

# CHAPITRE 16

## Contrôle de l'expression génique

### Aperçu du chapitre

- 16.1 Contrôle de l'expression génique
- 16.2 Les protéines de régulation
- 16.3 La régulation procaryote
- 16.4 La régulation eucaryote
- 16.5 Structure de la chromatine et expression génique
- 16.6 Contrôle post-transcriptionnel des eucaryotes
- 16.7 Dégradation des protéines



### Introduction

Dans une symphonie, les différents instruments jouent à des moments différents ; la partition détermine quels instruments jouent et à quel moment. Dans un organisme aussi, les différents gènes s'expriment à des moments différents, quand une « partition génétique » inscrite dans les régions régulatrices de l'ADN décide quels gènes sont actifs et quand. La figure montre un « puff » formé sur ce chromosome de drosophile : il représente des gènes qui s'expriment activement. L'expression génique et son contrôle sont le sujet abordé dans ce chapitre.

### 16.1 Contrôle de l'expression génique

#### Objectifs

1. Identifier l'endroit où s'exerce généralement le contrôle de l'expression génique.
2. Décrire le rôle habituel des protéines de régulation.
3. Donner une liste des différences entre le contrôle de l'expression génique chez les procaryotes et chez les eucaryotes.

Le contrôle de l'expression génique est essentiel pour tous les organismes. Chez les procaryotes, il permet à la cellule de tirer parti des

modifications de son environnement. Chez les eucaryotes pluricellulaires, il est essentiel pour diriger le développement et maintenir l'homéostasie. En fait, les différents types de cellules d'un organisme pluricellulaire ont des fonctions différentes selon les protéines exprimées. Toutes ces cellules possèdent un lot de gènes « d'entretien » nécessaires à leur fonctionnement, mais elles expriment en outre un ensemble spécifique de protéines caractéristique de chaque type cellulaire.

#### Le contrôle s'effectue généralement au niveau de l'initiation de la transcription

Au chapitre 15, nous avons vu que l'expression génique constitue la transformation du génotype en phénotype – c'est le flux d'information venant de l'ADN et produisant les protéines fonctionnelles qui contrôlent les activités cellulaires. Cet aspect traditionnel de l'expression génique

implique un contrôle du processus principalement au niveau de l'initiation de la transcription. Cette perspective reste valide, mais de plus en plus d'indices montrent que la portion du génome qui est transcrite et le contrôle de cette transcription sont plus complexes qu'on ne l'avait supposé chez les organismes pluricellulaires.

Dans ce chapitre, nous allons développer le contrôle de l'initiation de la transcription de façon assez détaillée en raison de son importance et parce que c'est encore le mécanisme de contrôle de l'expression génique le mieux étudié. Sur cette base, nous verrons aussi comment la structure de la chromatine influence l'expression génique et comment un contrôle peut également s'exercer après la transcription. Ce dernier sujet nous conduira à un nouveau monde étonnant, celui des molécules d'ARN de régulation.

L'ARN polymérase est la clé de la transcription ; elle doit avoir accès à l'hélice d'ADN et doit pouvoir s'unir au promoteur pour que la transcription débute. Le rôle des protéines de régulation consiste à moduler la faculté de l'ARN polymérase à s'unir au promoteur. Ce contrôle de l'accès de l'ARN polymérase à un promoteur est fréquent chez les procaryotes et les eucaryotes, mais nous allons voir qu'il diffère beaucoup dans ses détails.

Ces protéines régulatrices s'unissent à des séquences nucléotidiques spécifiques de l'ADN longues de 10 à 15 nucléotides seulement (même une longue protéine de régulation possède une « empreinte », ou zone d'union, réduite à une vingtaine de nucléotides). On a caractérisé des centaines de ces séquences de régulation, et chacune offre un site d'union à une protéine capable de reconnaître la séquence. Soit l'union de la protéine à la séquence régulatrice *bloque* la transcription en se plaçant sur le trajet de l'ARN polymérase, soit elle *stimule* la transcription en facilitant l'union de l'ARN polymérase au promoteur.

## Chez les procaryotes, les stratégies de contrôle sont mises en route pour s'adapter aux modifications de l'environnement

Le contrôle de l'expression génique est très différent chez les procaryotes et chez les eucaryotes. Les cellules procaryotes ont été modelées par l'évolution de manière à croître et se diviser le plus rapidement possible et leur permettre d'exploiter des ressources transitoires. La rotation des protéines est rapide chez les procaryotes, permettant à ces organismes de réagir rapidement aux changements de leur environnement externe en modifiant le mode d'expression des gènes.

Chez les procaryotes, la principale fonction du contrôle génique est l'adaptation des activités de la cellule à son environnement immédiat. Les modifications de l'expression génique concernent les enzymes présentes dans la cellule pour répondre à la quantité et au type de nutriments disponibles et à la quantité d'oxygène. Presque toutes ces modifications sont parfaitement réversibles, ce qui permet à la cellule d'adapter la quantité de ses enzymes dans les deux sens, en fonction des modifications de l'environnement.

## Chez les eucaryotes, les stratégies de contrôle sont destinées à maintenir l'homéostasie et à diriger le développement

Par contre, les cellules des organismes pluricellulaires ont été modelées par l'évolution pour les protéger des modifications temporaires de leur

environnement immédiat. La plupart sont soumises à des conditions assez constantes. En fait, l'*homéostasie* – conservation d'un environnement interne constant – est considérée par beaucoup comme la caractéristique des organismes pluricellulaires. Dans ces organismes, les cellules répondent à des signaux provenant de leur environnement immédiat (comme les facteurs de croissance et les hormones) en modifiant l'expression de leurs gènes, mais elles interviennent ainsi dans la régulation de l'ensemble de l'organisme.

C'est peut-être pour le développement de l'organisme lui-même que le contrôle de l'expression génique est le plus important. C'est le résultat de modifications coordonnées de l'expression des gènes au cours du développement, ainsi que dans les différents tissus. La compréhension de ce problème complexe a été un objectif essentiel de la génétique du développement, et les recherches ont mis en lumière des circuits d'expression génique qui ont persisté pendant de longues périodes évolutives (voir chapitre 19). Un même circuit de régulation peut aussi intervenir au sein d'un organisme pour modeler des structures différentes à des endroits différents et à des stades différents du développement. Par exemple, le gène *sonic hedgehog* intervient dans la différenciation du tube neural et, ultérieurement, dans celle des doigts de la main en développement.

Les eucaryotes unicellulaires disposent aussi de mécanismes de contrôle différents de ceux des procaryotes. Tous les eucaryotes possèdent un noyau entouré d'une membrane, utilisent les mêmes mécanismes de condensation de l'ADN en chromosomes et disposent du même système d'expression génique : toutes ces caractéristiques les distinguent des procaryotes.

### Questions d'apprentissage 16.1

L'expression génique est généralement contrôlée au niveau de l'initiation de la transcription. Les protéines de régulation s'unissent à des séquences d'ADN spécifiques et interviennent dans la fixation de l'ARN polymérase aux promoteurs. Une protéine particulière peut empêcher ou stimuler la transcription. Chez les procaryotes, la régulation se concentre sur l'adaptation des activités cellulaires à l'environnement pour garantir leur viabilité. Chez les eucaryotes pluricellulaires, la régulation est destinée à conserver l'homéostasie interne et, même dans les formes unicellulaires, ce contrôle comporte des mécanismes qui tiennent compte du noyau et des chromosomes multiples.

- Vous attendriez-vous à ce que le contrôle de l'expression génique d'un eucaryote unicellulaire comme la levure soit plus proche de celui des humains ou d'*E. coli* ?

## 16.2 Les protéines de régulation

### Objectifs

1. Expliquer comment des protéines peuvent interagir avec des paires de bases sans déroulement de l'hélice.
2. Décrire les caractéristiques communes aux motifs de liaison à l'ADN.

La possibilité pour certaines protéines de régulation de s'unir à des séquences *spécifiques* d'ADN et soit de bloquer la transcription, soit de la faciliter en s'unissant à l'ARN polymérase est à la base du contrôle de l'expression des gènes. Pour comprendre comment les cellules contrôlent l'expression génique, il faut d'abord avoir une vision claire de ce processus de reconnaissance moléculaire.

## Les protéines peuvent interagir avec l'ADN par le sillon majeur

Il n'est en fait pas nécessaire que l'hélice d'ADN soit déroulée pour que les protéines puissent identifier les séquences d'ADN. Quand une protéine se fixe à la surface de l'ADN, les bases sont exposées dans les deux sillons de la molécule. En raison de sa nature, l'hélice forme deux sillons dans la structure cylindrique de l'hélice : un majeur et un mineur (voir figure 14.8). Dans le sillon le plus profond, le sillon majeur, les donneurs et accepteurs de liaisons hydrogène des nucléotides sont accessibles. Ces groupements chimiques ont une forme spécifique pour chacune des quatre paires de bases possibles et représentent, pour une protéine logée dans le sillon, un moyen de lire la séquence de bases (figure 16.1).

## Les domaines de liaison à l'ADN interagissent avec des séquences d'ADN spécifiques

La reconnaissance entre protéines et ADN est essentielle pour toutes les fonctions de l'ADN, comme la réplication, l'expression des gènes et le contrôle de ces activités. Les protéines capables d'interagir avec l'ADN sont assez nombreuses, mais la détermination de la structure de certaines d'entre elles a montré que quelques motifs protéiques seulement se fixent effectivement à l'ADN. Ces motifs sont probablement apparus relativement tôt au cours de l'évolution et ils ont été réutilisés, après modification, dans des protéines très diverses. Ces motifs de liaison à l'ADN ont la propriété d'interagir avec des séquences de bases spécifiques, généralement par le sillon majeur de l'hélice d'ADN.

Les motifs de liaison à l'ADN constituent la structure clé du domaine de liaison à l'ADN de ces protéines. Par sa fonction, ce domaine est une partie distincte de la protéine nécessaire à la liaison à l'ADN par interaction entre séquences spécifiques. Les protéines de régulation doivent aussi interagir avec l'appareillage de transcription, grâce à un autre domaine de régulation.

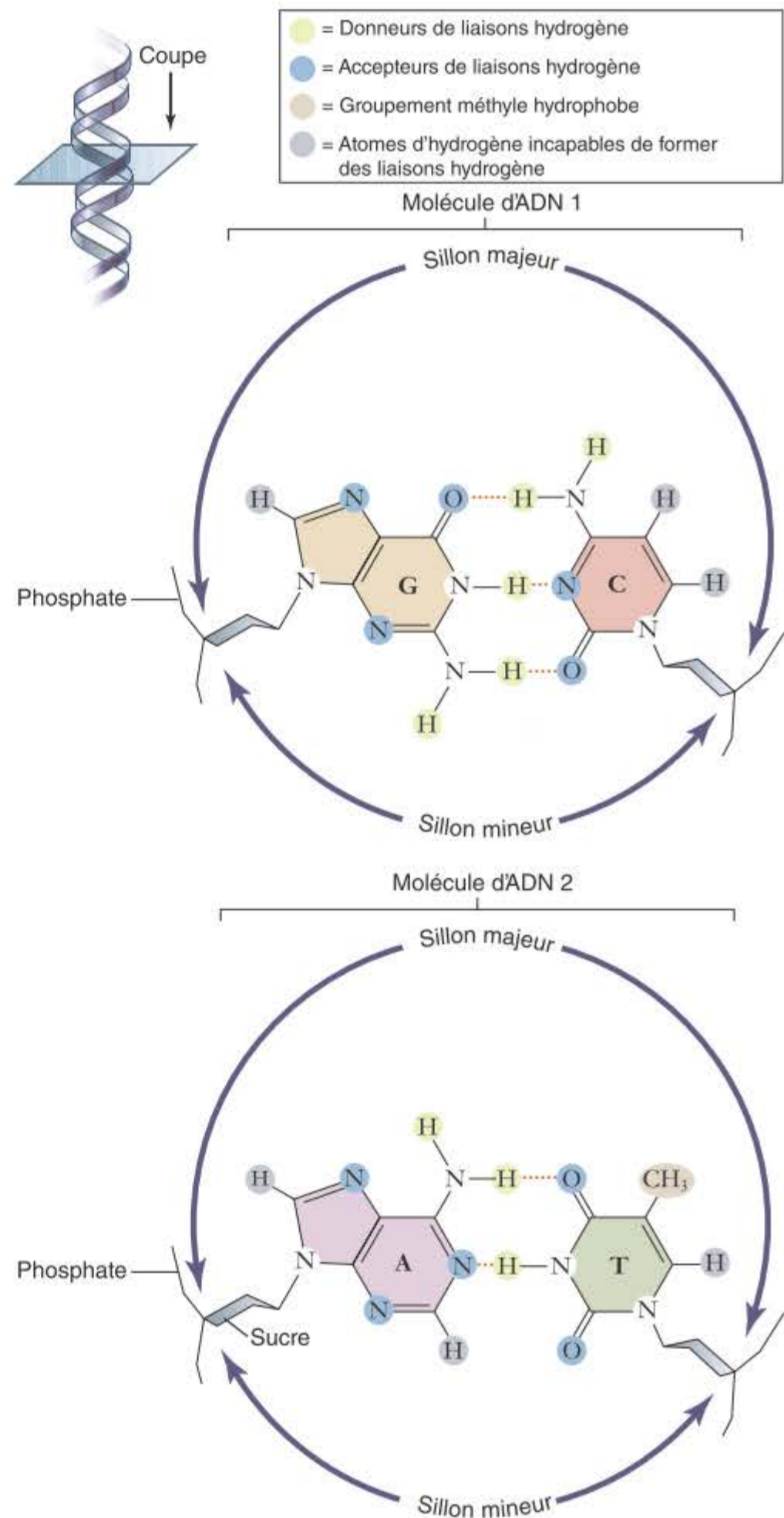
Notez que deux protéines possédant le même domaine de liaison à l'ADN ne s'unissent pas nécessairement à la même séquence d'ADN. Les motifs de liaison à l'ADN se ressemblent par leur structure tridimensionnelle, pas par les contacts spécifiques qu'ils ont avec l'ADN.

## Plusieurs motifs communs de liaison à l'ADN se retrouvent dans beaucoup de protéines

On trouve un nombre limité de motifs communs de liaison à l'ADN dans beaucoup de protéines différentes. Quatre, parmi les mieux connus, sont décrits dans les sections suivantes afin de démontrer comment les protéines de liaison interagissent avec l'ADN.

### Le motif hélice-coude-hélice

Le motif de liaison à l'ADN le plus commun est l'**hélice-coude-hélice**, formé de deux hélices  $\alpha$  de la protéine reliées par un court segment non hélicoïdal, le « coude » (figure 16.2a). Premier motif de liaison à l'ADN



**Figure 16.1** Lecture du sillon majeur de l'ADN. En regardant vers le bas dans le sillon majeur de l'ADN, on peut voir les bords des bases qui émergent dans le sillon. Chacune des quatre paires de bases possibles (deux sont visibles ici) propose un ensemble particulier de groupements chimiques à l'intérieur du sillon, ce que montrent les billes de couleurs différentes dans ce schéma. Une protéine de régulation peut identifier la disposition des paires de bases grâce à cette signature caractéristique.

connu, le motif hélice-coude-hélice a, depuis lors, été identifié dans des centaines de protéines se liant à l'ADN.

Observée de près, la structure d'un motif hélice-coude-hélice montre comment les protéines possédant ce motif sont capables d'interagir avec le sillon majeur de l'ADN. Les segments hélicoïdaux du motif interagissent entre eux pour former un angle à peu près droit. Quand ce motif est appliqué contre l'ADN, un des segments hélicoïdaux (l'*hélice de reconnaissance*) s'adapte parfaitement au sillon majeur, tandis que l'autre s'appuie à la surface externe de la molécule d'ADN, participant à la mise en place correcte de l'hélice de reconnaissance.

La plupart des séquences régulatrices de l'ADN reconnues par des motifs hélice-coude-hélice forment des paires symétriques. Ces séquences s'unissent à des protéines possédant deux motifs hélice-coude-hélice distants de 3,4 nanomètres (nm), distance correspondant à une spire de l'hélice d'ADN (figure 16.2a). La présence de *deux* sites de liaison protéine-ADN multiplie par deux la surface de contact entre la protéine et l'ADN et renforce nettement la liaison qui se forme entre eux.

### Le motif homéodomaine

Une catégorie particulière de motifs hélice-coude-hélice, l'homéodomaine, joue un rôle critique dans le développement d'organismes eucaryotes très divers, y compris les humains. Chez la drosophile, un groupe de mutations, dénommées homéotiques, entraînent la substitution d'une partie du corps par une autre. L'analyse des gènes affectés a montré une séquence de 60 acides aminés formant l'homéodomaine. La portion de l'homéodomaine la plus conservée contient une hélice de reconnaissance d'un motif hélice-coude-hélice. C'était la première indication de l'ancienneté du mécanisme de développement, et il se retrouvait à des distances phylogénétiques plus grandes qu'on ne le pensait. Pour ce qui concerne nos gènes, nous sommes beaucoup plus proches d'une mouche que vous ne pourriez l'imaginer.

### Le motif à doigt de zinc

Un type différent de motif de liaison à l'ADN fait appel à un ou plusieurs atomes de zinc pour coordonner sa liaison à l'ADN. Appelés **motifs à doigt de zinc**, ces motifs prennent plusieurs formes. Dans l'une d'elles, un atome de zinc unit un segment en forme d'hélice  $\alpha$  à un feuillet  $\beta$  (voir chapitre 3) en sorte que le segment hélicoïdal s'insère dans le sillon majeur de l'ADN.

Ce type de motif forme souvent des groupes, les feuilletts  $\beta$  séparant les segments hélicoïdaux, de telle sorte que chaque hélice soit au contact du sillon majeur. C'est un peu comme une main agrippant l'ADN, les doigts se plaçant dans le sillon majeur. Plus il y a de motifs à doigt de zinc dans le groupe, plus forte est la liaison de la protéine à l'ADN.

