWI/SNF ne soit pas requis pour la transcription de la totalité du génome indique soit qu'il n'intervient effectivement qu'au niveau d'un nombre restreint de gènes, soit que sa fonction est partiellement redondante avec celle d'un autre facteur (de remodelage ou de modification des histones).

Un ensemble d'approches, menées aussi bien *in vivo* qu'*in vitro*, a permis d'établir que la formation de la particule nucléosomale procède en deux étapes. Un tétramère d'histone (H3/H4)₂ entre d'abord en interaction avec l'ADN, puis deux dimères H2A/H2B viennent s'adjoindre (figure 04)

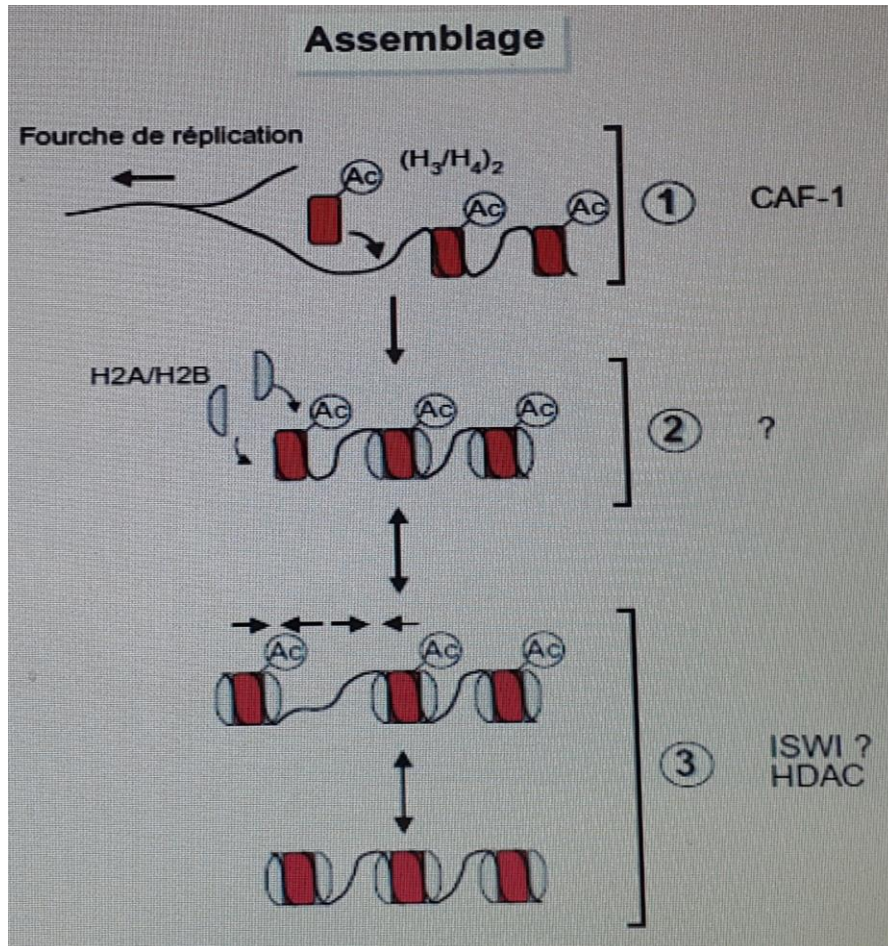


Figure 04 : L'assemblage des nucléosomes. Au niveau de la fourche de réplication, l'assemblage des nucléosomes de novo procède en plusieurs étapes qui sont schématisées dans cette figure. **A.** Un tétramère d'histones (H3/H4)₂ nouvellement synthétisées (sous forme acétylée) est d'abord associé à l'ADN naissant. Cette étape est stimulée par le facteur d'assemblage CAF-1 (chromatin assembly factor 1). **B.** Deux dimères (H2A/H2B) sont ensuite ajoutés, et complètent la particule nucléosomale. **C.** Enfin, une étape dite de «maturation» permet l'espacement régulier des nucléosomes et généralement la désacétylation des histones nouvellement incorporées. L'obtention d'une telle régularité pourrait impliquer, entre autres, des facteurs dits de «remodelage».

CAF-1 représente un candidat idéal pour promouvoir l'assemblage de nucléosomes *de novo* couplé à la réplication et à la réparation de l'ADN. Pourtant, des analyses génétiques, chez la levure, montrent que les gènes codant pour les sous-unités de CAF-1 (CAC1, CAC2, CAC3 – CAC pour *chromatin assembly complex*) ne sont pas essentiels. Ceci implique soit l'existence de voies d'assemblage alternatives, soit qu'un assemblage non assisté par des chaperons puisse suffire dans des conditions normales (figure 05).

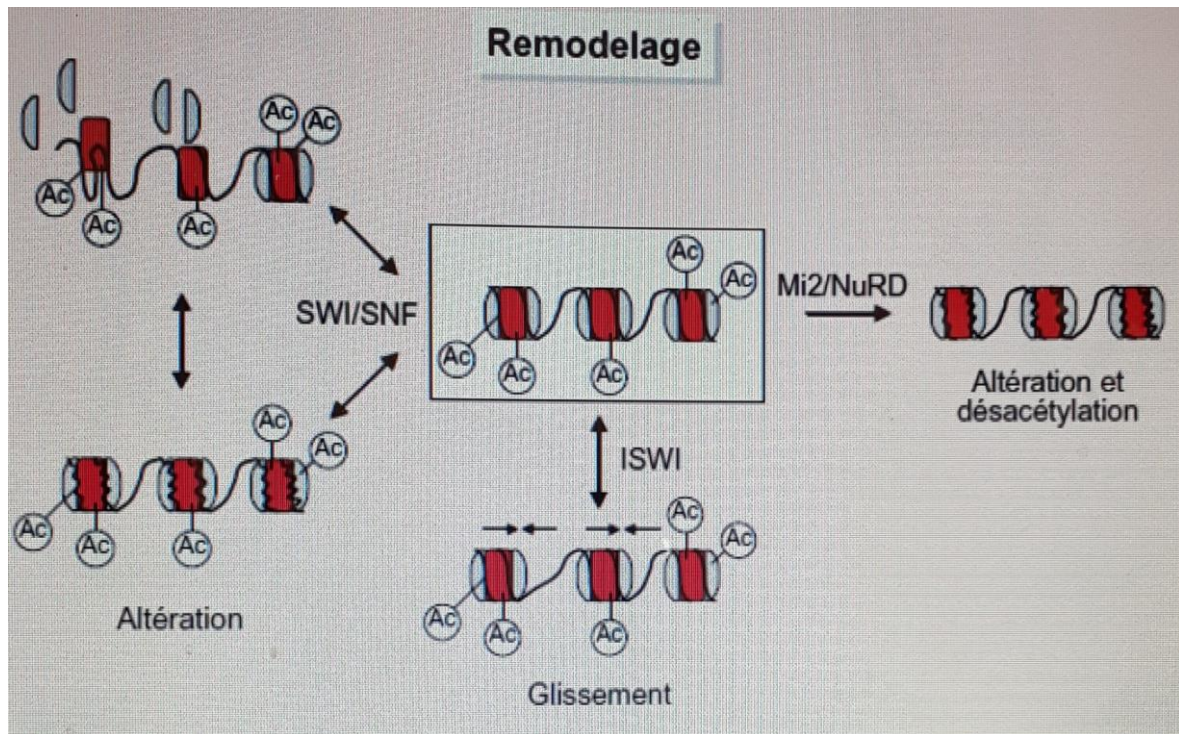


Figure 05. Mécanismes de remodelage des nucléosomes: des modèles. Le remodelage du nucléosome représente l'ensemble des réarrangements structuraux qui rendent l'ADN nucléosomal potentiellement plus accessible.

Sur ce schéma sont représentés les modèles correspondant aux principaux processus de remodelage catalysés par les facteurs du type SWI/SNF, ISWI et Mi2/NuRD. Ces facteurs sont des complexes multiprotéiques qui contiennent tous une activité ATPase critique pour le remaniement de l'organisation nucléosomale. Les facteurs de la famille SWI/SNF altéreraient les contacts histones/ADN. La structure ainsi plus fluide pourrait osciller entre une forme altérée et la forme d'origine. Un mécanisme de « glissement », entraînant le déplacement des octamères d'histones en cis le long de l'ADN, est proposé pour la famille ISWI. Enfin, dans le cas de la famille Mi2/NuRD, une altération de l'organisation nucléosomique serait accompagnée d'une désacétylation des histones, catalysée par les sous-unités histone désacétylases (HDAC1/2).

Université Frères Mentouri Constantine 1
Faculté Sciences de la Nature et de la Vie
Département de Microbiologie

Contrôle de l'arrêt de la transcription chez
Saccharomyces cerevisiae

Support pédagogique destiné aux étudiants Master 1, spécialité Biologie Moléculaire des Microorganismes

Présentée par P^R ALATOU R.

1. Introduction

Les mécanismes favorisant la transcription à travers la chromatine décrits dans le premier chapitre ne sont pas les seuls à aider l'ARN pol II à progresser le long de l'ADN. En effet la polymérase peut s'arrêter temporairement et parfois reculer ce qui mène à un état appelé 'pause transcriptionnelle », si la pause transcriptionnelle persiste, elle évolue progressivement vers le statut **d'arrêt transcriptionnel** (Gu et Reines, 1995). Donc soit l'état de pause se résout de lui-même et l'ARN pol II étant capable de repartir, soit la polymérase est incapable de reprendre l'élongation et passe à un état d'arrêt transcriptionnel. Une ARN pol II arrêtée nécessite l'intervention de facteurs additionnels pour reprendre une élongation progressive ou pour être éliminée. L'arrêt de l'ARN pol II peut être provoqué par de multiples événements tels que des dommages de l'ADN, des séquences difficile à transcrire, une déplétions en nucléotides, un obstacle physique sur l'ADN, ou la structure de la chromatine (revue : Wilson *et al.*, 2013).

2. Facteurs impliqués dans la résolution de l'état transcriptionnel

L'arrêt de la transcription est caractérisé par une « marche arrière » de l'ARN pol II le long de l'ADN matrice, généralement de 2 à 14pb (Gu and Reines, 1995b). Durant cette marche arrière l'appariement ARN-ADN est maintenu mais l'extrémité 3' du transcrit sort du site actif de la polymérase. Les complexes d'élongation en pause transcriptionnelle et arrêtés sont ciblés par des facteurs de transcription, dont **TFIIS** et **TFIIF** qui agissent de concert pour résoudre la pause ou l'arrêt transcriptionnel et redémarrer l'élongation une fois que l'obstacle est éliminé ou surmonté (figure 01). TFIIS aide notamment l'ARN pol II à couper l'ARN lorsqu'elle a reculé afin de replacer l'extrémité 3' de l'ARN dans le centre catalytique de la polymérase (revue : Svejstrup, 2007).