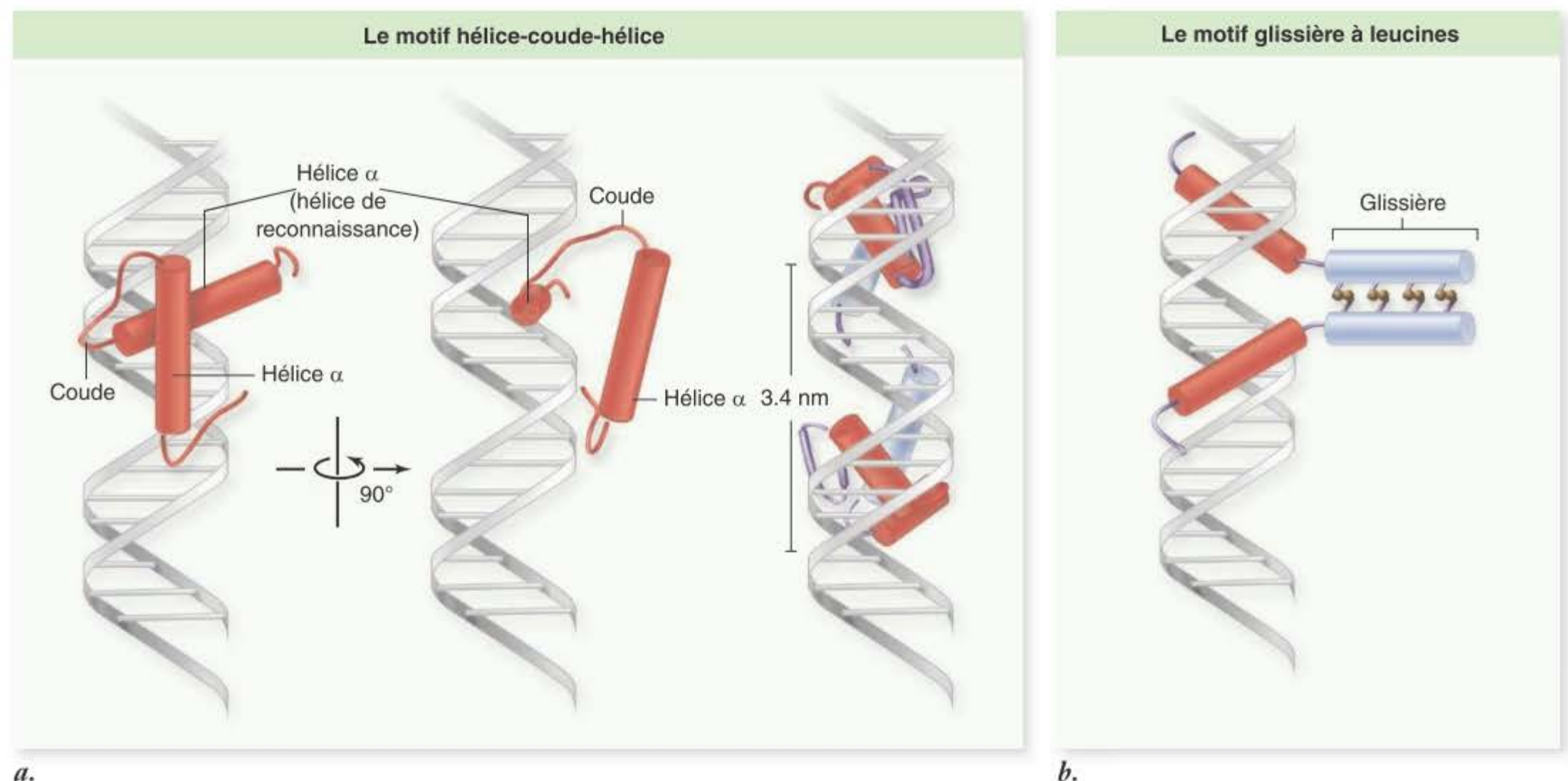
### Le motif glissière à leucines, ou « leucine zipper »

Dans cet autre motif d'union à l'ADN, le nom se réfère à la dimérisation d'un motif permettant l'association à l'ADN de différentes sous-unités d'une protéine. Cette **glissière à leucines** provient de l'interaction d'une région d'une sous-unité contenant plusieurs acides aminés hydrophobes (généralement des leucines) avec la même région d'une autre sous-unité. Cette interaction réunit les deux sous-unités et forme une structure en Y dont les deux bras sont des régions hélicoïdales s'adaptant au sillon majeur de l'ADN, mais des deux côtés de l'hélice (figure 16.2b) et tiennent l'ADN comme une pince.

### Questions d'apprentissage 16.2

L'hélice d'ADN possède un sillon majeur et un mineur ; les protéines de régulation interagissent avec l'ADN en accédant aux bases le long du sillon majeur. Toutes ces protéines possèdent des motifs de liaison à l'ADN, souvent avec un ou deux segments en hélice  $\alpha$ . Ces motifs forment la partie active du domaine de liaison à l'ADN, et un autre domaine de la protéine interagit avec l'appareil de transcription.

- Quelles seraient les conséquences d'une mutation de la protéine hélice-coude-hélice qui modifierait l'espace entre les deux hélices ?



**Figure 16.2** Les principaux motifs d'union à l'ADN. Deux motifs différents d'union à l'ADN sont illustrés avec leur interaction avec l'ADN. *a.* Le motif hélice-coude-hélice s'unit à l'ADN grâce à une hélice  $\alpha$ , l'hélice de reconnaissance, qui interagit avec le sillon majeur. L'autre hélice détermine la position de l'hélice de reconnaissance. Les protéines possédant ce motif sont généralement des dimères avec deux sous-unités identiques, contenant chacune le motif d'union à l'ADN. Les deux exemplaires du motif (en rouge) sont distants de 3,4 nm, exactement l'écart entre deux spires de l'hélice d'ADN. *b.* La glissière à leucines réunit deux sous-unités d'une protéine complexe et permet ainsi l'interaction entre les régions en forme d'hélice  $\alpha$  et l'ADN.

## 16.3 La régulation procaryote

### Objectifs

1. Montrer les différences entre contrôle par induction et contrôle par répression.
2. Expliquer le contrôle de l'expression génique dans l'opéron *lac*.
3. Expliquer le contrôle de l'expression génique par l'opéron *trp*.

On peut se représenter la régulation avec précision en examinant les mécanismes utilisés par les procaryotes pour contrôler l'initiation de la transcription. Certains aspects sont semblables chez les procaryotes et les eucaryotes, mais il existe aussi de profondes différences. Dans les sections 16.4 et 16.5, nous envisagerons les systèmes eucaryotes et nous verrons leurs différences par rapport aux systèmes procaryotes plus simples.

### Le contrôle de la transcription peut être positif ou négatif

Au niveau de l'initiation de la transcription, le contrôle peut être positif ou négatif. Un **contrôle positif** augmente la fréquence de l'initiation et un **contrôle négatif** la diminue. Ces deux formes de contrôle sont effectuées par des protéines de régulation dont les effets sont cependant opposés.

#### Contrôle négatif par répresseurs

Les **répresseurs** sont des protéines responsables du contrôle négatif. Ces protéines s'unissent à des sites régulateurs de l'ADN désignés comme **opérateurs**, ils empêchent ou réduisent l'initiation de la transcription. Ils fonctionnent comme des barrages routiers et empêchent l'initiation effective par la polymérase.

Les répresseurs n'agissent pas isolément ; tous réagissent à des molécules effectrices spécifiques. L'union d'un effecteur peut modifier la conformation du répresseur et améliorer ou empêcher son union à l'ADN. Ces répresseurs sont des protéines allostériques avec un site actif qui se lie à l'ADN et un site régulateur s'unissant aux effecteurs. La liaison à un effecteur modifie la faculté du répresseur à s'unir à l'ADN (voir les détails concernant les protéines allostériques au chapitre 6).

#### Contrôle positif par activateurs

Une autre classe de protéines régulatrices allostériques intervient dans le contrôle positif : ce sont des **activateurs**, capables de s'unir à l'ADN et de stimuler l'initiation de la transcription. Ces activateurs améliorent la liaison de l'ARN polymérase au promoteur et accroissent le niveau de l'initiation de la transcription.

Logiquement et physiquement, les activateurs et les répresseurs sont opposés. Des molécules d'effecteurs peuvent accroître ou réduire la liaison des activateurs.

### Les procaryotes adaptent l'expression génique en réaction aux conditions environnementales

Les modifications de l'environnement des bactéries et archées entraînent souvent des modifications de l'expression génique. En général, les gènes qui codent des protéines intervenant dans des voies cataboliques (dégra-

dition de molécules) ont une réponse opposée à celle des gènes intervenant dans des voies anaboliques (synthèse de molécules). Dans l'exposé qui suit, nous décrivons les enzymes de la voie catabolique transportant et utilisant un sucre, le lactose. Nous décrivons ensuite la voie anabolique de synthèse d'un acide aminé, le tryptophane.

Comme on l'a vu au chapitre 15, les gènes procaryotes sont souvent organisés en opérons, groupes de gènes faisant partie d'une même unité de transcription avec un seul promoteur. Les gènes intervenant dans une même voie métabolique sont souvent organisés ainsi. Les protéines nécessaires à l'utilisation du lactose sont codées par l'**opéron *lac*** et celles qui interviennent dans la synthèse du tryptophane sont codées par l'**opéron *trp***.



**Question** Quel est l'avantage, pour une bactérie, de réunir dans un même opéron plusieurs gènes dont tous les produits interviennent dans une même voie biochimique ?

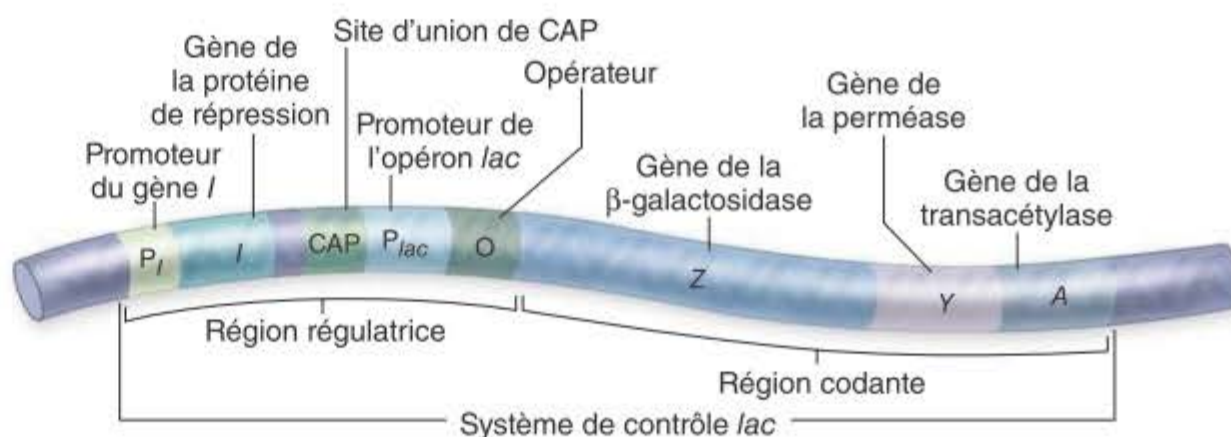
### Induction et répression

Si une bactérie trouve du lactose, elle commence à synthétiser les enzymes nécessaires à son utilisation. Quand il n'y a plus de lactose, cependant, la synthèse de ces protéines n'est plus nécessaire. On dira que la synthèse des protéines est **induite** par la présence du lactose. Il y a donc **induction** quand les enzymes d'une voie sont synthétisées en réponse à un substrat.

Quand il y a du tryptophane dans l'environnement, une bactérie ne produira pas les enzymes nécessaires à sa synthèse. Si le tryptophane n'est plus disponible, la bactérie commence à produire ces enzymes. On a une **répression** quand les bactéries capables de produire des enzymes de biosynthèse ne les synthétisent pas. En cas d'induction et de répression, la bactérie s'adapte et synthétise les enzymes les mieux adaptées à son environnement immédiat.

### Contrôle négatif

Le contrôle de l'initiation de la transcription peut être positif ou négatif. En apparence, la répression peut sembler négative et l'induction positive, mais ce n'est pas le cas. En fait, la régulation des deux opérons, *lac* et *trp*, est contrôlée négativement par des protéines de régulation. Ce fonctionnement en apparence opposé est possible parce que, dans les deux cas, les protéines de répression interagissent avec l'ADN et avec leurs effecteurs respectifs dans des voies opposées.



**Figure 16.3** La région *lac* du chromosome d'*Escherichia coli*.

L'opéron *lac* se compose d'un promoteur, d'un opérateur et de trois gènes (*lac Z*, *Y* et *A*) codant les protéines nécessaires au métabolisme du lactose. Il existe en outre un site de liaison pour la protéine activatrice du catabolite (CAP), qui intervient dans l'union de l'ARN polymérase au promoteur. Le gène *I* code une protéine de répression capable de s'unir à l'opérateur et de bloquer la transcription de l'opéron *lac*.

Pour que les deux mécanismes fonctionnent, la molécule externe, comme le lactose ou le tryptophane, doit agir adéquatement sur le gène à réguler. Pour l'induction de *lac*, la présence du lactose doit empêcher la liaison du répresseur à sa séquence régulatrice. Pour la répression de *trp* au contraire, la présence du tryptophane doit entraîner la liaison d'un répresseur à sa séquence régulatrice.

Ces réponses sont opposées parce que les besoins de la cellule sont opposés dans les voies anaboliques et cataboliques. Chaque voie est décrite en détail afin de montrer comment les interactions entre protéines et ADN permettent à la cellule de répondre aux conditions environnementales.

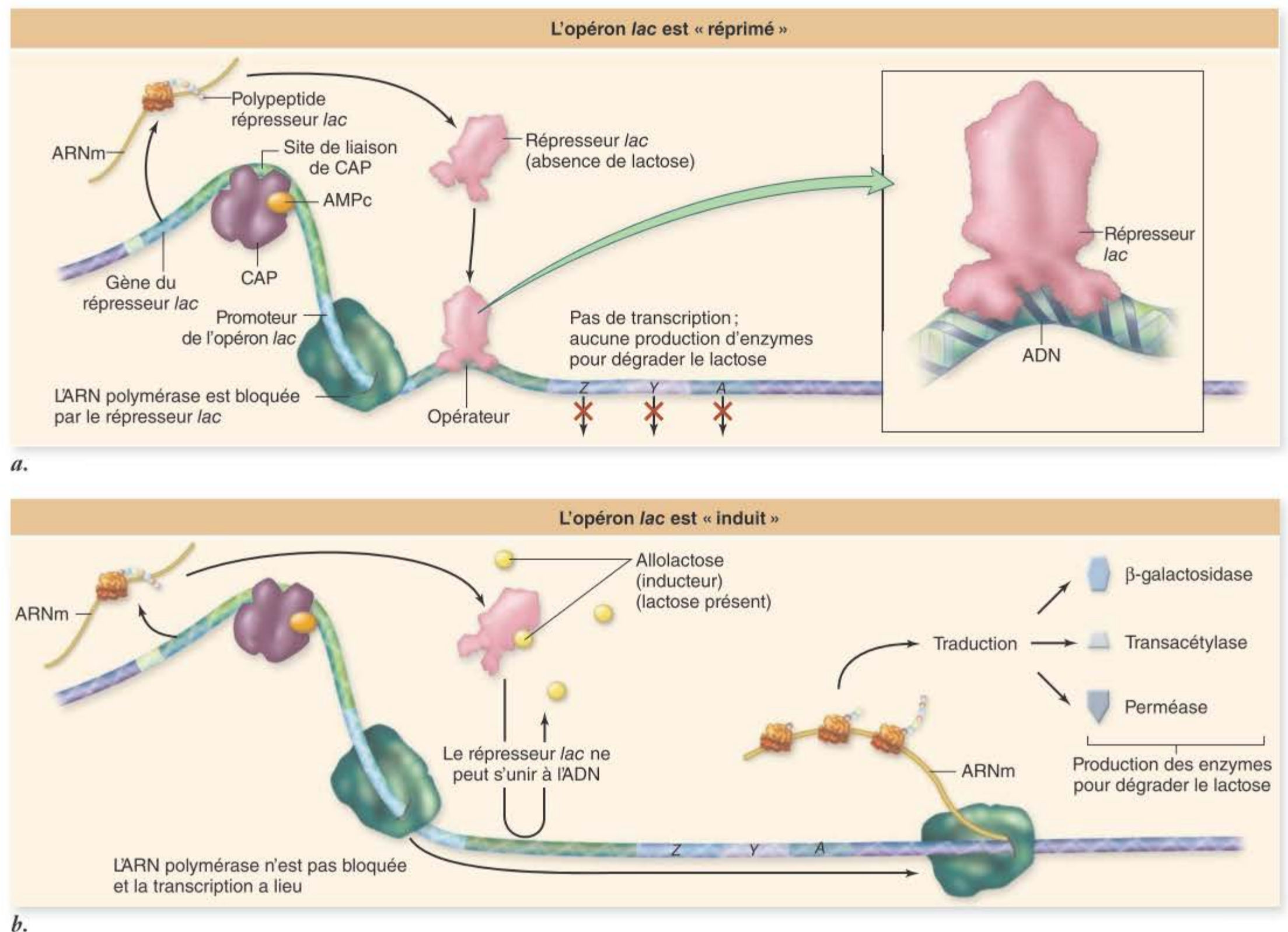
## L'opéron *lac* est contrôlé négativement par le répresseur *lac*

Le contrôle de l'expression génique dans l'opéron *lac* a été élucidé grâce aux travaux pionniers de Jacques Monod et François Jacob.

L'opéron *lac* comprend les gènes codant les fonctions nécessaires à l'utilisation du lactose : la  $\beta$ -galactosidase (*lacZ*), la perméase (*lacY*) et la lactose transacétylase (*lacA*), ainsi que les séquences régulatrices nécessaires au contrôle de l'expression de ces gènes (figure 16.3). En outre, le gène du répresseur *lac* (*lacI*) est uni au reste de l'opéron *lac* et l'on considère qu'il en fait partie, bien qu'il possède son propre promoteur. La disposition des régions de contrôle en amont de la région codante est typique de la plupart des opérons procaryotes, bien que le répresseur ne s'y trouve pas.

### Fonctionnement du répresseur

L'initiation de la transcription de l'opéron *lac* est contrôlée par le répresseur *lac*. Ce répresseur s'unit à l'opérateur, contigu au promoteur (figure 16.4a). Cette union empêche la fixation de l'ARN polymérase au promoteur. Cette liaison à l'ADN est sensible à la présence de lactose. Le répresseur s'unit à l'ADN en l'absence de lactose, mais pas en sa présence.



**Figure 16.4 Induction de l'opéron *lac*.** *a.* En l'absence de lactose, le répresseur *lac* s'unit à l'ADN au niveau du site opératoire et empêche ainsi la transcription de l'opéron. Quand la protéine du répresseur est unie au site opératoire, l'opéron *lac* est réprimé. *b.* L'opéron *lac* est transcrit (induit) quand la CAP est fixée et quand le répresseur ne l'est pas. L'union de l'allolactose au répresseur modifie la forme de ce dernier, qui ne peut plus s'unir à l'opérateur et bloquer l'activité de la polymérase.



**Analyse des données** Quel serait le phénotype d'une mutation du répresseur empêchant l'union à l'ADN ? L'union de l'inducteur ? Quelle serait la différence avec une mutation de l'opérateur empêchant l'union du répresseur ?

### Interaction entre répresseur et inducteur

En l'absence de lactose, le répresseur *lac* s'unit à l'opérateur et l'opéron est réprimé (voir figure 16.4a). L'effecteur contrôlant l'union de l'ADN au répresseur est un métabolite du lactose, l'allolactose, produit quand le lactose est présent. L'allolactose s'unit au répresseur et modifie sa conformation de telle sorte qu'il ne peut plus s'unir à l'opérateur (figure 16.4b). L'opéron est alors induit. Permettant l'induction de l'opéron, l'allolactose est généralement désigné comme l'inducteur.

Le taux de lactose diminuant, l'allolactose voit sa concentration décroître et il y en a trop peu pour s'unir au répresseur et permettre à nouveau l'union de celui-ci à l'ADN. Ce système de contrôle négatif par le répresseur *lac* et son inducteur, l'allolactose, permet donc à la cellule de répondre aux variations du taux de lactose dans l'environnement.

Même en l'absence de lactose, l'opéron *lac* s'exprime à un niveau très faible. Quand il y a du lactose, il entre dans la cellule et la quantité d'allolactose produite suffit à l'induction de l'opéron.