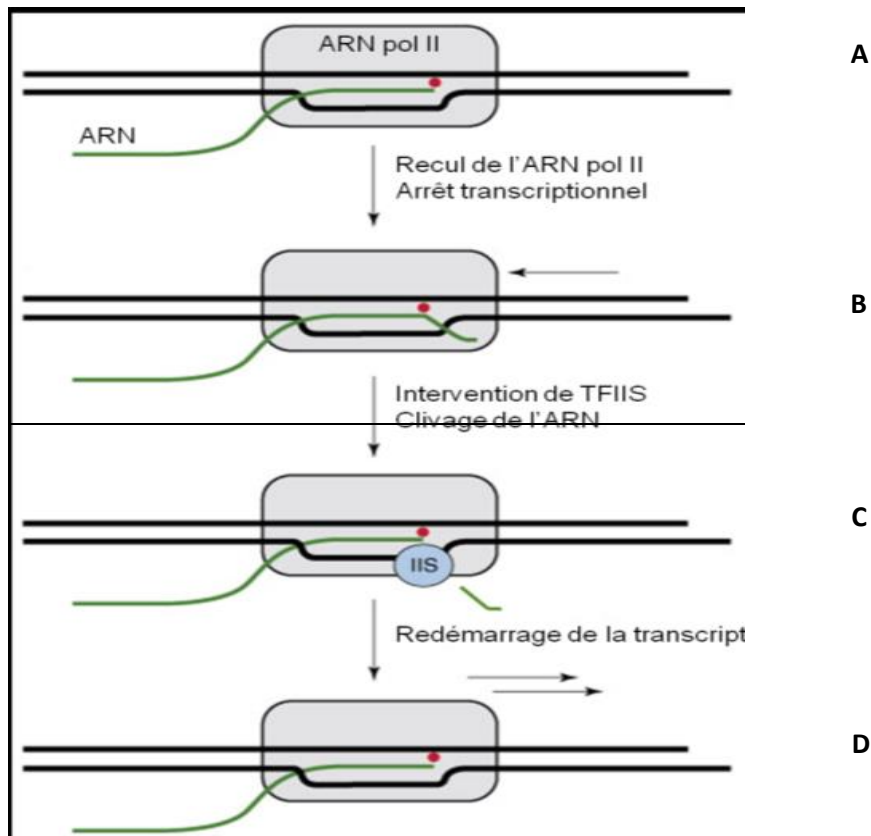


Figure 01 : Résolution de l'arrêt transcriptionnel par TFIIIS. Lors d'une pause de la transcription (A), l'ARN pol II peut reculer et passer dans un état d'arrêt transcriptionnel dans lequel l'extrémité 3' de l'ARN est déplacée du site catalytique (point rouge) de l'ARN pol II (B). TFIIIS (IIS) stimule le clivage endonucléolytique de l'ARN par l'ARN pol II, ce qui permet de repositionner l'extrémité 3' du transcript dans le site catalytique de l'ARN pol II (C). De cette manière, la transcription peut redémarrer permettant éventuellement à l'ARN pol II de surmonter l'obstacle initial (D). Figure adaptée de Svejstrup, 2007.

3. Rôle des facteurs THIIIS et TFIIF dans la résolution de l'arrêt transcriptionnel

➤ TFIIIS

TFIIIS est nommé Dst1 chez *S. cerevisiae*, est un facteur général d'élongation s'associant aux ARN pol II arrêtées connu pour favoriser le clivage endonucléolytique des ARN associés aux ARN pol II reculées, ce qui permet un redémarrage de la transcription (Izban and Luse, 1992a; Reines, 1992, figure 1). TFIIIS aide également l'ARN pol II à transcrire à travers les sites de pause sur l'ADN en favorisant la marche arrière des polymérase pausées (Schweikhard et al., 2014).

Dans le complexe d'élongation arrêté et reculé, 8 nucléotides de l'ARN sorti du site actif interagissent avec une région conservée chez les eucaryotes dans « l'entonnoir » de la polymérase provoquant des changements conformationnels de l'ARN pol II, stabilisant le

complexe polymérase-ADN-ARN et bloquant l'élongation. Pour résoudre ce blocage, TFIIIS entre dans l'entonnoir de la polymérase, rectifie les réarrangements conformationnels et stimule le clivage endonucléolytique du transcrit par l'ARN pol II permettant le réalignement de son extrémité 3' avec le site actif de la polymérase (Cheung and Cramer, 2011). Cette activité de clivage semble essentielle pour la survie cellulaire puisque l'utilisation d'un mutant de TFIIIS bloquant le clivage par l'ARN pol II est létal (Sigurdsson et al., 2010).

➤ **TFIIF**

TFIIF est un facteur général de la transcription associé au PIC (Pre-Initiation Complex). Il a également été montré que ce facteur est capable de se lier au complexe d'élongation.

Des études structurales menées plus récemment suggèrent que TFIIF se lie à la surface de l'ARN pol II, au niveau du tunnel d'entrée de l'ADN (Chen et al., 2010; Eichner et al., 2010), et stimule l'ouverture de la bulle de dénaturation de l'ADN (Gong et al., 2005; Zhang et al., 2005a). De cette manière, TFIIF pourrait aider les polymérases arrêtées à redémarrer.

Dans certain cas TFIIIS et TFIIF ne sont pas suffisant pour redémarrer la transcription une ARN pol II arrêtée, la polymérase est alors éliminée par un mécanisme de **poly-ubiquitination** menant à la dégradation de la grande sous-unité de l'ARN pol II

4. Ubiquitination et dégradation de l'ARN pol II

L'ubiquitination est définie comme étant une protéine de 76 acides aminés servant, elle-même de marqueur de protéines à éliminer. Elle est retrouvée au niveau de tous les compartiments subcellulaires dans toutes les cellules des organismes.