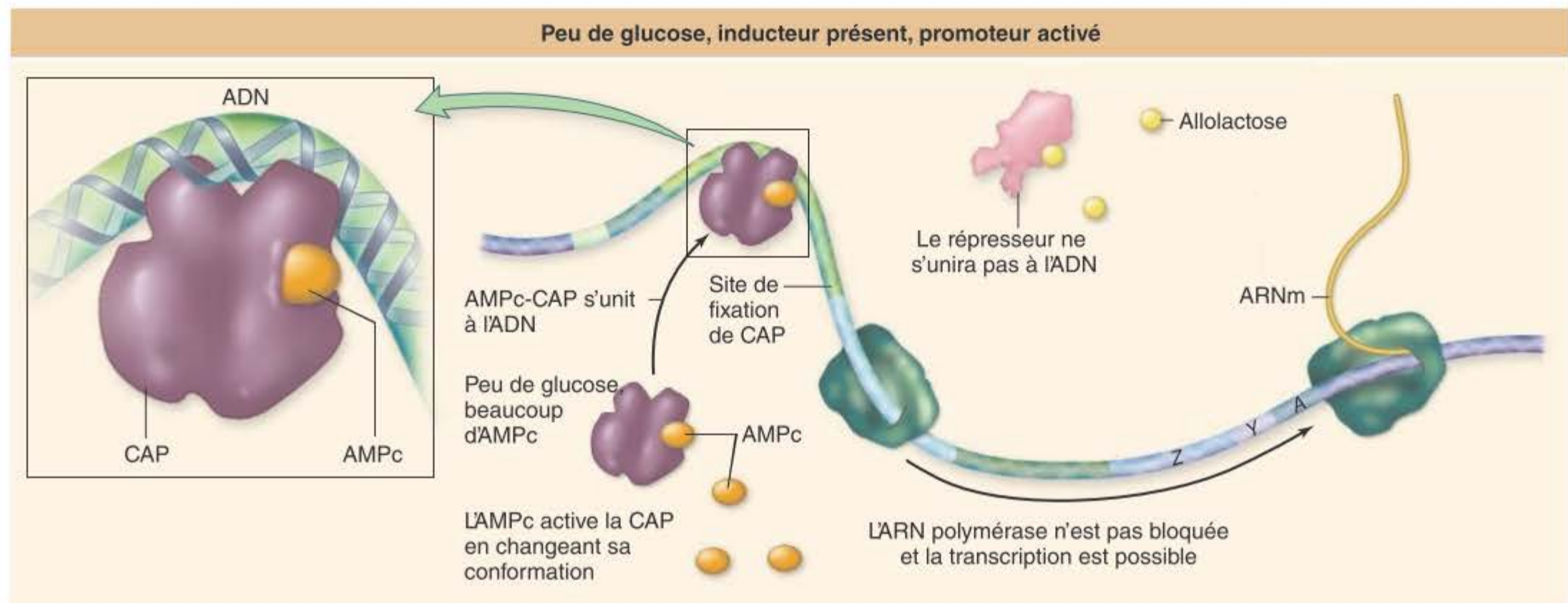
### La présence de glucose empêche l'induction de l'opéron *lac*

La répression par le glucose est due à l'utilisation préférentielle du glucose en présence d'autres sucres comme le lactose. Si les bactéries se déve-

loppent en présence de glucose et de lactose, l'opéron *lac* n'est pas induit. Quand le glucose est épuisé, cet opéron est induit et le lactose peut servir de source d'énergie.

En dépit du terme *répression par le glucose*, ce mécanisme implique une protéine activatrice capable de stimuler la transcription de nombreux opérons cataboliques, y compris l'opéron *lac*. Cet activateur, la **protéine activatrice catabolique (CAP, pour catabolic activator protein)** est une protéine allostérique avec l'AMPc comme effecteur. Cette protéine est aussi appelée **protéine de réponse à l'AMPc (CRP : cAMP response protein)** parce qu'elle s'unit à l'AMPc, mais nous utiliserons CAP pour mettre l'accent sur son rôle de régulateur positif. Seule, la CAP ne s'unit pas à l'ADN, mais son union à l'AMPc modifie sa conformation et permet sa liaison à l'ADN (figure 16.5). Le taux d'AMPc dans la cellule diminue en présence de glucose et la transcription n'est plus stimulée pour les opérons qui réagissent à la CAP.

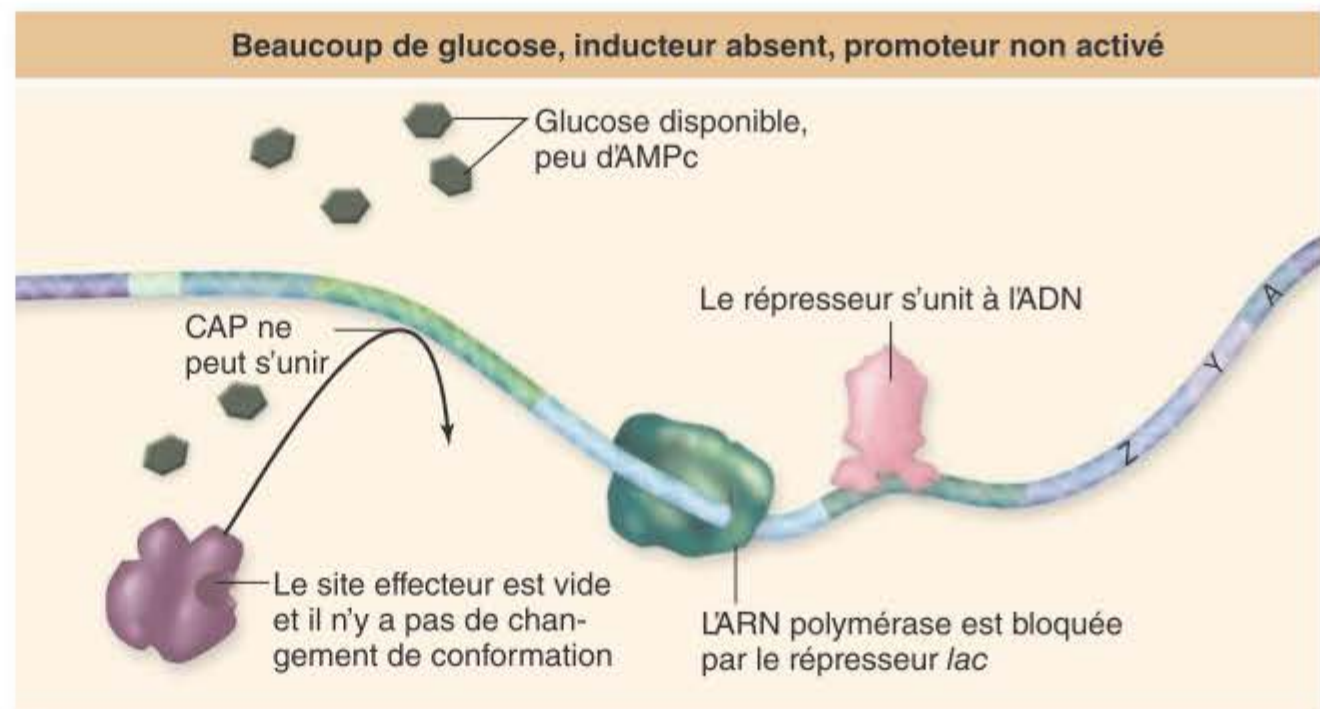
On a longtemps pensé que le système CAP-AMPc était le seul mécanisme de répression par le glucose. Des recherches récentes ont cependant montré que la présence de glucose inhibe le transfert du lactose dans la cellule. La cellule est ainsi privée de l'inducteur de l'opéron *lac*, l'allolactose, et permet l'union du répresseur à l'opérateur. Ce mécanisme, l'**exclusion de l'inducteur**, est actuellement considéré comme la principale forme de répression de l'opéron *lac* par le glucose.



a.

### Figure 16.5 Effet du glucose sur l'opéron *lac*.

L'expression de l'opéron *lac* est contrôlée par un régulateur négatif (répresseur) et un régulateur positif (CAP). Le fonctionnement de la CAP est sensible au taux de glucose. **a.** Pour que la CAP s'unisse à l'ADN, elle doit être unie à l'AMPc. Quand il y a peu de glucose, l'AMPc est abondant et s'unit à la CAP. Le complexe CAP-AMPc provoque l'enroulement de l'ADN autour de lui. La CAP est ainsi au contact de l'ARN polymérase (non représentée), qui s'unit plus efficacement au promoteur. **b.** Une concentration élevée en glucose a deux effets : l'AMPc est rare et la CAP ne peut activer le promoteur, le transport du lactose est bloqué (exclusion de l'inducteur).



b.

À cause de l'exclusion de l'inducteur, le rôle de la CAP en l'absence de glucose semble superflu. Mais, en réalité, le contrôle positif de CAP-AMPC est nécessaire parce que le promoteur de l'opéron *lac* seul fixe l'ARN polymérase de manière peu efficace. Cette efficacité limitée est corrigée par le contrôle positif de l'activateur CAP-AMPC. L'expression est donc maximale en l'absence de glucose et en présence de lactose. Dans ce cas, la présence de l'activateur et l'absence du répresseur se combinent pour aboutir à l'expression maximale (voir figure 16.5).

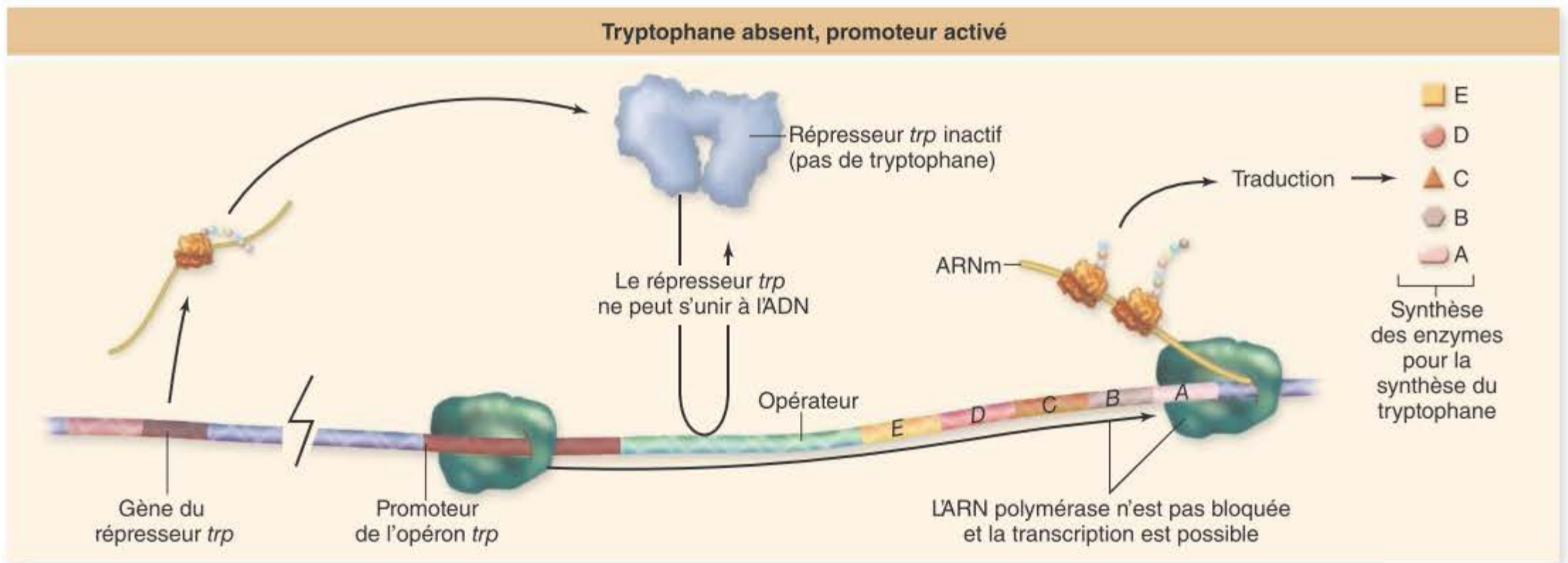
### L'opéron *trp* est contrôlé par le répresseur *trp*

Comme l'opéron *lac*, l'opéron *trp* comporte une série de gènes codant les enzymes qui interviennent dans une même voie biochimique. Dans le cas de l'opéron *trp*, ces enzymes sont nécessaires à la synthèse du

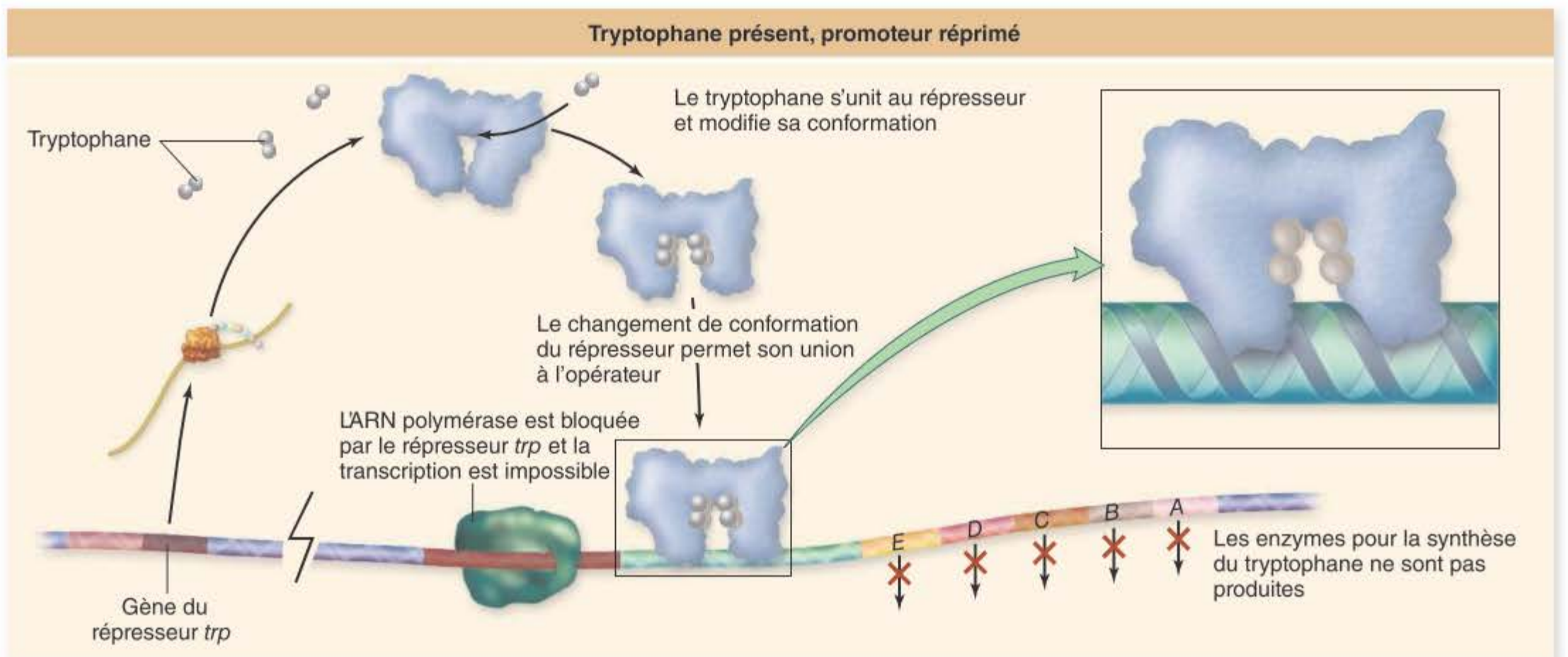
tryptophane. La région régulatrice contrôlant la transcription est située en amont des gènes. L'opéron *trp* est contrôlé par un répresseur codé par un gène localisé en dehors de l'opéron. Cet opéron s'exprime continuellement en l'absence de tryptophane et ne s'exprime pas quand il est présent.

Le répresseur *trp* est une protéine hélice-coude-hélice qui s'unit au site opérateur contigu au promoteur *trp* (figure 16.6). En l'absence de tryptophane, le répresseur *trp* ne s'unit pas à son opérateur, il permet l'expression de l'opéron et la production des enzymes nécessaires à la synthèse du tryptophane.

Quand son taux augmente, le tryptophane (*corépresseur*) s'unit au répresseur et modifie sa conformation, permettant sa liaison à l'opérateur. L'union du complexe répresseur-corépresseur à l'opérateur empêche la liaison de l'ARN polymérase au promoteur. La modification



a.



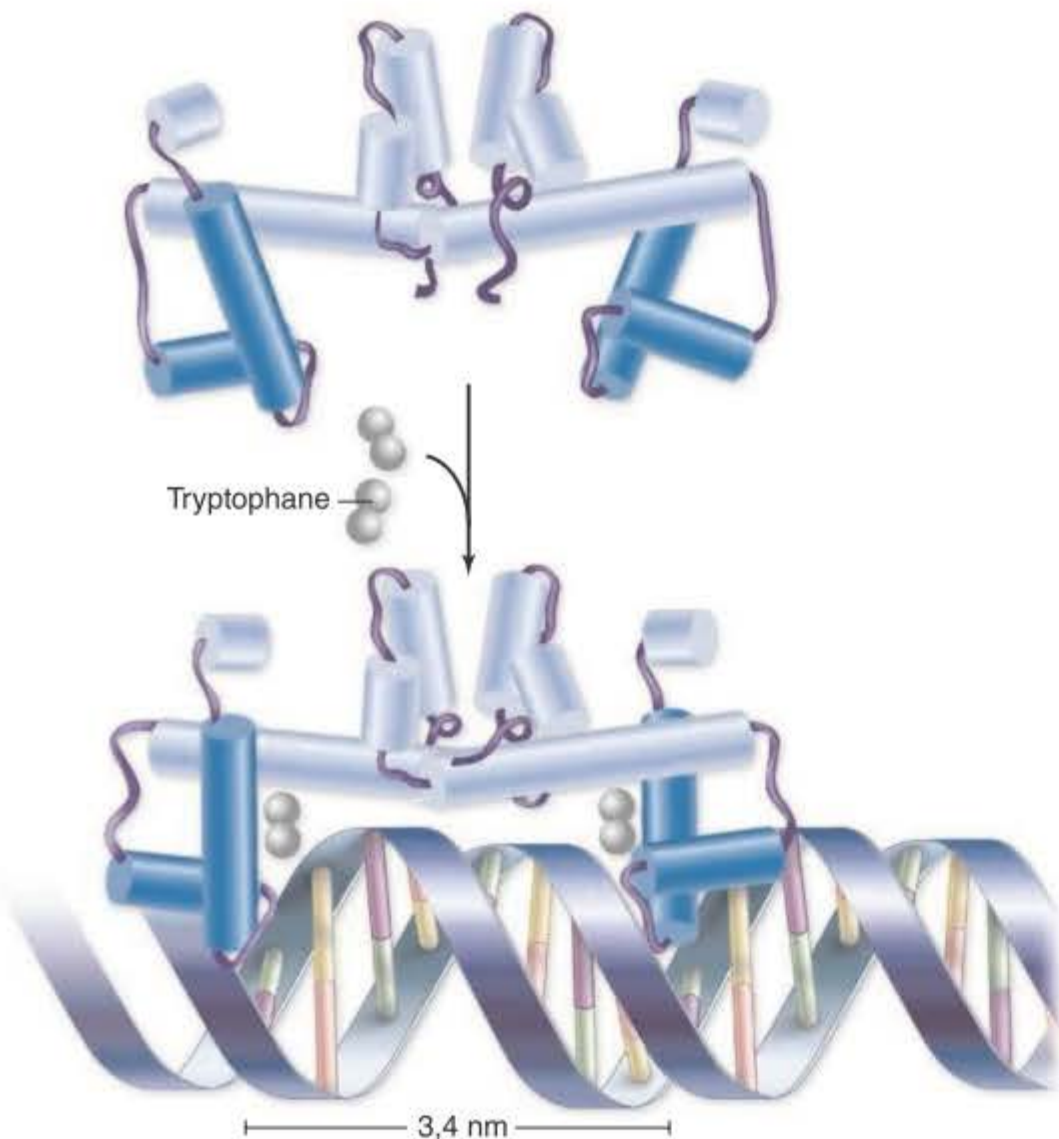
b.

**Figure 16.6 Comment est contrôlé l'opéron *trp*.** L'opéron tryptophane code les enzymes nécessaires à la synthèse du tryptophane. *a.* Le répresseur de tryptophane isolé ne peut s'unir à l'ADN. Le promoteur peut fonctionner librement et l'ARN polymérase transcrit l'opéron. *b.* Si le tryptophane est présent, il s'unit au répresseur, dont la conformation est modifiée, permettant son union à l'ADN. Le complexe tryptophane-répresseur s'unit fermement à l'opérateur et empêche l'ARN polymérase d'entamer la transcription.

de la structure du répresseur par le tryptophane concerne l'orientation d'une paire de motifs hélice-coude-hélice et permet l'insertion de leurs hélices de reconnaissance dans les sillons majeurs contigus de l'ADN (figure 16.7).

Quand le tryptophane est présent et uni au répresseur et quand ce complexe est uni à l'opérateur, on dit que l'opéron est *réprimé*. Quand le taux de tryptophane diminue, le répresseur seul ne peut s'unir à l'opérateur et l'opéron peut s'exprimer. À ce stade, on dit que l'opéron est *déréprimé*, pour distinguer ce stade de l'induction (voir figure 16.6).

Pour comprendre comment l'induction et la répression peuvent résulter d'une régulation négative, il faut connaître le comportement des protéines de répression et de leurs effecteurs. Dans l'induction, le répresseur seul peut s'unir à l'ADN et l'inducteur empêche cette union. En cas de répression, le répresseur ne s'unit à l'ADN que s'il est lié au corépresseur. Induction et répression sont d'excellents exemples de la manière dont les interactions entre molécules peuvent affecter leur structure et montrent l'importance de la structure de la molécule pour son fonctionnement.



**Figure 16.7** L'union du tryptophane modifie la conformation du répresseur. L'union du tryptophane au répresseur augmente la distance entre les deux hélices de reconnaissance du répresseur et permet à celui-ci de se glisser entre deux portions voisines du sillon majeur de l'ADN.

### Questions d'apprentissage 16.3

L'induction résulte du démarrage de l'expression des gènes dans une voie en réponse à un substrat ; la répression est la conséquence d'un blocage de l'expression en réponse à un substrat. Le contrôle négatif de l'opéron *lac* est assuré par une protéine de répression s'unissant à l'ADN et empêchant ainsi la transcription. En présence de lactose, l'opéron est enclenché ; l'allolactose s'unit au répresseur, qui ne peut plus s'associer à l'ADN. Cet opéron est aussi régulé négativement par une protéine activatrice. L'opéron *trp* est contrôlé négativement par un répresseur protéique qui doit être lié au tryptophane pour s'unir à l'ADN. En l'absence de tryptophane, le répresseur ne peut s'unir à l'ADN et l'opéron est déréprimé.

- *Quelle serait la conséquence, pour la régulation de l'opéron *trp*, d'une mutation du répresseur *trp* qui resterait capable de s'unir à *trp*, mais plus à l'ADN ?*

## 16.4 La régulation eucaryote

### Objectifs

1. *Distinguer le rôle des facteurs de transcription généraux et spécifiques.*
2. *Décrire la formation d'un complexe d'initiation Pol II.*
3. *Expliquer comment les facteurs de transcription peuvent agir à distance dans l'ADN.*

Le contrôle de la transcription est beaucoup plus complexe chez les eucaryotes que chez les procaryotes. Les concepts de base des interactions entre protéines et ADN restent valides, mais les protéines qui interagissent sont différentes et plus nombreuses. La séparation de la transcription et de la traduction dans des compartiments cellulaires différents entraîne des complications ; c'est aussi le cas pour l'organisation de l'ADN en chromatine. Avec le temps, nous sommes arrivés à une représentation précise du contrôle de l'initiation de la transcription. Cette représentation reste valable, mais le nouveau mode de séquençage de l'ADN et l'analyse des éléments fonctionnels de l'ADN au niveau du génome est en train de modifier ce schéma (voir chapitre 18). Ainsi, on a constaté que la partie du génome qui est transcrite est plus importante qu'on ne le pensait, mais la signification de cette transcription n'est pas encore claire. Dans ce chapitre, nous nous concentrons sur ce que nous connaissons du mode d'interaction des protéines avec l'ADN pour le contrôle de la transcription, et sur la manière dont peut intervenir la structure de la chromatine.

### Les facteurs de transcription peuvent être généraux ou spécifiques

Au chapitre 15, nous avons introduit le concept des facteurs de transcription. La transcription eucaryote exige plusieurs de ces protéines, qui

appartiennent à deux catégories : les *facteurs de transcription généraux* et les *facteurs de transcription spécifiques*. Les facteurs généraux sont nécessaires à l'assemblage d'un appareil de transcription et au recrutement de l'ARN polymérase II pour un promoteur. Les facteurs spécifiques augmentent le taux de transcription dans certains types de cellules ou en réponse à des signaux.

### Facteurs de transcription généraux

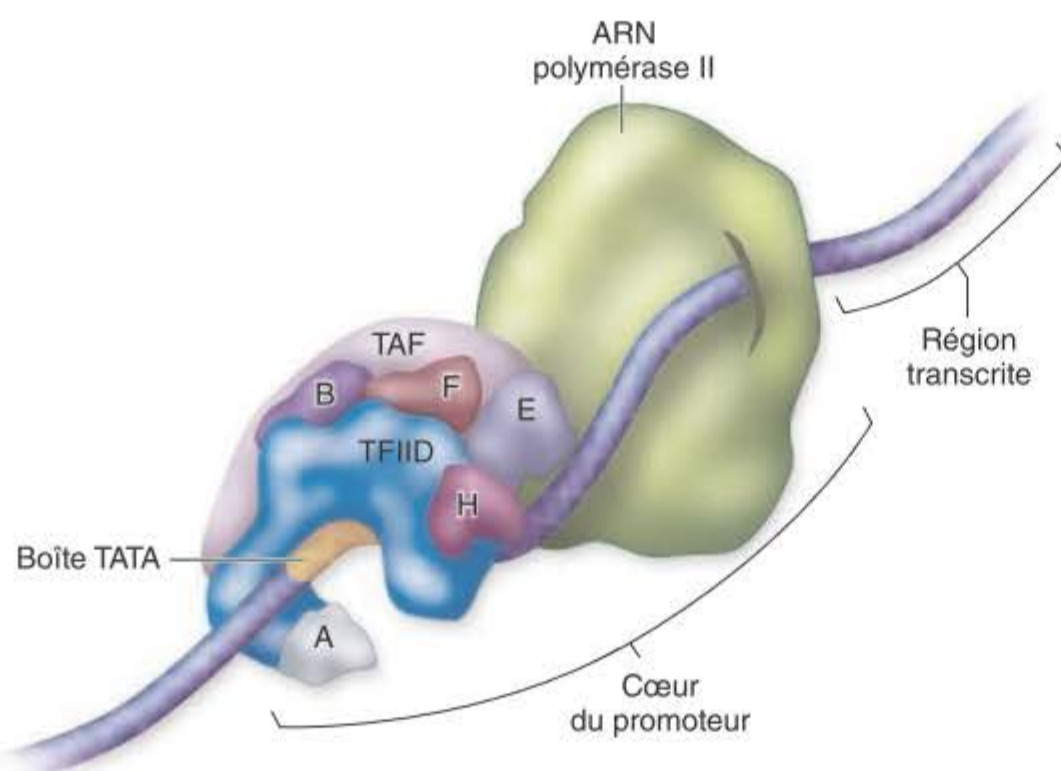
À elle seule, l'ARN polymérase II n'est pas capable d'assurer l'initiation de la transcription des gènes codant des protéines. Une foule de **facteurs de transcription généraux** sont en outre nécessaires pour une initiation efficace. Ces facteurs sont nécessaires pour la transcription, mais ils ne vont pas au-delà de ce niveau de base.

On désigne les facteurs de transcription généraux par le sigle TFII, pour « facteur de transcription de l'ARN polymérase II » suivi d'une lettre. Le plus important de ces facteurs, TFIID, comprend la protéine de liaison à TATA qui reconnaît la séquence de la boîte TATA présente dans beaucoup de promoteurs eucaryotes.

L'union de TFIID est suivie par celle de TFIIIE, TFIIIF, TFIIA, TFIIB et TFIIH, ainsi que de nombreux facteurs accessoires, les *facteurs associés à la transcription*, ou TAF (*transcription associated factors*). Le *complexe d'initiation* qui en résulte (figure 16.8) est visiblement beaucoup plus complexe que l'ARN polymérase holoenzyme bactérienne s'unissant à un promoteur. Et il y a encore un autre niveau de complexité : bien que capable d'initier la synthèse au niveau de base, le complexe d'initiation ne peut assurer un niveau élevé de transcription sans la participation d'autres facteurs spécifiques.

### Facteurs de transcription spécifiques

Les **facteurs de transcription spécifiques** agissent en fonction du tissu – ou du moment – pour stimuler une transcription supérieure au



**Figure 16.8** Formation du complexe d'initiation eucaryote.

Le facteur général de transcription TFIID s'unit à la boîte TATA et est rejoint par les autres facteurs généraux, TFIIIE, TFIIIF, TFIIA, TFIIB et TFIIH. Ce complexe est complété par plusieurs facteurs associés à la transcription (TAF) qui amènent ensemble la molécule d'ARN Pol II au sein du promoteur.

niveau de base. Le nombre et la diversité de ces facteurs sont énormes. On peut se faire une idée de cette prolifération de facteurs en se concentrant sur le motif d'union à l'ADN plutôt que sur les facteurs spécifiques.

L'étude de ces facteurs fait apparaître un thème commun important : des facteurs de transcription spécifiques, les *activateurs*, possèdent une organisation en domaines. Chacun possède un domaine d'union à l'ADN et un autre domaine d'activation qui interagit avec l'appareil de transcription, et ces domaines sont essentiellement indépendants dans la protéine. L'« échange » de domaines de liaison à l'ADN entre facteurs différents modifie la spécificité de liaison de ces facteurs sans que leur faculté d'activer la transcription soit affectée.

### Les promoteurs et les amplificateurs sont des sites de liaison pour les facteurs de transcription

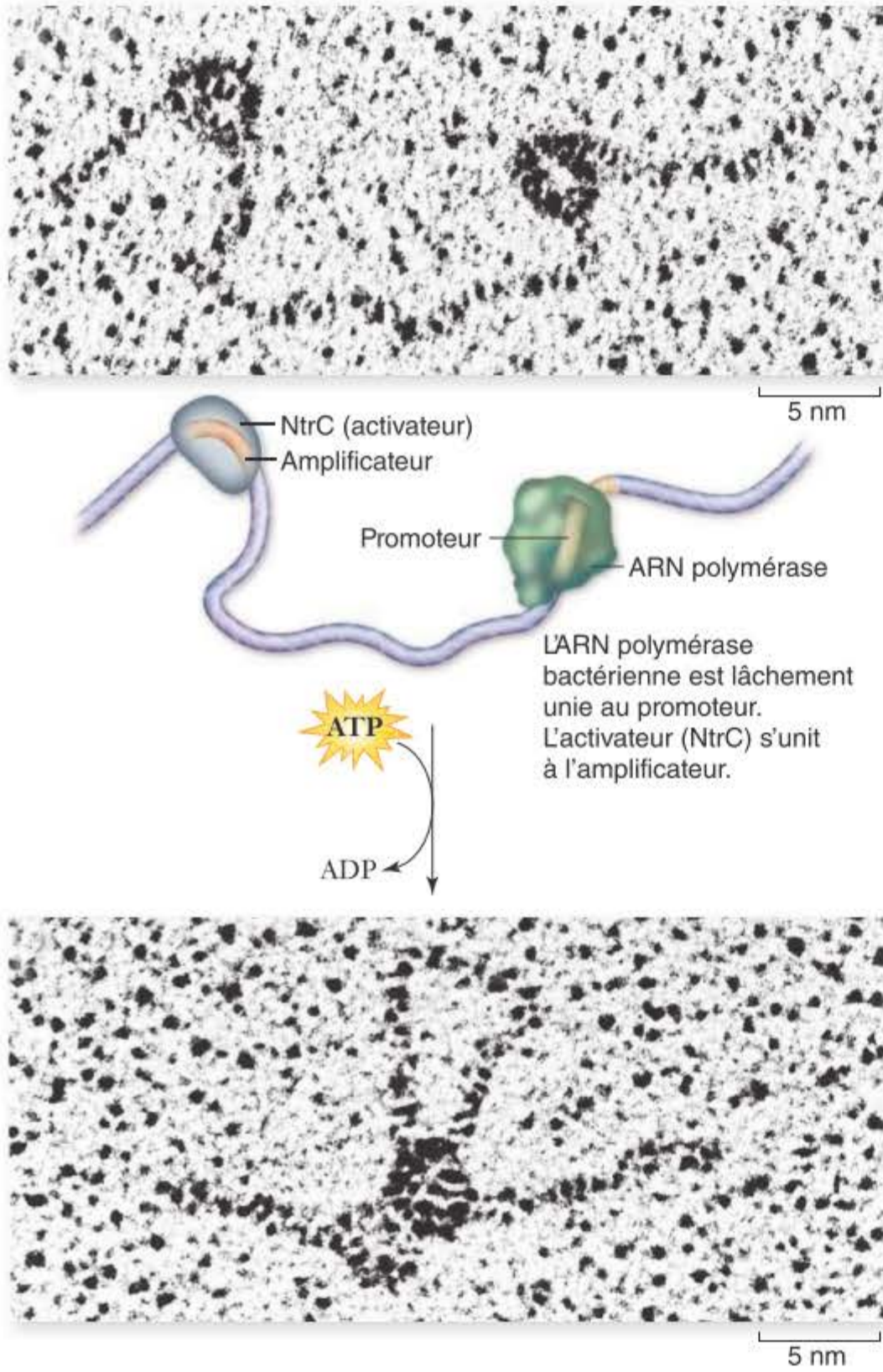
On a vu au chapitre précédent que les promoteurs possèdent des sites de fixation à l'ADN pour les facteurs de transcription généraux. Ces facteurs interviennent ensuite dans l'union de l'ARN polymérase II au promoteur (et aussi dans l'union des ARN polymérases I et III à leurs promoteurs spécifiques). Au contraire, la portion holoenzyme de l'ARN polymérase des procaryotes peut directement reconnaître un promoteur et s'y unir.

À l'origine, on a défini les **amplificateurs** comme des séquences d'ADN nécessaires pour une transcription de haut niveau, capables de fonctionner indépendamment de leur position ou de leur orientation. Au début, cela paraissait difficile à concevoir, surtout parce que les biologistes moléculaires étaient conditionnés par les systèmes procaryotes et s'attendaient à trouver des régions de contrôle immédiatement en amont de la région codante. Les amplificateurs sont en fait les sites d'union pour des facteurs de transcription spécifiques. La possibilité pour les amplificateurs d'agir à longue distance fut d'abord une énigme, mais les chercheurs pensent aujourd'hui que cette action passe par la formation d'une boucle d'ADN qui rapproche l'amplificateur du promoteur.

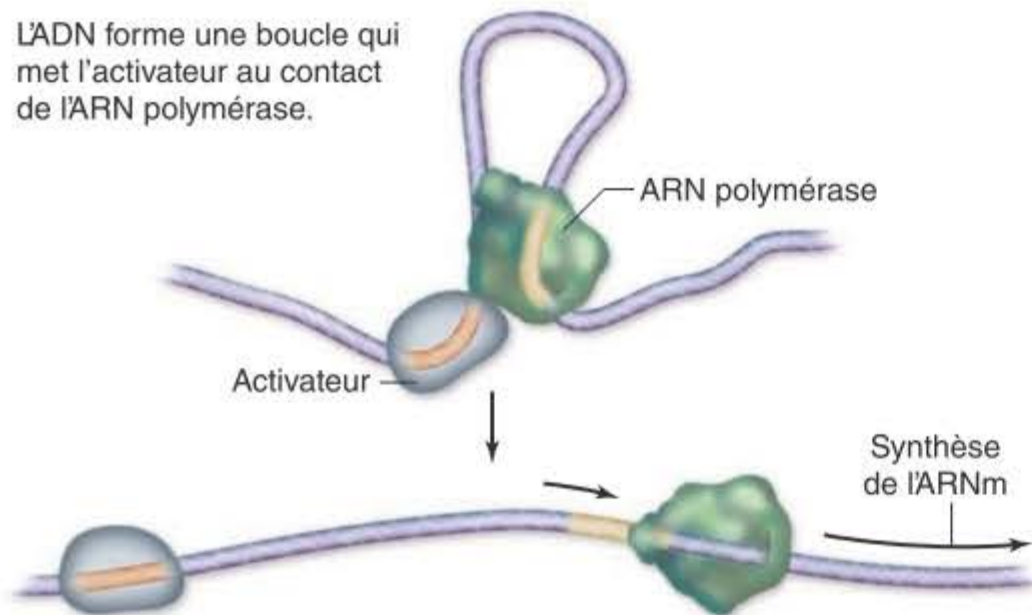
Cette formation de boucles est plus importante dans les systèmes eucaryotes, mais on l'a d'abord démontrée pour des protéines procaryotes de liaison à l'ADN (figure 16.9). Le point important est qu'une longue distance entre deux sites ne correspond pas à une grande distance physique parce que la flexibilité de l'ADN permet la formation de boucles. Un activateur uni à un amplificateur peut ainsi entrer en contact avec des facteurs de transcription unis à un promoteur éloigné (figure 16.10).