Lorsqu'une ARN pol II est bloquée par une lésion sur l'ADN elle peut avoir deux effets antagonistes sur la réparation de la lésion. D'une part elle favorise le recrutement de facteurs de réparation de l'ADN par le mécanisme du « transcription-coupled nucléotide excision repair » (TC-NER). Mais d'autre part, lorsque ce système de réparation n'est pas suffisant pour réparer la lésion, l'ARN pol II bloque l'accès à l'ADN des autres machineries de réparation requises. Pour libérer l'accès à l'ADN, les ARN pol II arrêtées sont ubiquitinées et désassemblées, ce qui favorise la réparation des lésions de l'ADN (Wilson *et al.*, 2013).

Il a été montré que l'ubiquitination et la dégradation de l'ARN pol II est une réponse générale aux complexes d'élongation arrêtés. Par exemple une délétion de TFIIIS entraîne une augmentation de l'ubiquitination de l'ARN pol II (Somesh *et al.*, 2005).

Université Frères Mentouri Constantine 1
Faculté Sciences de la Nature et de la Vie
Département de Microbiologie

Rôle et mécanismes de dégradation des ARNm chez *E. coli*
(Régulation post-transcriptionnelle)

Support pédagogique destiné aux étudiants Master 1, spécialité Biologie Moléculaire des Microorganismes

Réalisée par P^R ALATOU R.

Il existe plusieurs niveaux de régulation de l'expression génique qui permettent à la bactérie de s'adapter aux conditions environnementales. En effet, selon ces conditions, des variations de stabilité des ARNm ont été identifiées. Par exemple, il a été démontré que le stress induit par la phase stationnaire peut également influencer la régulation de la stabilité de certains messagers (Albertson & Nyström, 1994; Kuzj *et al.*, 1998) tout comme la croissance en anaérobie (Georgellis *et al.*, 1993). Par ailleurs, les mécanismes de dégradation des ARNm jouent un rôle incontournable dans la régulation post-transcriptionnelle car ils ont le pouvoir de répression d'un gène donnée.

Dans cette partie, les acteurs principalement les enzymes, les mécanismes et les régulations de la dégradation des ARNm chez *E. coli* seront abordés.

1. Dégradation des ARNm chez *E. coli*

1.1 Les principales ribonucléases

Selon leur activité les ribonucléases sont classées en deux familles : endoribonucléases et des exoribonucléases. Les endoribonucléases sont capables d'induire un clivage intramoléculaire au sein même des séquences d'ARN au niveau des liaisons phosphodiester. Les exonucléases sont des enzymes qui dégradent les ARN par leur extrémité. Chez *E. coli*, seules des exonucléases 3'-5' ont pu être identifiées (Zuo & Deutscher, 2001).

a. Les endoribonucléases

L'endoribonucléase la plus importante pour la dégradation des ARNm est la **RNase E**, codée par le gène essentiel *rne* (Carpousis *et al.*, 2009), elle clive les séquences d'ARN simple brin. C'est une protéine localisée à la membrane (Khemici *et al.*, 2008). La RNase E est une protéine de 1061 acides aminés possédant deux domaines fonctionnels distincts. En effet, l'extrémité N-terminale de la protéine (1 – 529) forme une structure globulaire qui correspond au siège de l'activité catalytique de la RNase alors que la partie C-terminale (530 – 1061) plutôt non structurée sert de support pour la formation d'un complexe protéique appelé '**dégradosome**' (Callaghan *et al.*, 2004, 2005). L'enzyme clive préférentiellement les ARNm dont l'extrémité 5' est monophosphorylée.

Parmi les endoribonucléases ayant également un rôle dans la dégradation des ARNm, on trouve la **RNase III**, codée par le gène *mec*, qui clive spécifiquement les doubles brins d'ARN (Nicholson, 1999).

Enfin, certains ARNm peuvent être spécifiquement clivés par l'endoribonucléase **RNase G**, codée par le gène *mng*. Cette enzyme non essentielle est considérée comme un paralogue de la RNase E en raison d'une très forte homologie avec son domaine catalytique globulaire (Okada *et al.*, 1994). RNase G est, elle aussi, plus sensible à la présence d'une extrémité 5' monophosphate.

b. Les exonucléases

Parmi les exoribonucléases identifiées chez *E. coli*, la **RNase II**, la **PNPase** et la **RNase R**, codées respectivement par les gènes *rnb*, *pnp* et *rnr* sont les principales exonucléases impliquées dans la dégradation des ARNm (Régnier & Arraiano, 2000; Cheng & Deutscher, 2005). Ces enzymes sont des exoribonucléases 3'-5' processives. Cela signifie qu'elles dégradent l'ARNm par son extrémité 3' vers son extrémité 5'. Alors que la RNase II et la RNase R libèrent des ribonucléotides monophosphates, la PNPase fournit des ribonucléotides diphosphates.

1.2 Les mécanismes de dégradation des ARNm

A partir des nombreuses études réalisées chez *E. coli* sur la dégradation des ARNm, il a été possible de définir un modèle global décrivant les mécanismes mis en jeu durant ce processus (Figure 01).

La dégradation des ARNm est initiée par un clivage endoribonucléolytique impliquant principalement la RNase E (Carpousis et al., 2009). Cette coupure peut également être réalisée par la RNase G ou par la RNase III. Cette opération entraîne la libération de produits de dégradation correspondant à des fragments d'ARNm. Ces fragments sont ensuite pris en charge par les exonucléases 3'-5' pour être réduits à l'état de monoribonucléotides (Carpousis, 2002). En 2011 Bouvier et Carpousis proposent un modèle de dégradation des ARNm par la RNase E. Ce modèle stipule qu'il peut exister deux voies d'entrées pour cette enzyme :

- soit le transcrit présente une extrémité 5' monophosphate favorisant un premier clivage de la RNase E de façon 5'-dépendante (par exemple suite à l'action RppH 'RNA 5Pyrophosphohydrolase').
- soit l'extrémité 5' n'est pas reconnue (à cause d'une triphosphorylation ou d'une structure secondaire piégeant l'extrémité 5') et la RNase E utilise la voie d'entrée « directe » en clivant le messager de façon indépendante de l'extrémité 5'.