### Les coactivateurs et les médiateurs unissent les facteurs de transcription à l'ARN polymérase II

D'autres facteurs interviennent spécifiquement dans le fonctionnement des facteurs de transcription. Ces *coactivateurs* et *médiateurs* sont également nécessaires à l'activation de la transcription par le facteur de transcription. Ils agissent en s'unissant au facteur de transcription et en se

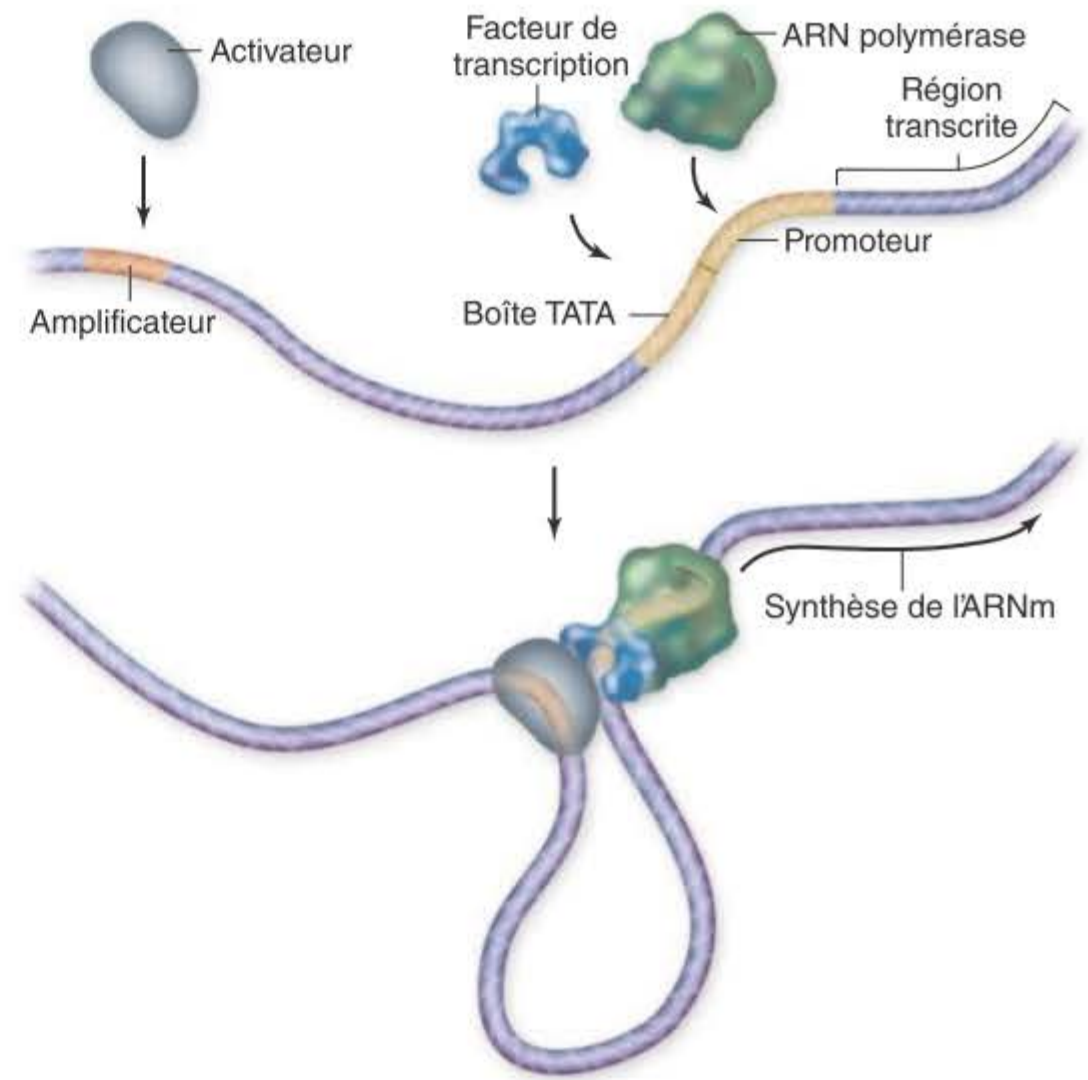


L'ADN forme une boucle qui met l'activateur au contact de l'ARN polymérase.



L'activateur déclenche l'activation de l'ARN polymérase et la transcription débute. La boucle d'ADN se déroule.

**Figure 16.9** Boucle d'ADN provoquée par des protéines. Quand l'activateur bactérien NtrC s'unit à un amplificateur, l'ADN forme une boucle jusqu'à un site éloigné auquel l'ARN polymérase est unie et active ainsi la transcription. Ces amplificateurs sont rares chez les procaryotes, mais communs chez les eucaryotes.



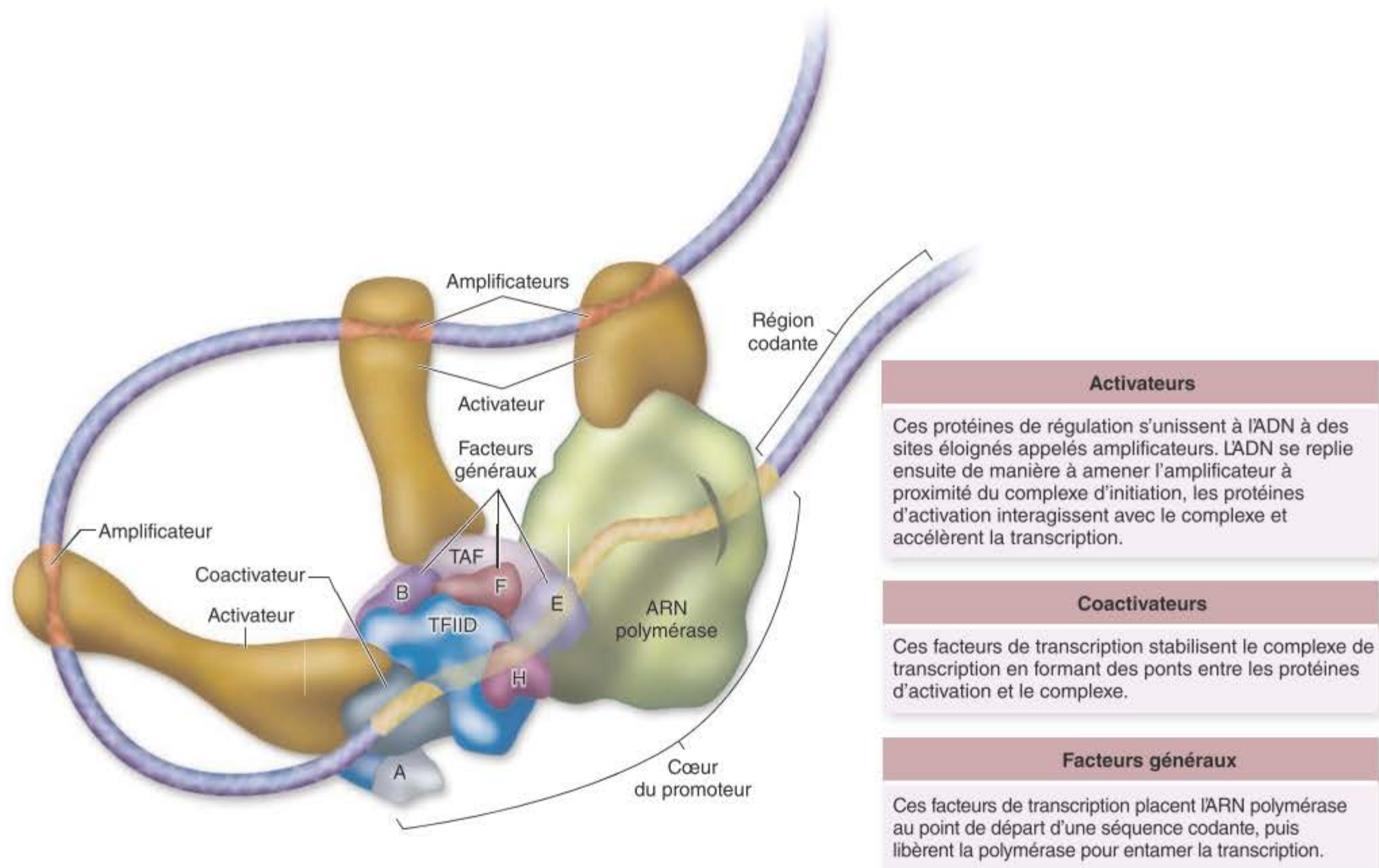
**Figure 16.10** Comment fonctionnent les amplificateurs. Le site amplificateur se trouve loin du gène à contrôler. L'union d'un activateur (*en gris*) à l'amplificateur permet son interaction avec les facteurs de transcription (*en bleu*) associés à l'ARN polymérase et elle active la transcription.

fixant ensuite à une autre partie de l'appareil de transcription. Les médiateurs sont indispensables au fonctionnement de certains facteurs de transcription, mais tous n'en ont pas besoin. Il existe beaucoup moins de coactivateurs que de facteurs de transcription parce qu'un même coactivateur peut servir à de nombreux facteurs de transcription

### Le complexe de transcription rassemble tout

Quelques principes généraux s'appliquent à une vaste gamme de situations, mais presque chaque gène eucaryote – ou groupe de gènes régulé de façon coordonnée – représente un cas particulier. Pratiquement tous les gènes transcrits par l'ARN polymérase II ont besoin de la même série de facteurs généraux pour assembler un complexe d'initiation, mais l'assemblage de ce complexe et son niveau final de transcription dépendent des facteurs spécifiques de transcription dont la combinaison permet la mise en forme du **complexe de transcription** (figure 16.11).

La mise en forme des promoteurs eucaryotes peut donc être très simple, si l'on tient compte uniquement de ce qui est nécessaire au complexe d'initiation, ou très compliquée, si l'on considère tous les facteurs susceptibles de s'unir en un complexe et d'influencer la transcription. Ce type de régulation génique combinatoire aboutit à une grande souplesse parce qu'il répond aux nombreux signaux affectant la transcription que peut recevoir une cellule, en permettant l'intégration de ces signaux.



**Figure 16.11 Interactions entre les différents facteurs au sein du complexe de transcription.** Tous les facteurs de transcription spécifiques s'unissent à des séquences d'amplification qui peuvent être éloignées du promoteur. Ces protéines peuvent ensuite interagir avec le complexe d'initiation en formant des boucles d'ADN qui amènent les facteurs à proximité du complexe d'initiation. Comme il est précisé dans le texte, certains facteurs de transcription, appelés activateurs, peuvent interagir directement avec l'ARN polymérase II ou avec le complexe d'initiation, alors que d'autres ont besoin de coactivateurs supplémentaires.

**?** **Question** Comment les eucaryotes coordonnent-ils l'activation de nombreux gènes dont la transcription doit être simultanée ?

### Questions d'apprentissage 16.4

Chez les eucaryotes, l'initiation a besoin de facteurs de transcription qui s'unissent au promoteur et recrutent l'ARN polymérase II pour former un complexe de transcription. Les facteurs généraux produisent le niveau de base de transcription. Les facteurs de transcription spécifiques, qui s'unissent à des séquences amplificatrices, peuvent accroître le taux de transcription. Les amplificateurs peuvent agir à distance parce que l'ADN est capable de former des boucles et de rapprocher un amplificateur d'un promoteur. Des coactivateurs et médiateurs supplémentaires unissent certains facteurs de transcription spécifiques à l'ARN polymérase II.

- *Quelle serait la conséquence d'une mutation entraînant la perte d'un facteur de transcription général ou d'un facteur spécifique ?*

## 16.5 Structure de la chromatine et expression génique

### Objectifs

1. *Décrire au moins deux types d'empreinte épigénétique.*
2. *Expliquer le rôle des complexes de remodelage de la chromatine.*

Dans une cellule diploïde humaine, quelque  $6 \times 10^9$  pb sont emballées dans  $3 \times 10^7$  nucléosomes. La chromatine est le résultat d'une condensation supplémentaire et permet de loger quelque 2 m d'ADN dans un noyau de 5 à 10  $\mu\text{m}$ . À l'évidence, cela complique le processus de transcription. La cellule a résolu ce problème en modulant sélectivement la structure de la chromatine par modification de l'ADN et des histones, afin de permettre les interactions de l'ADN nécessaires à la vie. On considère que ces modifications de la chromatine sont à la base des

modifications épigénétiques-héréditaires du phénotype ne provenant pas de modifications de la séquence d'ADN.

## La méthylation de l'ADN, la modification des histones et l'ARN non codant peuvent affecter la structure de la chromatine

Une *altération épigénétique* peut se définir comme une modification qui persiste en l'absence du stimulus qui l'a engendrée et qui est héritée au cours des divisions cellulaires. Cette définition peut convenir pour l'altération de la structure de la chromatine, bien que les mécanismes de l'hérédité mitotique ne soient pas encore élucidés. La structure de la chromatine est affectée par des modifications très diverses des histones aussi bien que par la méthylation de l'ADN, et même par des ARN non-codants.

### Méthylation de l'ADN

La méthylation de l'ADN fut la première modification de la structure de la chromatine ayant un effet épigénétique. L'addition d'un groupement méthyle à la cytosine donne la 5-méthylcytosine, mais ce changement n'affecte pas l'appariement avec la guanine (figure 16.12). Il existe une corrélation entre les niveaux élevés de méthylation de l'ADN et les gènes inactifs, et l'expression d'un allèle particulier observée dans l'empreinte génomique (voir chapitre 13) est, au moins en partie, liée à la méthylation de l'ADN.

La méthylation satisfait à la définition précise de l'effecteur épigénétique par des mécanismes bien connus. La réplication semi-conservative de l'ADN méthylé sur les deux brins donne de l'ADN à demi-méthylé qui devient entièrement méthylé grâce à une méthylase de maintenance. Chez les humains, DMNT1 est la méthylase de maintenance. D'autres méthylases peuvent induire le processus, mais elles n'ont pas besoin de rester actives.

### Inactivation du chromosome X

Chez les mammifères, les femelles inactivent un chromosome X par une sorte de compensation de la dose pour que l'expression du chromosome X soit égale dans les deux sexes. Ce mécanisme, proposé d'abord par Mary Lyon en 1961, reste aujourd'hui un sujet de recherche active. C'est aussi le premier exemple de régulation de l'expression génique par un long ARN non codant. La région du chromosome qui

déclenche le processus d'inactivation renferme un gène appelé *transcrit spécifique de l'inactivation de X*, ou *Xist* (*X-inactivation-specific transcript*), qui ne code pas de protéine. Le mécanisme d'inactivation est assez complexe, mais, selon une version courante, *Xist* « enrobe » l'ensemble du chromosome X inactif et recrute le complexe régulateur de la chromatine 2 (PRC2). Il en résulte une modification des histones associées à la chromatine inactive.

### Modification des histones

La modification des histones est un sujet complexe parce que les modifications potentielles sont nombreuses, et quatre histones peuvent être modifiées. Les modifications peuvent être l'acétylation et la méthylation de la lysine, ainsi que la phosphorylation de la sérine, de la thréonine et de la tyrosine. En général, l'acétylation, en particulier celle de H3, est liée aux sites de transcription, dans les régions régulatrices comme dans la région transcrite du gène lui-même. Mais la méthylation de cette même histone (H3) peut avoir l'effet opposé, en fonction de la lysine méthylée.

## Certains activateurs de la transcription modifient la structure de la chromatine

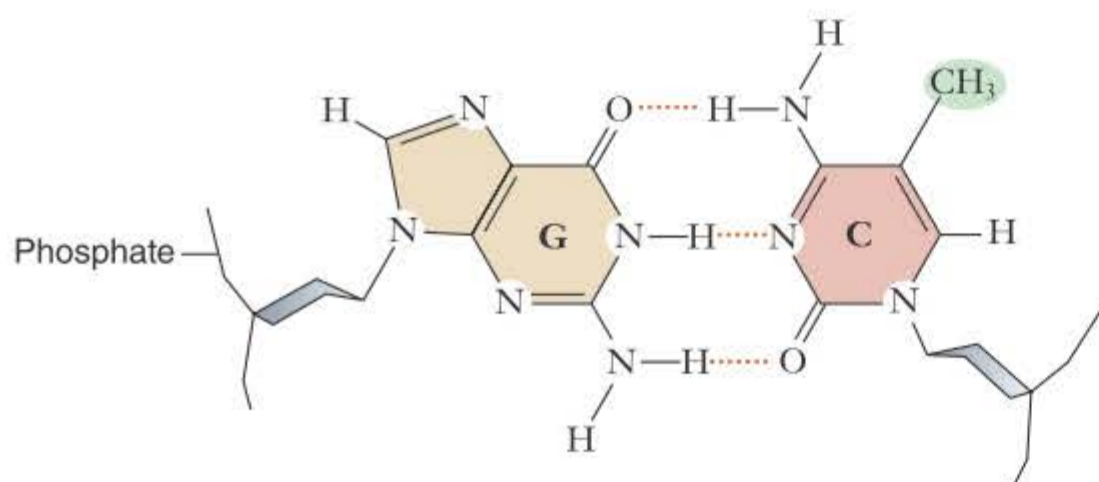
Le contrôle de la transcription eucaryote exige la présence de nombreux facteurs différents qui l'activent. On a vu que certains activateurs semblent agir par une interaction directe avec le complexe d'initiation ou avec les coactivateurs qui interagissent eux-mêmes avec ce complexe. D'autres cas sont moins clairs. On a montré récemment que certains coactivateurs sont des histone acétylases. On constate alors une augmentation de la transcription suite à la suppression des structures de niveau supérieur de la chromatine susceptibles de l'empêcher (figure 16.13). On a aussi montré que certains corépresseurs sont des histone désacétylases.

Ces observations ont fait croire à l'existence d'un « code d'histone » analogue au code génétique. Ce code d'histone contrôlerait la structure de la chromatine, et donc l'accès à l'ADN de l'équipement de transcription.

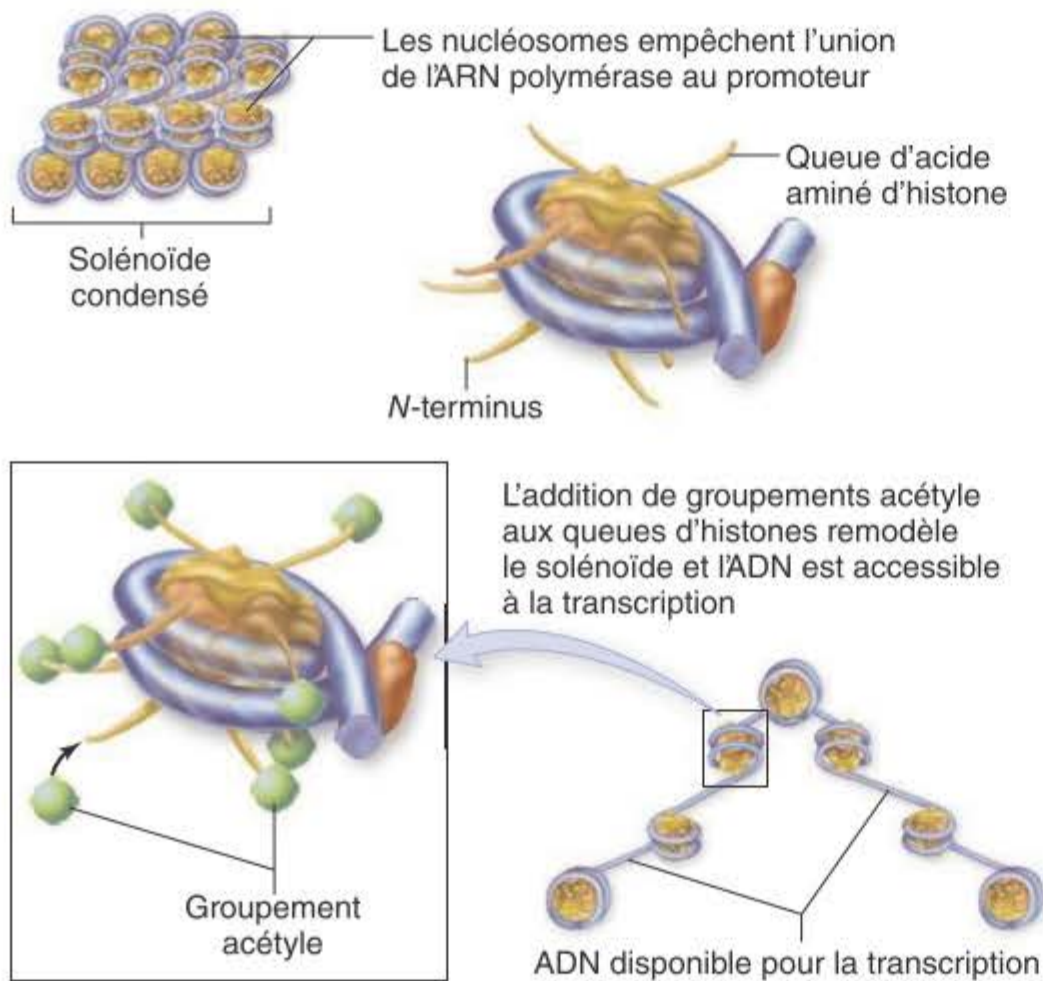
## Les complexes de remodelage de la chromatine modifient également la structure de la chromatine

On commence à comprendre la manière dont les modifications de la structure de la chromatine peuvent contrôler l'expression génique. Une découverte clé est l'existence de ce que l'on appelle les **complexes de remodelage de la chromatine**. Dans ces grands complexes de protéines se trouvent des enzymes qui modifient les histones et l'ADN, ainsi que la structure même de la chromatine.