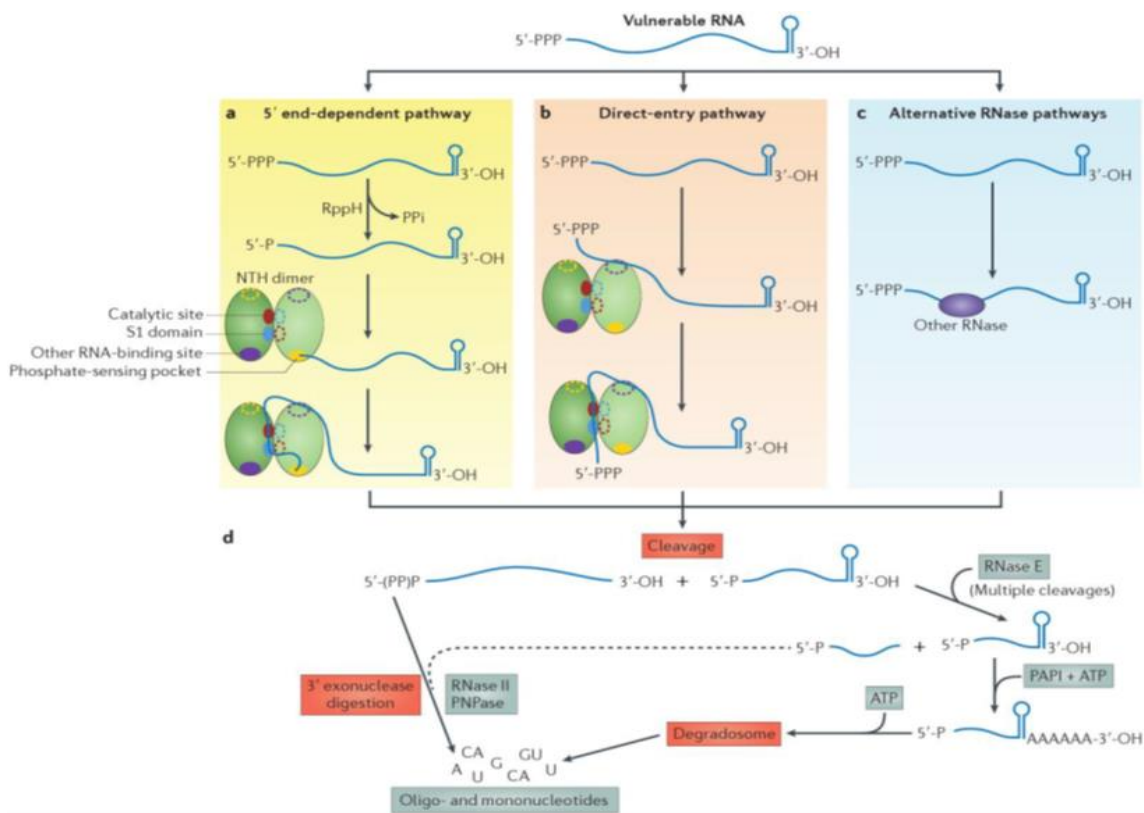


Figure 01 : Mécanismes généraux de dégradation des ARNm(Mackie, 2013)

- a) La « voie d'entrée 5'-dépendante » de laRNase E est liée à la détection de l'extrémité 5' monophosphate par la RNase E. b) « Entrée directe » de la RNase E indépendante de l'extrémité 5' de l'ARNm. c) Initiation de la dégradation par d'autres endoribonucléases (RNase G et RNase III). d) Les trois « voies » convergent vers les mêmes étapes tardives de dégradation des messagers faisant intervenir les exonucléases et un complexe protéique : le dégradosome. Seuls deux monomères de la RNase E sont représentés en vert. Le site catalytique est représenté en rouge, le domaine S1 qui forme un canal pour l'ARNm lors d'une reconnaissance 5'-dépendante du transcrite est représenté en bleu, un deuxième site d'interaction avec l'ARNm est représenté en violet et le site spécifique de reconnaissance de l'extrémité 5' monophosphate est représenté en jaune. Les sites sont colorés pour un monomère et représentés en pointillés pour l'autre.

1.3. Le dégradosome

La dégradation des ARNm chez *E. coli* montre que plusieurs partenaires protéiques interviennent. Bien qu'ils puissent agir de façon séparée et indépendante, il a été démontré que certaines de ces protéines interagissent et coopèrent pour former un complexe protéique appelé dégradosome.

Le dégradosome est constitué par l'association de quatre protéines (figure 02):

- L'RNase E qui est utilisée comme un support global grâce à son domaine C-terminal.
- La PNPase (polynucleotide phosphorylase) qui dispose d'une activité exoribonucléolytique.
- La Dead-Box RhlB qui ouvre la molécule d'ARN grâce à son action d'hélicase.

- L'élonase qui est une enzyme glycolytique. Elle pourrait jouer un rôle dans la stabilité du complexe (Py et al., 1996), car son action au niveau du dégradosome n'est pas connue. Donc l'interaction de l'endoribonucléase RNase E, de l'hélicase RhlB et de l'exonucléase PNPase au sein du dégradosome est supposée faciliter la dégradation des ARNm.