Une catégorie de ces facteurs de remodelage, les facteurs de remodelage de la chromatine dépendant de l'ATP, fonctionnent comme des moteurs moléculaires affectant l'ADN et les histones. Grâce à l'énergie de l'ATP, ces facteurs modifient les rapports entre les histones et l'ADN. Ils peuvent catalyser quatre modifications différentes des liaisons entre histones et ADN (figure 16.14) : (1) faire glisser les nucléosomes le long de l'ADN, (2) créer une forme remodelée où l'ADN est plus accessible, (3) éliminer les nucléosomes de l'ADN et (4) remplacer les histones par des formes variantes. Avec ces modifications, l'ADN est plus accessible aux protéines de régulation qui, à leur tour, affectent l'expression génique.



**Figure 16.12 Méthylation de l'ADN.** La cytosine est méthylée en 5-méthylcytosine. Étant situé latéralement, le groupement méthyle (*en vert*) n'interfère pas avec les liaisons hydrogène d'une paire de bases G-C, mais il peut être identifié par les protéines.



**Figure 16.13.** La modification des histones affecte la chromatine. L'ADN des eucaryotes est d'abord organisé en nucléosomes, puis en autres structures de la chromatine. Les histones du cœur des nucléosomes portent des queues amine protubérantes. Ces queues peuvent être modifiées par addition de groupements acétylé. L'acétylation modifie la structure de la chromatine, qui devient accessible au système de transcription.

### Questions d'apprentissage 16.5

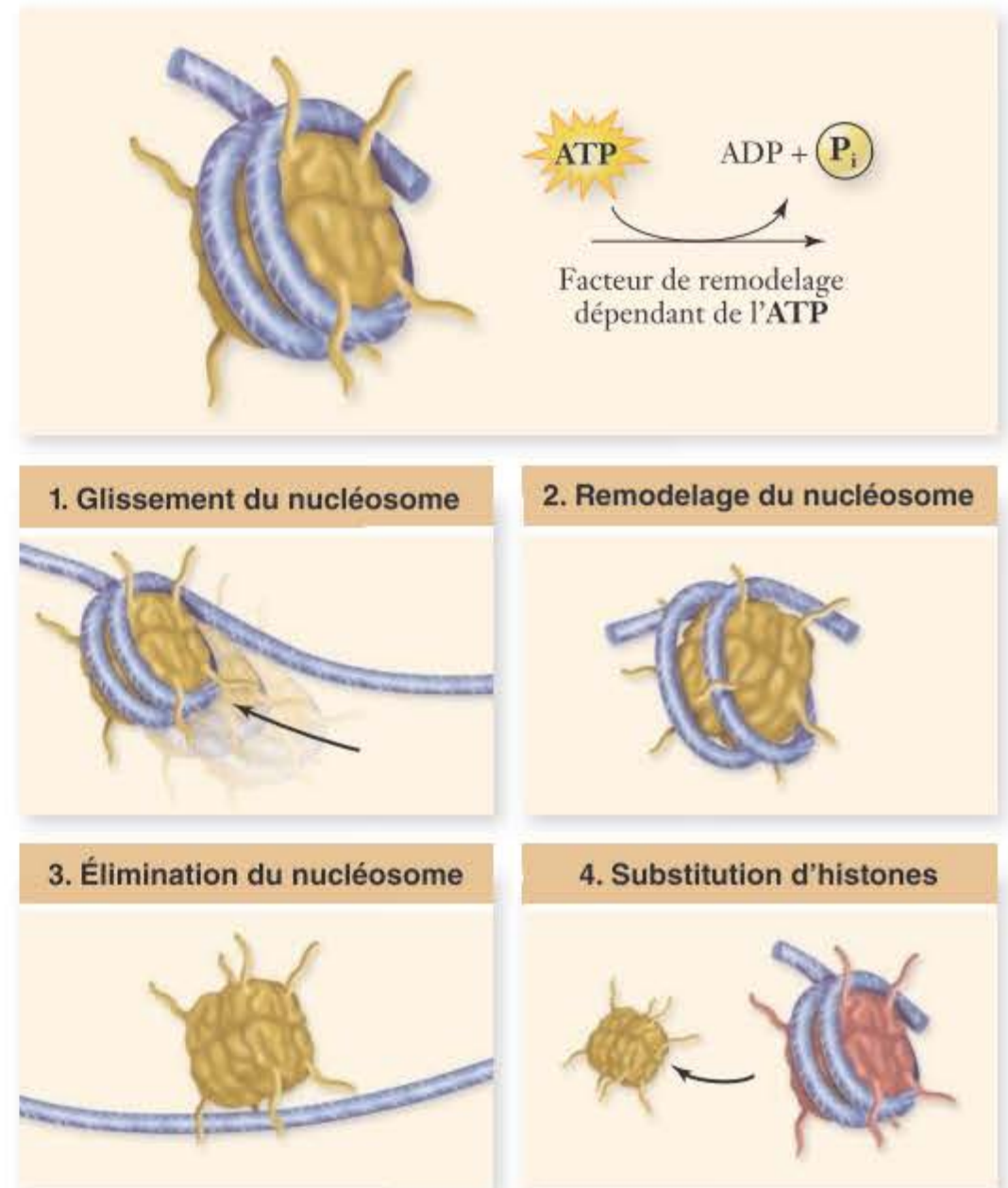
L'ADN eucaryote est empaqueté sous forme de chromatine, nouveau défi structural pour la transcription. Les modifications de la structure de la chromatine sont en rapport avec une modification de l'ADN et des histones, et paraissent être à l'origine des modifications épigénétiques. Certains activateurs de la transcription modifient les histones par acétylation. De grands complexes de remodelage de la chromatine contiennent des enzymes qui modifient la structure de la chromatine et rendent l'ADN plus accessible aux protéines régulatrices.

- On parle de « gènes domestiques » pour désigner les gènes en activité dans toutes les cellules. Expliquer l'idée qui sous-tend ce nom.

## 16.6 Contrôle post-transcriptionnel des eucaryotes

### Objectifs

1. Expliquer comment les petits ARN peuvent affecter l'expression génique.
2. Distinguer les différents types de régulation post-transcriptionnelle.



**Figure 16.14** Rôle des facteurs de remodelage dépendant de l'ATP. Les facteurs de remodelage dépendant de l'ATP utilisent l'énergie de l'ATP pour modifier la structure de la chromatine. Ils peuvent (1) faire glisser les nucléosomes le long de l'ADN et libérer des sites de fixation pour les protéines ; (2) remodeler la chromatine et créer des sites où l'ADN est plus accessible aux protéines ; (3) éliminer complètement les nucléosomes de l'ADN ; (4) remplacer les histones des nucléosomes par des formes variantes d'histones.

La séparation entre la transcription dans le noyau et la traduction dans le cytoplasme, chez les eucaryotes, est à l'origine de points de régulation qui n'existent pas chez les procaryotes. On pense depuis longtemps à ces formes « alternatives » de régulation, mais il apparaît à présent que leur rôle est beaucoup plus important qu'on ne le supposait auparavant. Dans cette section, nous allons envisager plusieurs de ces mécanismes de contrôle de l'expression génique en partant du domaine nouveau et passionnant des petits ARN.

### Les petits ARN contrôlent l'expression génique après la transcription

La génétique du développement a permis d'importants progrès dans la compréhension de la régulation de l'expression génique. Un exemple frappant est la découverte des petits ARN qui interviennent dans l'expression des gènes. On savait que le mutant *lin-4* modifiait le timing du nématode *C. elegans*, et les études génétiques ont montré que *lin-4* contrôle un autre gène, *lin-14*. Quand Ambros, Lee et Feinbaum eurent isolé le gène *lin-4* en 1992, ils constatèrent qu'il ne codait pas une protéine. Le gène *lin-4* ne code que deux petites molécules d'ARN, l'une de 22 et l'autre de 61 nucléotides. En outre, l'ARN de 22 nucléotides dérive de l'ARN de 61 nucléotides. On a montré ensuite que ce

petit ARN était complémentaire d'une région de *lin-14*. On a proposé un schéma montrant le rôle de l'ARN *lin-4* comme répresseur de la traduction de l'ARNm *lin-14* (figure 16.15). C'étaient les premiers **microARN**, ou **miARN** identifiés, mais on ne les appelait pas ainsi à cette époque.

Une autre voie de recherche totalement différente impliquait le blocage de l'expression des gènes par des ARN bicaténaires. On a montré

que ce phénomène était dû à une autre catégorie de petits ARN, les **petits ARN d'interférence**, ou **siARN**. Ils peuvent être introduits expérimentalement, provenir de virus invasifs, ou même être codés dans le génome. L'utilisation des siARN pour contrôler l'expression génique a montré l'existence de mécanismes cellulaires contrôlant cette expression par les petits ARN.

Depuis que l'on a découvert que les petits ARN rendaient les gènes silencieux, on s'est beaucoup intéressé à ce mécanisme pour ses applications expérimentales et aussi pour expliquer le contrôle post-traductionnel de l'expression génique. On a découvert une masse de nouveaux types de petits ARN, mais nous nous limiterons aux deux classes bien connues, les miARN et les siARN pour illustrer le mécanisme rendant l'ARN silencieux.

### Les gènes des miARN

Le rôle des miARN dans l'expression génique a d'abord semblé limité aux nématodes, le gène *lin-4* n'ayant pas d'homologues évidents dans d'autres systèmes. Sept années après leur découverte, on a trouvé un second gène, *let-7*, dans la même voie chez *C. elegans*. Le gène *let-7* codait aussi un ARN de 22 nucléotides capable d'influencer la traduction. Dans ce cas, on a immédiatement trouvé des homologues de *let-7* chez la drosophile et les humains.

Suite à la découverte d'un nombre croissant de miARN dans des organismes différents, on est passé à une recherche informatique et à des méthodes plus efficaces, comme les microalignements et le séquençage à haute capacité. Une banque de données spécialisée comprend actuellement 695 séquences de miARN humains.

On trouve des gènes de miARN à des endroits divers comme les introns de gènes qui s'expriment, et ils sont souvent groupés dans une même unité de transcription. On les trouve aussi dans des régions du génome auparavant considérées comme silencieuses au point de vue de la transcription. Cette découverte est particulièrement passionnante parce qu'une autre recherche sur la transcription dans les génomes animaux a montré que beaucoup de ce que nous pensions être silencieux au point de vue transcription ne l'est en réalité pas.

### Biogenèse et fonction des petits ARN

La synthèse d'un miARN débute dans le noyau et aboutit, dans le cytoplasme, à un ARN d'environ 22 nucléotides qui réprime l'expression génique (figure 16.16). À partir d'un gène de miARN, l'ARN polymérase II synthétise un transcrit initial appelé pri-miARN. La région de ce transcrit contenant le miARN peut se replier sur elle-même et former, par appariement de bases, une structure avec tige et boucle. Cette structure est scindée, dans le noyau, par une nucléase appelée Drosha qui ne garde que la structure en tige et boucle, le pré-miARN. Ce pré-miARN sort du noyau par un pore nucléaire, uni à une protéine, l'exportine 5. Dans le cytoplasme, le pré-miARN est encore scindé par une autre nucléase appelée Dicer et donne un court ARN bicaténaire contenant le miARN. Celui-ci est chargé dans un complexe protéique dénommé **RISC** (*RNA induced silencing complex*). Le RISC comprend la protéine de liaison à l'ARN Argonaute (Ago) qui interagit avec le miARN. Le brin complémentaire est éliminé par une nucléase ou pendant le chargement.

À ce stade, le RISC est orienté afin de réprimer l'expression d'autres gènes en fonction de la complémentarité des séquences du miARN. La région complémentaire se trouve généralement dans la région 3' non traduite des gènes et le résultat peut être la scission de l'ARNm ou l'inhibition de la traduction. Chez les animaux, cette

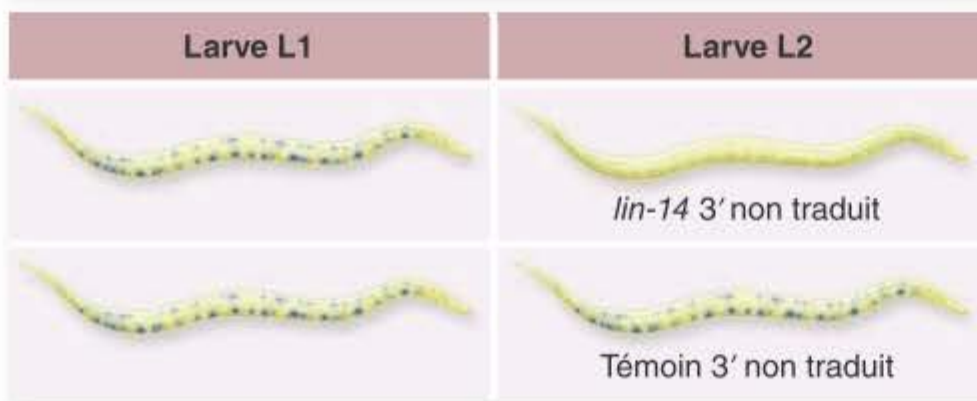
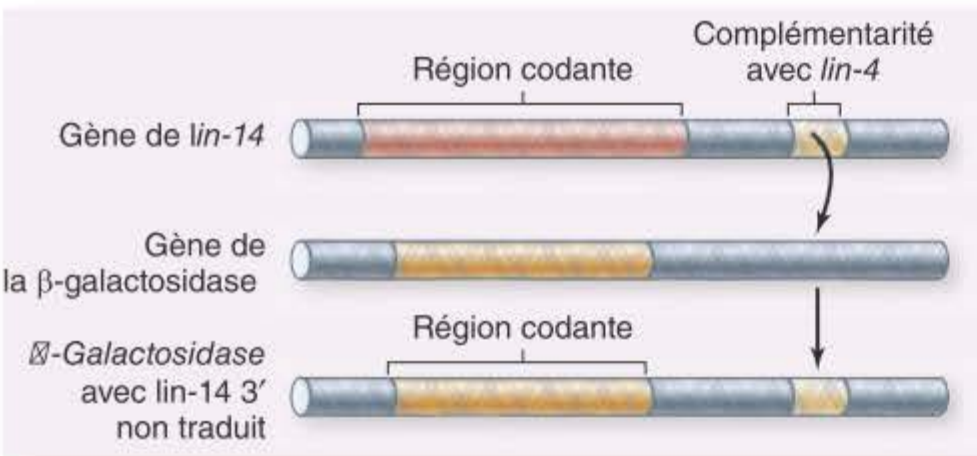
## DÉMARCHE SCIENTIFIQUE

**Hypothèse:** La région du gène *lin-14* complémentaire du miARN *lin-4* contrôle l'expression de *lin-14*.

**Prédiction:** Si la région de *lin-4* complémentaire du gène *lin-14* est épissée dans un gène rapporteur, celui-ci devrait entraîner la même régulation que *lin-14*.

**Test:** On utilise l'ADN recombinant pour construire deux versions d'un gène rapporteur ( $\beta$ -galactosidase). Dans les vers transgéniques (*C. elegans*), le gène rapporteur s'exprime en donnant une couleur bleue.

1. Le gène de la  $\beta$ -galactosidase avec la région 3' non traduite de *lin-14* contenant la région complémentaire de *lin-4*
2. Le gène de la  $\beta$ -galactosidase avec une région 3' non traduite témoin dépourvue de la région complémentaire de *lin-4*



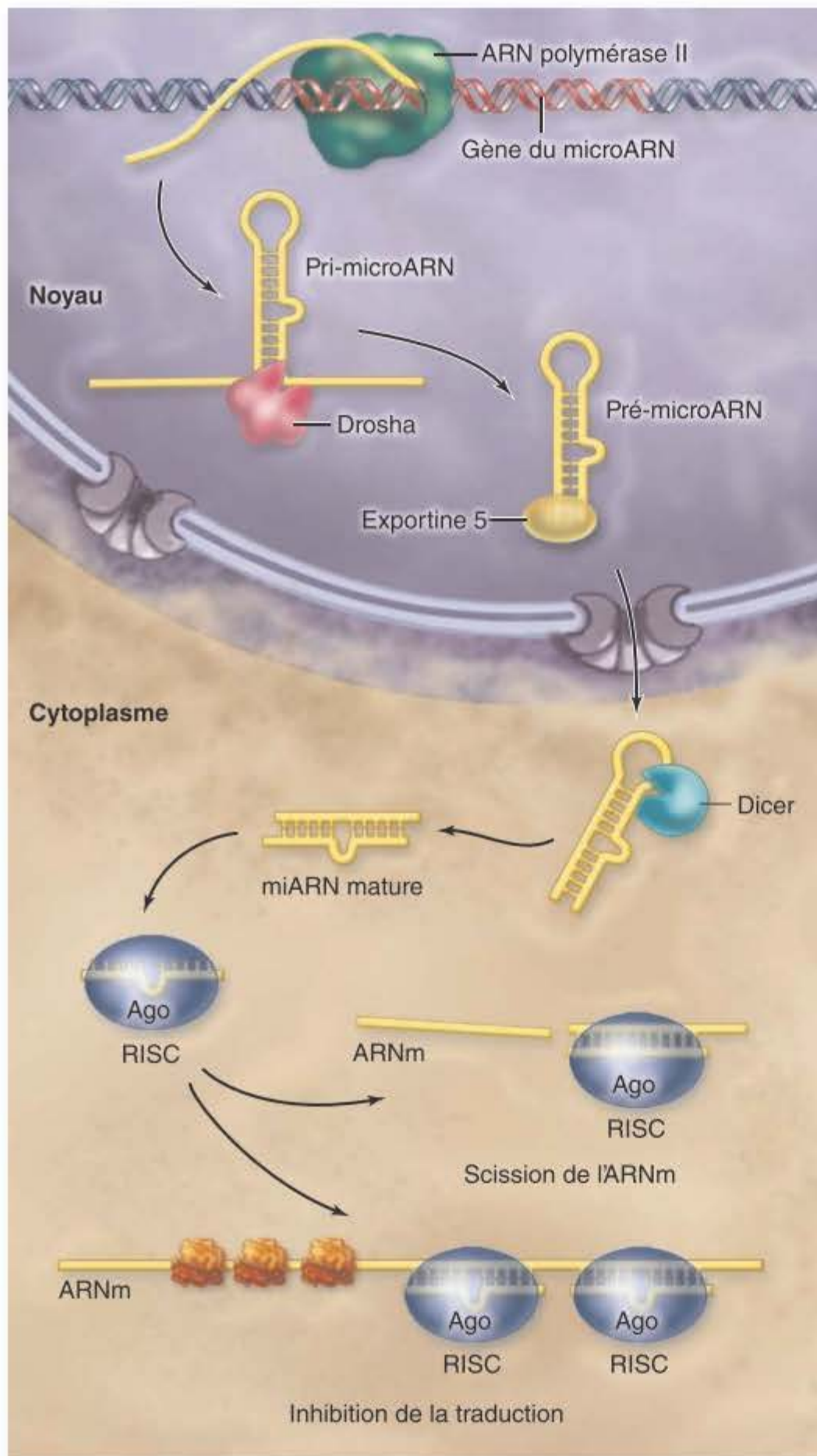
### Résultats:

1. Les vers transgéniques avec le gène rapporteur plus la région non traduite 3' de *lin-14* montrent une expression dans les larves L1, mais pas dans les L2. C'est ce qui est prévu pour le gène *lin-14*, contrôlé par *lin-4*.
2. Les vers transgéniques avec le gène rapporteur plus la région 3' témoin ne montrent pas l'expression attendue d'un contrôle par *lin-4*.

**Conclusion:** La région non traduite 3' de *lin-14* suffit pour empêcher l'expression génique dans les larves L2.

**Autres expériences:** Quel type d'expression prévoyez-vous pour ces constructions dans un mutant dépourvu de la fonction de *lin-4* ?

**Figure 16.15** Contrôle de l'expression du gène *lin-14*. Le gène *lin-14* est contrôlé par le gène *lin-4*. Ce contrôle est exercé par une région 3' non traduite de l'ARNm *lin-14* complémentaire du miARN *lin-4*.



**Figure 16.16 Biogenèse et fonctionnement du miARN.** Un gène de miARN est transcrit par l'ARN polymérase II pour donner un pri-miARN. Celui-ci est transformé par la nucléase Drosha en pré-miARN qui est exporté du noyau uni au facteur d'exportation Exportine 5. Dans le cytoplasme, le pré-miARN est modifié par la nucléase Dicer en miARN mature. Ce miARN est chargé dans un RISC qui peut scinder des ARNm cibles ou empêcher leur traduction.

inhibition semble plus fréquente que la scission de l'ARNm, mais le mécanisme précis de cette inhibition n'est pas encore élucidé. Chez les plantes, la scission de l'ARNm par RISC est fréquente et semble en rapport avec une complémentarité plus précise que dans les systèmes animaux entre les miARN végétaux et leurs cibles.

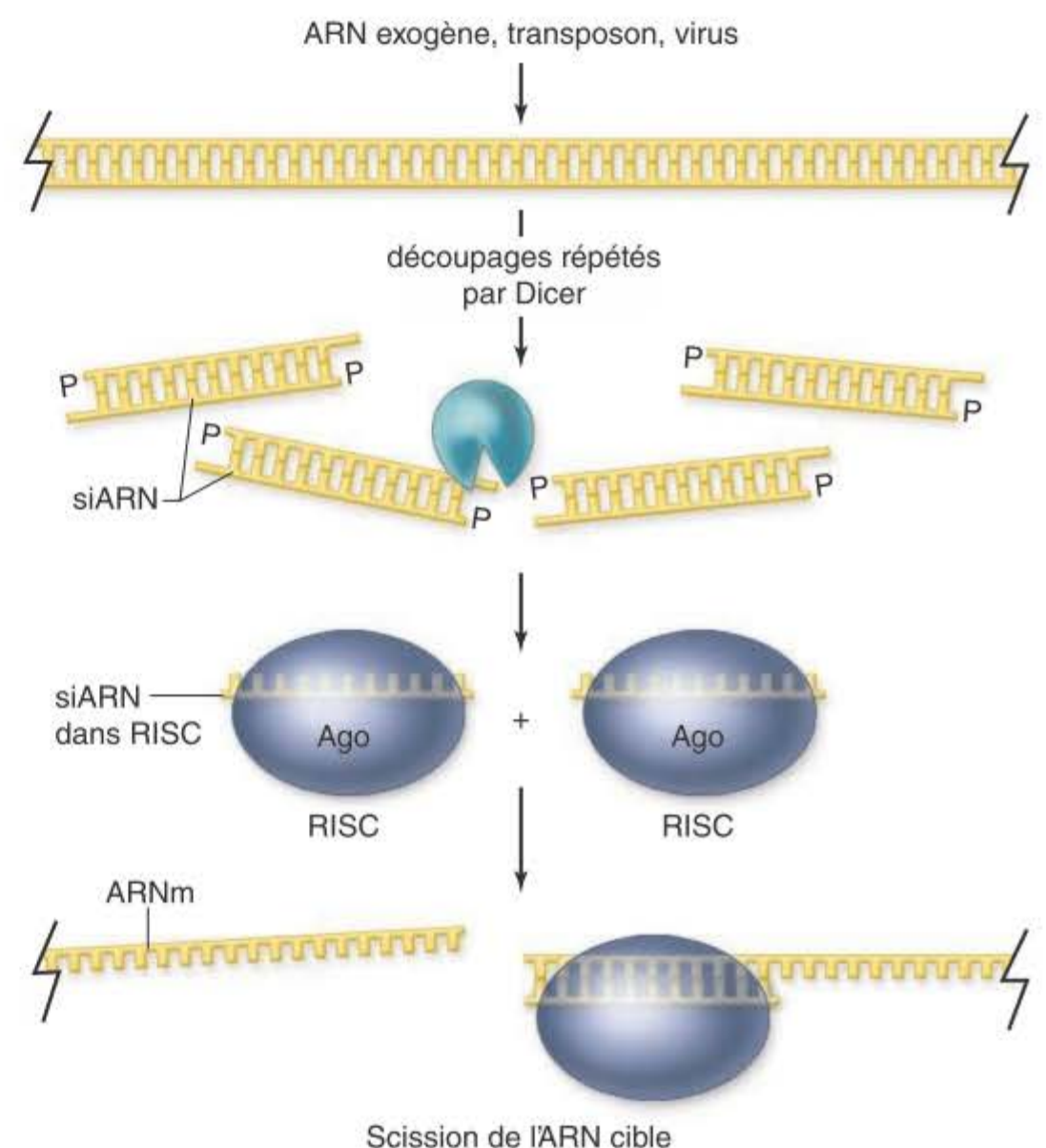
### Interférence de l'ARN

On connaît l'inhibition des gènes par les petits ARN depuis plusieurs années. Une certaine confusion provient de la nomenclature

utilisée parce que les observations réalisées sur des systèmes différents ont conduit à une profusion de termes. Cependant, l'interférence de l'ARN, la cosuppression et l'inhibition des gènes après la traduction découlent de mécanismes biochimiques semblables. Interférence de l'ARN est le terme le plus courant et implique la synthèse de siARN.

Cette synthèse est semblable à celle des miARN, sauf qu'ils proviennent d'un ARN bicaténaire (figure 16.17). Celui-ci peut être une très longue région complémentaire avec elle-même ou provenir de deux ARN complémentaires. Ces longs ARN bicaténaires sont modifiés par Dicer pour donner de nombreux siARN qui sont chargés dans un RISC contenant Ago. Les siARN sont en général presque parfaitement complémentaires de leurs ARNm cibles et le résultat est une scission de l'ARNm par le RISC contenant le siARN.

L'ARN bicaténaire à l'origine des siARN peut provenir de la cellule ou de l'extérieur. Dans la cellule même, des gènes peuvent produire des ARN avec de longues régions autocaténares qui se replient en un substrat pour Dicer dans le cytoplasme. Ils peuvent aussi provenir de régions répétées du génome contenant des éléments transposables. Les ARN bicaténaires exogènes peuvent être introduits expérimentalement ou par infection virale.



**Figure 16.17 Biogenèse et fonctionnement du siARN.** Les siARN peuvent provenir de différentes sources qui donnent toutes de longues régions d'ARN bicaténaire. L'ARN bicaténaire est modifié par la nucléase Dicer pour donner plusieurs siARN qui sont chargés chacun sur leur propre RISC. Le RISC coupe ensuite l'ARNm cible.

## Les petits ARN ont pu évoluer pour protéger le génome

Le fait que l'ARN viral peut être dégradé par le système de blocage de l'ARN peut donner des indications sur l'origine des petits ARN. Une autre observation du même ordre est la possibilité pour ce système de contrôler aussi l'action des transposons. Chez les souris et les drosophiles, il existe des preuves en faveur de l'implication du système de l'interférence de l'ARN dans la lignée germinale, où une classe spécifique de petits ARN paraît intervenir dans le blocage des transposons au cours de la spermatogenèse et de l'ovogenèse. Ce mécanisme peut donc dériver d'une voie ancienne destinée à protéger le génome d'attaques provenant de l'intérieur et de l'extérieur. La conservation des protéines clés fait penser que l'ancêtre de tous les eucaryotes possédait déjà une forme de voie de blocage par l'ARN.

### Distinction entre miARN et siARN

La biogenèse des miARN et des siARN implique une scission par Dicer et une incorporation dans un complexe RISC. La principale différence entre ces deux types de molécules est leurs cibles : les miARN ont tendance à réprimer des gènes différents de leur origine, alors que les siARN endogènes répriment plutôt des gènes dont ils dérivent. On utilise en outre les siARN expérimentalement pour bloquer l'expression des gènes. Cette technique utilise l'équipement cellulaire de blocage par l'ARN pour rendre un gène silencieux en introduisant un ARN bicaténaire complémentaire du gène.

Il y a d'autres différences entre les deux catégories de petits ARN. L'étude de nombreuses espèces a montré que les miARN ont en général été bien conservés au cours de l'évolution, ce qui n'est pas le cas des siARN. La biogenèse est semblable pour ce qui concerne les nucléases impliquées, mais la structure des ARN bicaténaires est différente. Le transcrite des gènes de miARN forme des structures en tige et boucle, alors que les ARN bicaténaires à l'origine des siARN peuvent être bimoléculaires ou formés de boucles très longues. Ces longues régions bicaténaires donnent de nombreux siARN, alors qu'un pré-miARN ne donne qu'un seul miARN.

## Les petits ARN peuvent intervenir dans la formation de l'hétérochromatine

On a aussi impliqué les systèmes d'inhibition par l'ARN dans la formation de l'hétérochromatine chez la levure scissipare, les plantes et la drosophile. Chez la levure scissipare, l'hétérochromatinisation du centromère est contrôlée par le siARN produit par la nucléase Dicer. Cette production d'hétérochromatine implique aussi une modification des histones et relie donc l'interférence par l'ARN aux complexes de remodelage de la chromatine dans ce système. On ne sait pas encore jusqu'à quel point ce phénomène est répandu.

Les plantes représentent un cas intéressant parce qu'elles possèdent des espèces différentes de petits ARN. Le système d'interférence par l'ARN est plus complexe que chez les animaux avec de nombreuses formes des protéines de la nucléase Dicer et des protéines de liaison à l'ARN Argonaute. Une catégorie de siARN endogène peut entraîner l'hétérochromatinisation par méthylation de l'ADN et modification des histones.

## L'épissage alternatif peut produire de nombreuses protéines à partir d'un seul gène

Les estimations les plus récentes de la fréquence de l'épissage alternatif citées au chapitre 15 montrent son importance. La signification fonctionnelle de ces données reste cependant peu claire. Nous allons envisager ici quelques exemples bien caractérisés.

L'épissage alternatif peut avoir des effets différents à des stades différents du développement ou dans des tissus différents. On trouve un exemple de différences dans le développement de la drosophile : la détermination du sexe découle d'une série complexe d'épissages alternatifs différents chez les mâles et les femelles.

On trouve un bel exemple d'épissage alternatif spécifique des tissus dans deux organes humains différents, la glande thyroïde et l'hypothalamus (voir chapitre 45). Ces deux organes produisent deux hormones différentes : la *calcitonine* et *PRGC* (peptide relié au gène de la calcitonine). La calcitonine contrôle l'accès au calcium et l'équilibre du calcium dans des tissus comme les os et les dents. PRGC intervient dans plusieurs fonctions neurales et endocrines. Bien que leur rôle physiologique soit très différent, ces deux hormones sont synthétisées à partir du même transcrite (figure 16.18).

La synthèse de l'un ou l'autre produit est déterminée par des facteurs spécifiques propres aux tissus qui contrôlent la maturation du transcrite primaire. Dans le cas de la calcitonine et du CGRP, l'épissage du préARNm est contrôlé par des facteurs différents présents dans la glande thyroïde et l'hypothalamus.

## L'édition de l'ARN modifie l'ARNm après la transcription

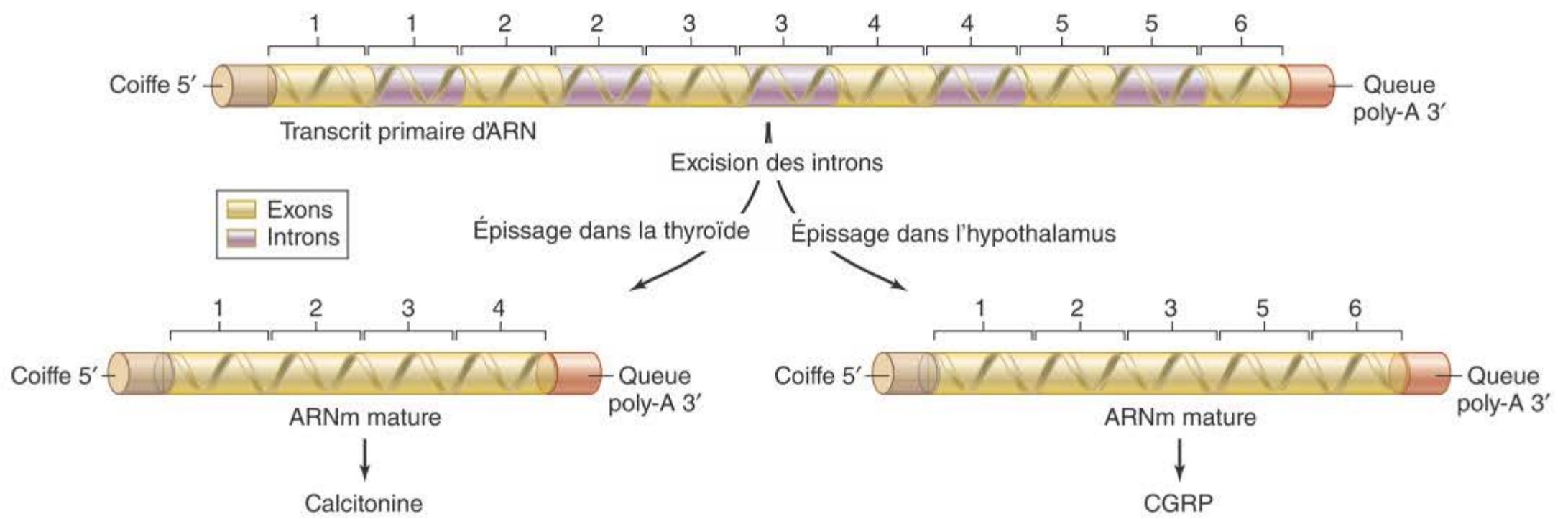
L'édition des transcrits d'ARNm matures produit parfois un ARNm modifié qui n'est pas réellement codé dans le génome – on ne s'y attendait pas. On a d'abord constaté l'insertion de résidus uracile dans certains transcrits d'ARN chez les protozoaires et l'on a cru à une anomalie.

Depuis lors, on a découvert un autre type d'édition de l'ARN dans des espèces de mammifères, y compris chez l'homme. Dans ce cas, l'édition implique l'altération chimique d'une base qui modifie ses propriétés d'appariement, généralement par désamination. On a, par exemple, observé la désamination de la cytosine en uracile et de l'adénine en inosine (l'inosine s'apparierait avec G pendant la traduction).

### Le récepteur de sérotonine 5-HT

On a aussi observé l'édition de l'ARN chez l'homme pour certains récepteurs du cerveau pour les opiacés. L'un d'eux, le récepteur de la sérotonine (5-HT) est édité à de nombreux endroits et donne au total 12 isoformes différentes de la protéine.

On ne connaît pas bien la fréquence de ces formes d'édition, mais il existe d'autres arguments montrant que l'information codée dans les gènes n'est pas le point final de la synthèse des protéines.



**Figure 16.18 L'épissage alternatif.** Beaucoup de transcrits primaires peuvent être épissés de plusieurs façons pour donner de nombreux ARNm. Dans cet exemple, le transcrit primaire est épissé dans la glande thyroïde en quatre exons codant une protéine, la calcitonine. Dans l'hypothalamus, le quatrième exon, contenant le site poly-A utilisé dans la thyroïde est sauté et deux autres exons sont ajoutés pour coder un peptide proche de la calcitonine (CGRP).

## L'ARNm doit sortir du noyau pour la traduction

Les transcrits d'ARNm transformés sortent du noyau par les pores (décrits au chapitre 4). La traversée de la membrane nucléaire par un transcrit est un processus actif qui exige sa reconnaissance par les récepteurs qui tapissent l'intérieur des pores. Des portions spécifiques du transcrit, comme la queue poly-A, paraissent jouer un rôle dans cette reconnaissance.

Il existe peu d'arguments rigoureux en faveur d'une régulation de l'expression génique à ce niveau, mais elle pourrait exister. En moyenne, environ 10 % des transcrits primaires sont formés d'exons qui donneront des séquences d'ARNm, mais environ 5 % seulement de tous les ARNm produits sous forme de transcrits primaires arrivent au cytoplasme. La moitié environ des exons des transcrits primaires ne quitteraient donc jamais le noyau, mais il n'est pas évident que la disparition de cet ARNm soit sélective.

## L'initiation de la traduction peut être contrôlée

La traduction, par les ribosomes du cytoplasme, d'un transcrit d'ARNm transformé, fait intervenir un complexe de protéines, les *facteurs de traduction*. Au moins dans certains cas, l'expression des gènes est contrôlée par la modification d'un ou plusieurs de ces facteurs. Dans d'autres cas, des **protéines réprimant la traduction** arrêtent la traduction en s'unissant au début du transcrit, qui ne peut plus s'attacher au ribosome.

Chez l'homme, la production de ferritine (protéine de stockage du fer) est normalement bloquée par un répresseur de la traduction, l'aconitase. L'aconitase s'unit à une séquence de 30 nucléotides au début de l'ARNm de la ferritine et forme une boucle stable à laquelle les ribosomes ne peuvent s'unir. Quand du fer pénètre dans la cellule, son union

à l'aconitase sépare celle-ci de l'ARNm de la ferritine qui est libéré, est traduit et produit 100 fois plus de ferritine.