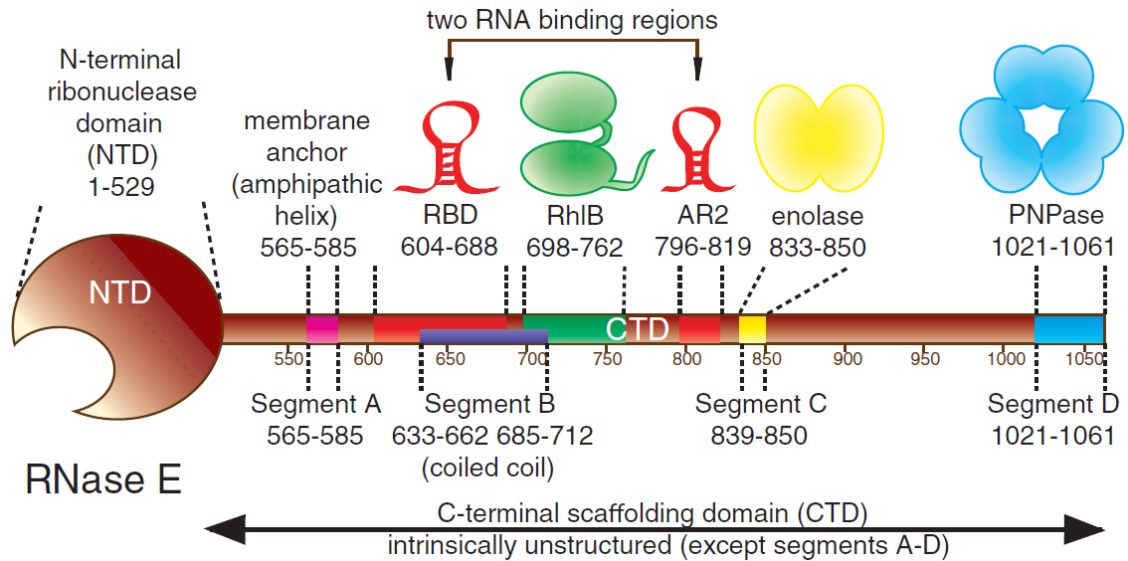


Figure 02 : Schéma du domaine C-terminal de la RNase E et des différents domaines d'interaction protéine-protéine permettant la formation du dégradosome. D'après (Górna et al., 2012).

Un seul monomère de RhlB se fixe au complexe alors que l'Enolase qui est un dimère se lie à la RNase E avec un ratio de 2:1. Le cas de la PNPase est plus compliqué car la protéine fonctionne en trimère et chaque monomère contient un site de liaison à la RNase E. En conséquence, un trimère de PNPase pourrait interagir avec trois RNase E simultanément.

Université Frères Mentouri Constantine 1
Faculté Sciences de la Nature et de la Vie
Département de Microbiologie

Le Silencing

Support pédagogique destiné aux étudiants Master 1, spécialité Biologie Moléculaire des Microorganismes

Présentée par P^R. R. ALATOU

Historique

Le phénomène d'ARN interférence a été découvert par Rochard Jorgensen des les années 1990 lors de ses recherches sur les mécanismes de coloration de *Petunias*. Les auteurs souhaitant obtenir un renforcement de la coloration des fleurs ont introduit une deuxième copie du gène de la pigmentation dans le génome de la plante, et ont obtenus contre toute attente, une extinction de la coloration (Napoli *et al.*, 1990). Ces expériences appelées de co-suppression ont mis en évidence un mécanisme très répandu chez les plantes : la protection du génome contre les éléments transposables mobiles et les virus à ARN (Vaucheret and Fagard, 2001). Ce mécanisme a été appelé « *Post Transcriptional Gene Silencing* » ou PTGS.

Plus tard Fire *et al* découvrent que l'introduction d'un ARN double brin (ARNdb) dans les cellules du nematode *Coenorhabditis elegans* pouvait réduire spécifiquement l'expression des protéines en se liant à leur ARN messenger (ARNm) et ceci pendant plusieurs générations de vers. Cela a permis de comprendre la fonction des nombreuses protéines (Fire *et al.*, 1998). Chez les animaux, ce phénomène a été appelé ARN interférence ou siRNA pour « short interference RNA ».

1. Le Silencing

L'ARN interférence (**ARNi**) est un terme qui se réfère à un silencing génique post – transcriptionnel induit soit par la dégradation soit par l'arrêt de la traduction de l'ARNm cible (voir (Ryther *et al.*, 2004).

Le silencing se réfère à un groupe de mécanismes par lequel l'expression d'un ou plusieurs gènes est régulée négativement ou entièrement supprimée par l'introduction d'un ARN antisens.

2. EXTINCTION POST-TRANSCRIPTIONNELLE

L'extinction post-transcriptionnelle de l'expression des gènes constitue le mécanisme principal du phénomène d'ARN interférence, et se déroule en deux étapes (figure 01) qui ont lieu au niveau du cytoplasme cellulaire.

En premier lieu, il y a reconnaissance d'un ARN double brin qui est alors dégradé en petits ARN double brin, et dans un second temps, ces derniers vont s'associer à un complexe protéique afin d'exercer leur action au niveau de l'ARN messenger. On trouve deux types de petits ARN double brin selon la nature de leur précurseur : les siRNA, qui vont entraîner la

dégradation de l'ARN messenger, et les miRNA, qui vont inhiber la traduction.

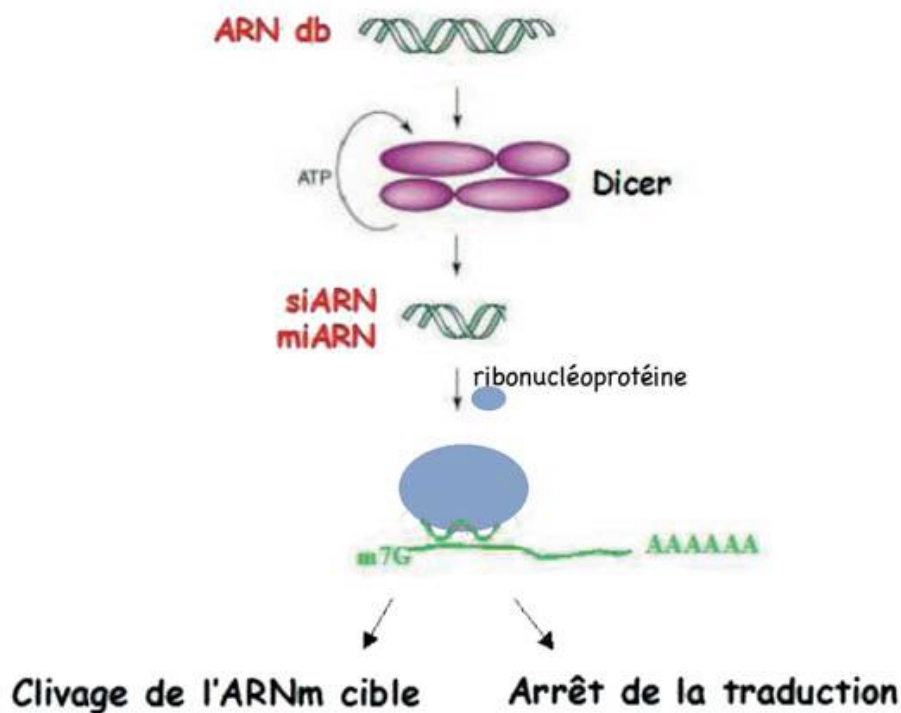


Figure 01 : Mécanisme général de l'extinction post-transcriptionnelle lors de l'ARN interférence.

A) Cas des siRNA

Les **siRNA** (short interfering RNAs) représentent des petites molécules d'ARN double brin de 21-22 nucléotides générées à partir d'un long double brin d'ARN.

Lorsqu'un ARN double brin est présent dans le cytoplasme des cellules eucaryotes, il est pris en charge par une enzyme à activité RNase III, qui le clive en petits ARN double brin, longs de 21 nucléotides, et présentant à chaque extrémité 3' deux nucléotides non appariés, tandis qu'ils sont phosphorylés en 5' (figure 02). Cette enzyme, majoritairement cytoplasmique et appelée **DICER**, a été identifiée en 2001 par Hammond chez *Drosophila melanogaster*. Il s'agit d'une ribonucléase reconnaissant spécifiquement tout ARN double brin, indépendamment de sa séquence, afin de produire de petits ARN de manière ATP dépendante. De telles endonucléases ont été identifiées dans tous les organismes compétents pour l'ARN interférence. Cette protéine est constituée d'un domaine ARN hélicase dans sa région N-terminale, d'un domaine, de deux domaines RNase III et d'un motif de liaison de l'ARN double brin dans sa région C-terminale

Des études structurales sur des RNases III bactériennes ont permis d'élaborer un modèle de

fonctionnement en dimère antiparallèle pour cette enzyme. Classiquement, ces enzymes possèdent quatre sites actifs. Mais dans le Dicer, un site sur deux seulement est actif, ce qui permet d'obtenir des ARN double brin de 22 nucléotides de long au lieu des petits ARN de 11 nucléotides produits habituellement par les RNases bactériennes.

Une fois formés, les siRNA s'associent à un complexe protéique effecteur, le RNA-Induced Silencing Complex (RISC). Il semblerait qu'il existe une protéine, R2D2, qui lie entre eux le Dicer et le complexe RISC, faisant ainsi le lien entre les deux étapes et accélérant le processus. Le complexe RISC a une composition protéique variable selon les espèces, mais il est toujours doté d'une activité hélicase ATP-dépendante et d'une activité RNase.

Les fonctions du RISC semblent nombreuses mais à ce jour, elles ne sont connues que dans les grandes lignes. Tout d'abord, ce complexe enzymatique déroule le petit ARN, de façon dépendante de l'ATP, pour dissocier les deux brins. La plus faible stabilité en 5' de l'antisens permettrait le déroulement sélectif et l'incorporation de ce brin dans le RISC. Le complexe activé ne conserve que le brin antisens, qui lui sert de guide pour reconnaître l'ARNm cible.

Le complexe RISC activé s'apparie à l'ARNm par l'intermédiaire des siRNA, qui ont une correspondance de séquence parfaite avec leur cible. Cela déclenche alors la dégradation de l'ARN messager. Lorsque l'ARN messager est clivé par le complexe RISC, les fragments générés, longs d'une vingtaine de nucléotides, sont ultérieurement dégradés par des RNases cellulaires, alors que le complexe endonucléolytique RISC est recyclé pour aller cliver d'autres ARNm.

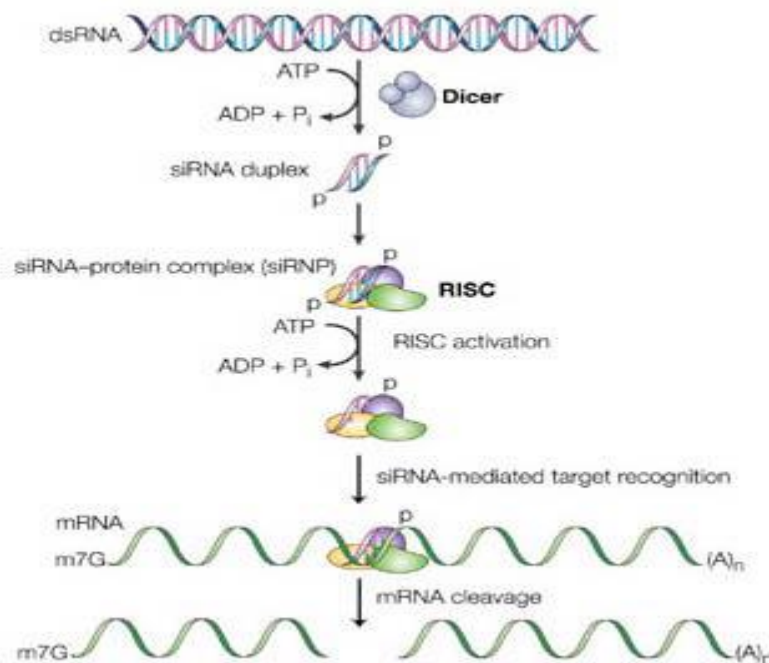


Figure 02 : Mécanisme d'action des siRNA. (d'après Dykxhoorn, 2003)

B) Cas des miRNA

Les miRNA sont des molécules endogènes de taille identique (19-25 nucléotides), que l'on a identifiées, entre autres, chez les nématodes, les drosophiles, les souris et les hommes. Ces petits ARN sont abondamment retrouvés dans les cellules (de quelques milliers à 40 000 par cellule et ils jouent un rôle important au cours du développement. Ils sont chargés de la régulation de l'expression des gènes, et un même miRNA peut réguler jusqu'à cent gènes différents.

Les miRNA dérivent du clivage successif de grands ARN en épingle à cheveux, encodés dans le génome et exprimés à partir d'un regroupement de gènes ou à partir d'une région unique. En premier lieu, on trouve les miRNA primaires (pri-miRNA), qui ont une complémentarité de séquence imparfaite (figure 03). A l'aide d'une enzyme appelée **Drosha**, ils sont clivés en plus petits ARN en épingle à cheveux de 70kb (short-hairpin RNA ; **shRNA**), que l'on désigne sous le nom de précurseurs de miRNA (pre-miRNA). Ces derniers sont exportés dans le cytoplasme et pris en charge par le Dicer, la même enzyme que dans le cas des siRNA. Le Dicer va couper les pre-miRNA en miRNA matures de 21-22 nucléotides, puis les dissocier en simple brin.

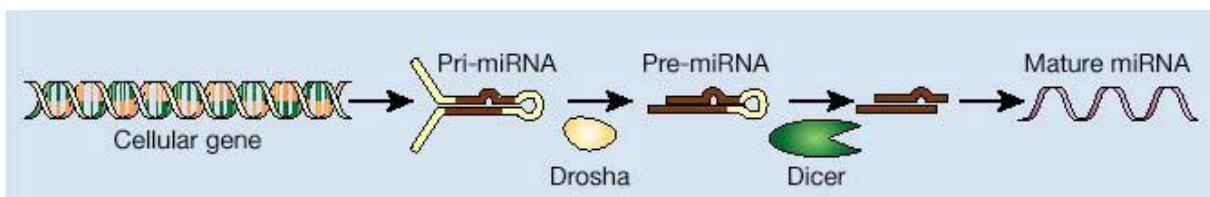


Figure 03 : Formation endogène des miRNA. (d'après Novina, 2004)

Ces miRNA simple brin s'associent en un complexe ribonucléoprotéique appelé miRNP (équivalent du RISC), et peuvent déclencher deux effets distincts selon leur complémentarité de séquence (figure 04). Les miRNA ayant une complémentarité de séquence partielle se lient à la région 3' non traduite des ARNm, ce qui a pour conséquence l'inhibition de la traduction.

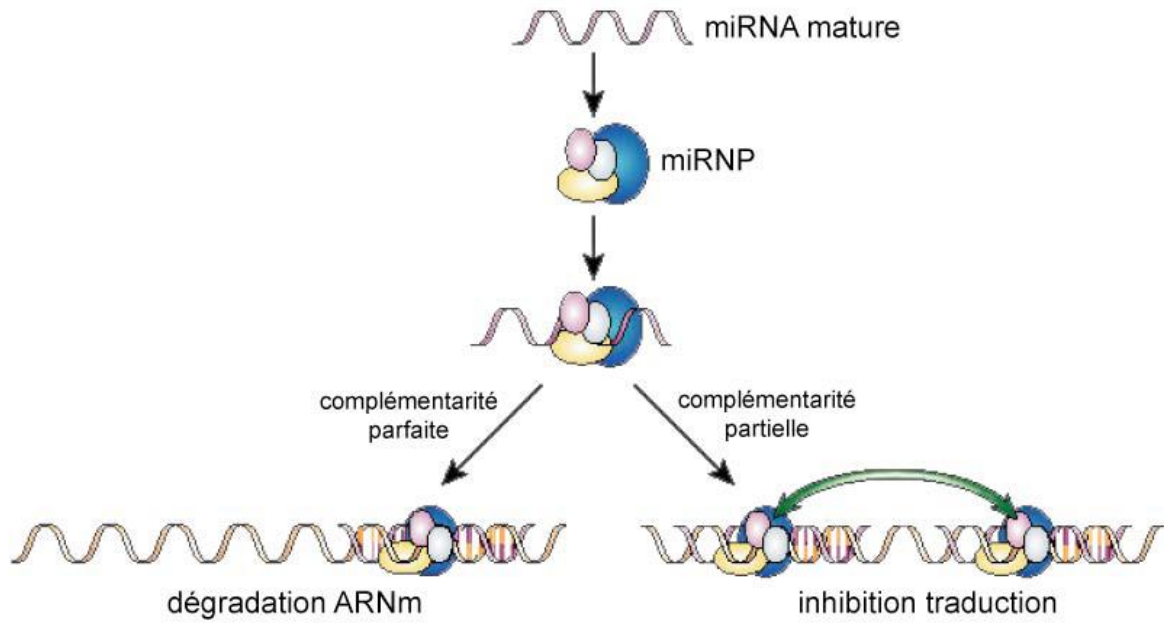


Figure 04 : Mécanisme d'action des miRNA