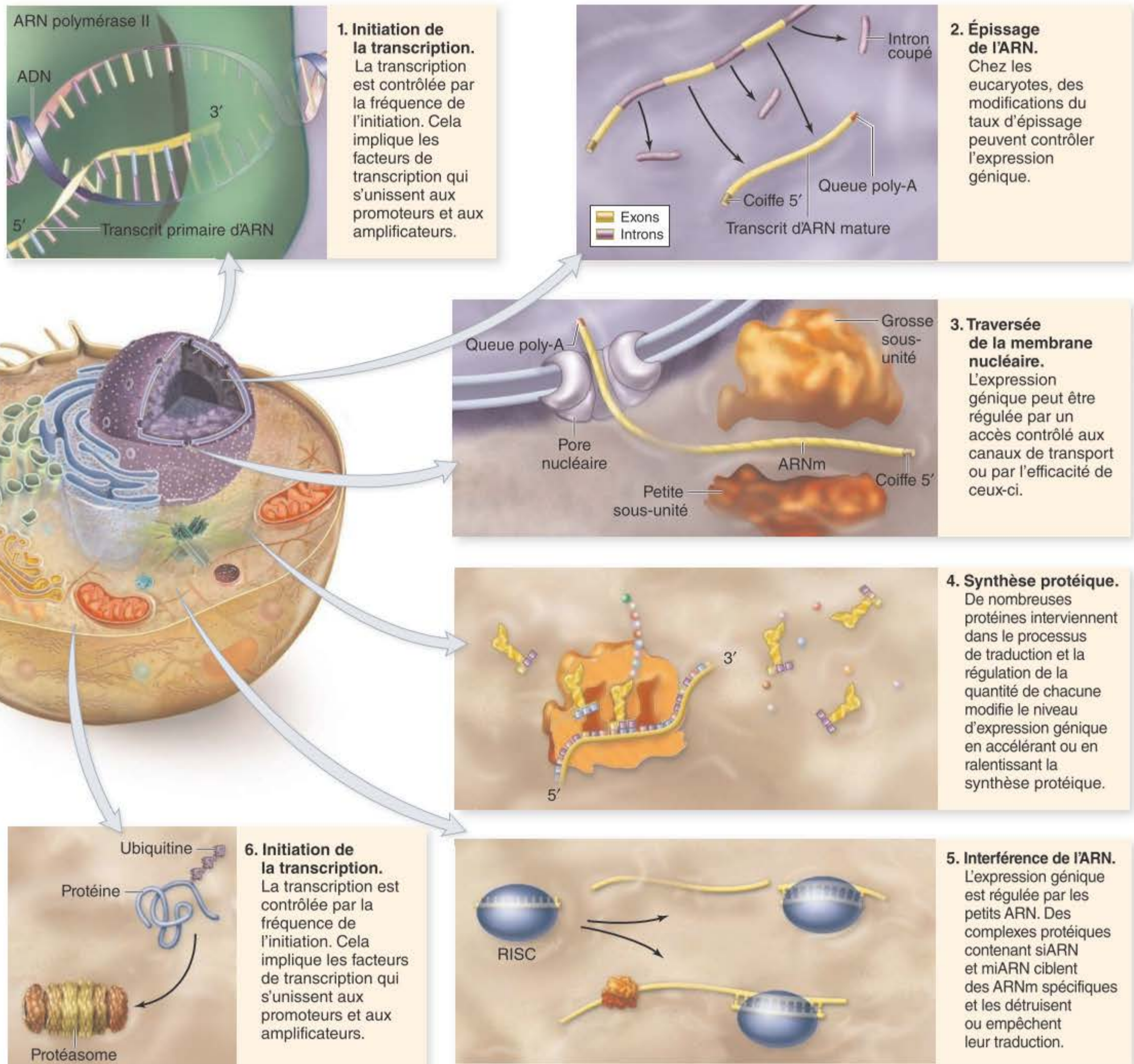
## La dégradation de l'ARNm est contrôlée

Un autre aspect des modifications de l'expression génique est la stabilité des transcrits d'ARNm dans le cytoplasme. Contrairement aux ARNm procaryotes, dont la demi-vie est normalement d'environ 3 minutes, les ARNm eucaryotes sont très stables. Par exemple, la demi-vie des transcrits du gène de la globine  $\beta$  dépasse 10 heures, une éternité dans la vie métabolique toujours changeante d'une cellule. Les transcrits codant les protéines de régulation et les facteurs de croissance sont cependant en général beaucoup moins stables, avec des demi-vies de moins d'une heure. Pour quelle raison ces transcrits particuliers sont-ils aussi instables ? Très souvent, ils contiennent des séquences spécifiques proches des extrémités 3' qui en font des cibles pour les enzymes qui dégradent les ARNm. Une séquence de nucléotides A et U proche de la queue poly-A d'un transcrit favorise l'élimination de la queue, qui déstabilise l'ARNm.

La perte de la queue poly-A entraîne une dégradation rapide par les exonucléases 3'-5'. Une autre conséquence de cette perte est la stimulation d'enzymes qui éliminent la coiffe 5' et aboutit à une dégradation par les exonucléases 5'-3'.

D'autres transcrits d'ARNm contiennent des séquences proches de leur extrémité 3' qui sont des sites de reconnaissance pour des endonucléases qui digèrent rapidement ces transcrits. La courte durée de vie des transcrits d'ARNm de nombreux gènes de régulation est vitale pour le fonctionnement de ces gènes parce qu'elle permet une modification rapide du taux des protéines régulatrices dans la cellule.

La figure 16.19 illustre les différents modes de contrôle de l'expression génique agissant après la transcription.



**Figure 16.19** Mécanismes du contrôle de l'expression génique des eucaryotes.

### Questions d'apprentissage 16.6

Les petits ARN contrôlent l'expression génique en dégradant sélectivement les ARNm, en inhibant la traduction ou en modifiant la structure de la chromatine. De multiples ARNm peuvent provenir d'un même gène par l'épissage alternatif parfois spécifique du tissu et du développement. L'édition de l'ARN peut aussi modifier la séquence d'un ARNm.

- Comment l'interférence de l'ARN pourrait-elle servir à la création d'un médicament ?

## 16.7 Dégradation des protéines

### Objectifs

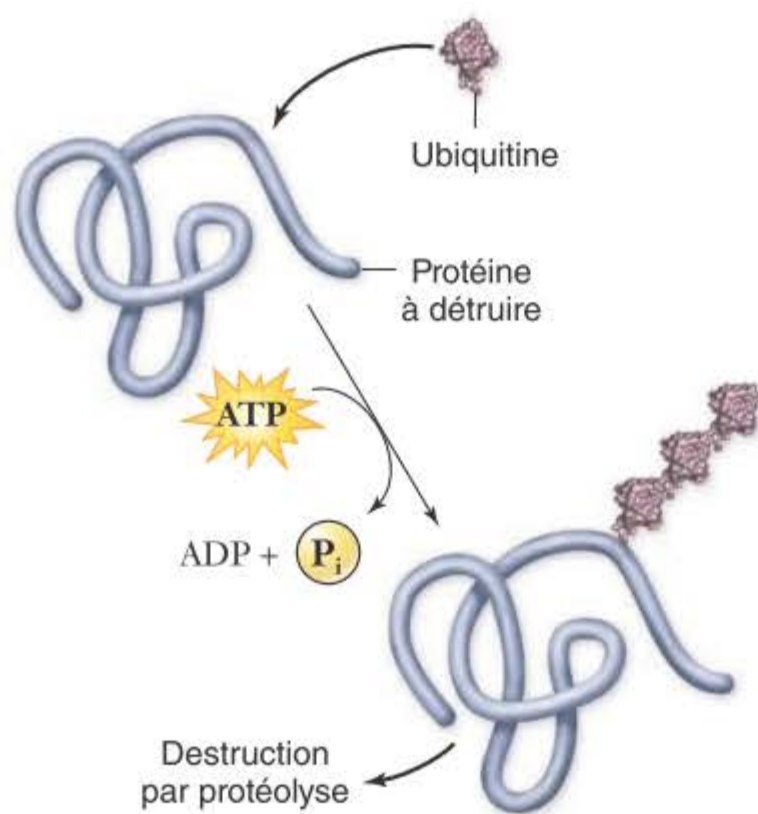
1. Décrire le rôle de l'ubiquitine dans la dégradation des protéines.
2. Expliquer le fonctionnement du protéasome.

La persistance de toutes les protéines produites par une cellule pendant sa durée de vie entraînerait de sérieux problèmes. Dans les années 1970, les recherches sur l'étiquetage des protéines ont montré que les cellules eucaryotes changent de protéines d'une manière contrôlée. C'est que des protéines sont constamment synthétisées et dégradées. Cette rotation des protéines n'est pas aussi rapide que chez les procaryotes, mais elle montre l'importance d'un système contrôlant leur renouvellement.

Les protéines peuvent être modifiées chimiquement et devenir non fonctionnelles ; en outre, la nécessité d'une protéine particulière peut être transitoire. Les protéines ne se replient pas non plus toujours correctement ou, avec le temps, elles peuvent être mal repliées. Ces modifications entraînent éventuellement la perte d'une fonction ou d'autres réactions chimiques, comme l'agglutination en complexes insolubles. En fait, plusieurs maladies neurodégénératives, comme l'Alzheimer, la maladie de Parkinson et la maladie de la vache folle, proviennent d'une agglutination de protéines formant des plaques caractéristiques dans les cellules du cerveau. Outre un renouvellement normal des protéines, les cellules ont besoin d'un mécanisme qui les débarrasse des protéines âgées, inutilisées ou incorrectement repliées.

Des enzymes, les **protéases**, peuvent dégrader les protéines par rupture des liaisons peptidiques et transformation en acides aminés. Ces enzymes sont évidemment nécessaires, mais elles ne peuvent certainement pas rester actives tout le temps dans le cytoplasme.

Les cellules eucaryotes disposent d'un moyen pour résoudre ces problèmes : c'est le confinement des enzymes de dégradation dans un compartiment cellulaire particulier. Rappelez-vous (chapitre 4) que les lysosomes sont des vésicules contenant des enzymes de digestion, y compris des protéases. Les lysosomes servent à éliminer les protéines et les organites âgés ou non fonctionnels, mais ce système n'est pas spécifique pour des protéines particulières. Les cellules ont besoin d'un autre système contrôlé pour éliminer les protéines âgées ou inutilisées et laisser intactes les autres protéines cellulaires.



**Figure 16.20** Marquage des protéines par l'ubiquitine. Les protéines à dégrader sont marquées à l'ubiquitine. Une enzyme, l'ubiquitine ligase, utilise l'ATP pour ajouter l'ubiquitine à une protéine. Quand une série d'ubiquitines a été ajoutée, la protéine marquée est détruite.

## L'addition d'ubiquitine marque les protéines à détruire

Les cellules eucaryotes résolvent ce problème en marquant les protéines à détruire, et en les dégradant ensuite. Le marquage utilisé par les cellules est la fixation d'une molécule d'**ubiquitine**. L'ubiquitine doit son nom à sa présence dans la plupart des cellules eucaryotes (elle est ubiquiste) ; c'est une protéine de 76 acides aminés isolée ou formant de longues chaînes attachées à d'autres protéines.

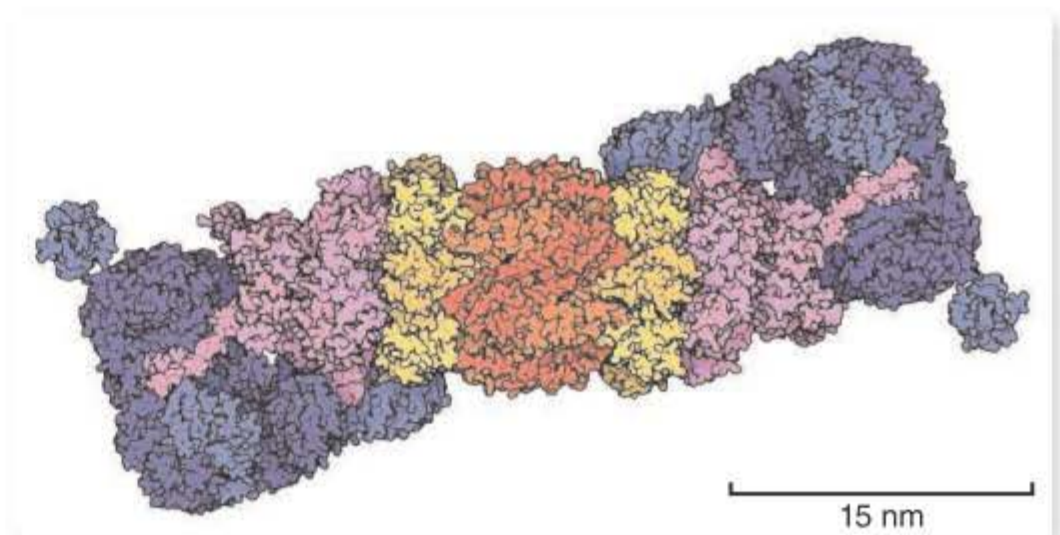
Une enzyme, l'ubiquitine ligase, fixe les longues chaînes aux protéines par étapes (figure 16.20). Cette réaction a besoin d'ATP et d'autres protéines ; ce processus comporte de nombreuses étapes et une régulation. Pour les protéines, la fixation d'une chaîne de polyubiquitines est un signal cellulaire signifiant la destruction de la protéine.

Deux grandes catégories de protéines sont marquées à l'ubiquitine : celles qui doivent être éliminées parce qu'elles sont mal repliées ou non fonctionnelles et celles qui sont produites et détruites par la cellule de façon contrôlée. Un exemple de ces dernières est celui des cyclines, protéines qui interviennent dans le contrôle du cycle cellulaire (chapitre 10). Quand elles ont rempli leur rôle dans la division active de la cellule, ces protéines sont marquées à l'ubiquitine et éliminées. De cette façon, la cellule peut contrôler son entrée en division ou ne pas se diviser.

## Le protéasome dégrade les protéines marquées par la polyubiquitine

Le **protéasome** est l'organite cellulaire qui dégrade les protéines marquées à l'ubiquitine ; c'est un grand complexe cylindrique dans lequel les protéines entrent d'un côté et sortent de l'autre sous la forme d'acides aminés ou de fragments peptidiques (figure 16.21).

Le complexe du protéasome comporte une région centrale qui a une activité de protéase, avec des éléments de régulation aux deux bouts. Bien que dépourvu de membrane, on peut considérer que cet organite représente une forme de compartimentation à très petite échelle. Par un mécanisme à deux étapes, d'abord un marquage des protéines à détruire, puis le passage par un grand complexe, les protéines à dégrader sont isolées du reste du cytoplasme.



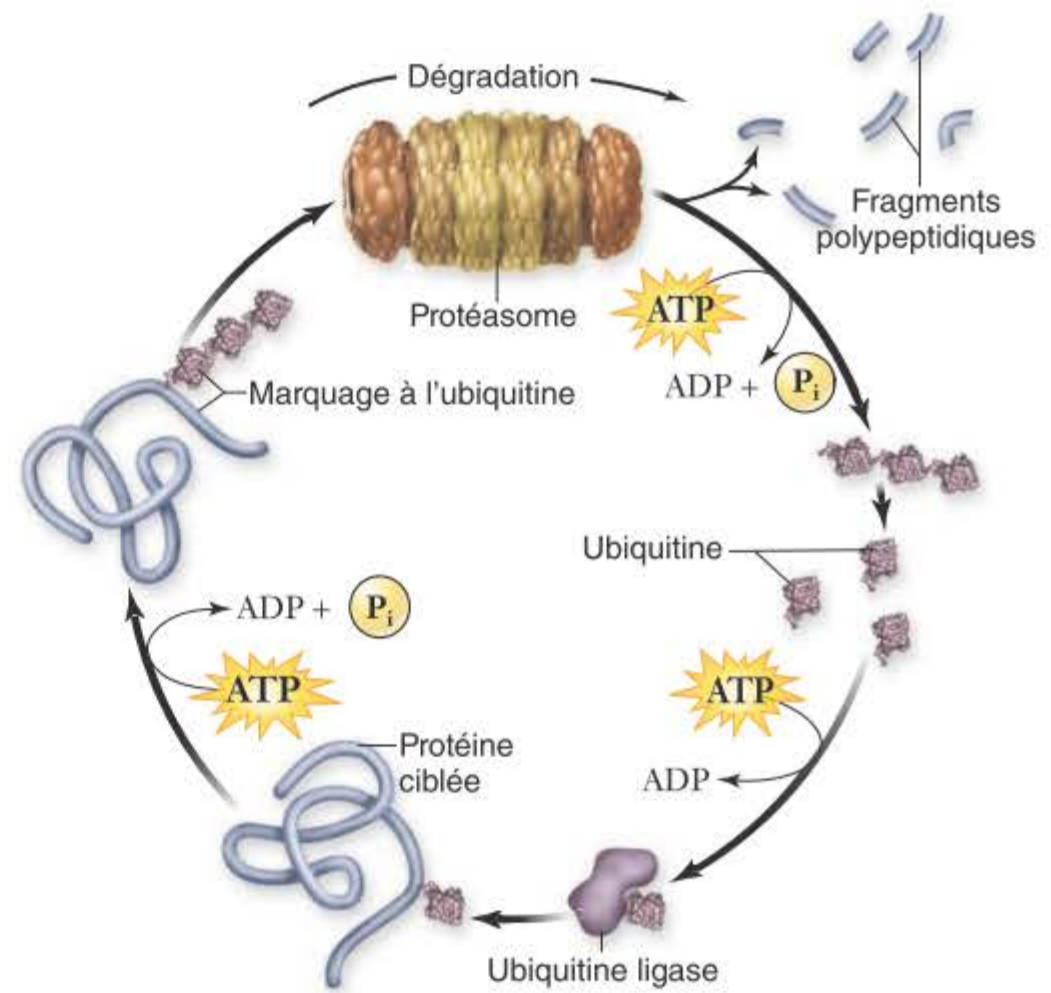
**Figure 16.21** Le protéasome de la drosophile. Le complexe central a une activité protéolytique et les extrémités fonctionnent comme régulateurs. Les protéines entrent par un bout du cylindre et sont coupées en fragments peptidiques sortant par l'autre bout.

Le marquage par l'ubiquitine suivie de la dégradation par le protéasome est la *voie ubiquitine-protéasome*. On peut la considérer comme un cycle dans lequel l'ubiquitine fixée aux protéines n'est pas elle-même détruite dans le protéasome. Quand les protéines sont dégradées, la chaîne d'ubiquitine elle-même est simplement découpée en ubiquitines séparées réutilisables (figure 16.22).

### Questions d'apprentissage 16.7

Le contrôle de la dégradation des protéines chez les eucaryotes implique l'addition d'une protéine, l'ubiquitine, qui marque la protéine à détruire. Le protéasome, complexe cylindrique dont le centre a une activité de protéase, identifie les protéines marquées et les découpe, un peu comme un broyeur détruit des documents. L'ubiquitine est recyclée sans modification.

- Quelle conséquence aurait, pour la cellule, un marquage par l'ubiquitine mal contrôlé ?



**Figure 16.22** Dégradation par la voie ubiquitine-protéasome. Les protéines sont d'abord marquées à l'ubiquitine, puis elles entrent dans le protéasome pour y être dégradées. Dans le protéasome, la polyubiquitine est enlevée, puis dépolymérisée en molécules simples d'ubiquitine pouvant être réutilisées.

**?** **Question** Quelles sont les deux raisons justifiant le marquage par la cellule d'un polypeptide par l'ubiquitine ?

## Résumé

### 16.1 Contrôle de l'expression génique

Le contrôle s'effectue généralement au niveau de l'initiation de la transcription.

La transcription est contrôlée par des protéines de régulation qui ajustent la faculté de liaison de l'ARN polymérase au promoteur. La transcription peut ainsi être empêchée ou stimulée.

Chez les procaryotes, les stratégies de contrôle sont mises en route pour s'adapter aux modifications de l'environnement.

Chez les eucaryotes, les stratégies de contrôle sont destinées à maintenir l'homéostasie.

### 16.2 Les protéines de régulation

Les protéines peuvent interagir avec l'ADN par le sillon majeur.

La double hélice d'ADN forme un sillon majeur et un mineur ; les bases du sillon majeur sont accessibles aux protéines de régulation.

Les domaines de liaison à l'ADN interagissent avec des séquences d'ADN spécifiques.

Un motif de liaison à l'ADN est une région de la protéine de régulation capable de s'unir à l'ADN (figure 16.2).

Plusieurs motifs communs de liaison à l'ADN se retrouvent dans beaucoup de protéines.

Les motifs communs sont le motif hélice-boucle-hélice, le motif homéodomaine, le motif à doigt de zinc et la glissière à leucines.

### 16.3 La régulation procaryote

Le contrôle de la transcription peut être positif ou négatif.

Le contrôle négatif est assuré par des protéines, des répresseurs, qui interfèrent avec la transcription. Le contrôle positif est assuré par une catégorie de protéines de régulation, les activateurs, qui stimulent la transcription.

Les procaryotes adaptent l'expression génique en réaction aux conditions environnementales.

L'opéron *lac* est induit par la présence de lactose : les enzymes qui utilisent le lactose ne sont produites qu'en présence de lactose. L'opéron *trp* est réprimé : les enzymes nécessaires à la synthèse du tryptophane sont bloquées en présence de tryptophane.

L'opéron *lac* est contrôlé négativement par le répresseur *lac*.

L'opéron *lac* est induit quand l'effecteur (l'allolactose) s'unit au répresseur et modifie sa conformation, en sorte qu'il ne s'unit plus à l'ADN (figure 16.4).

La présence de glucose empêche l'induction de l'opéron *lac*.

L'expression maximale de l'opéron *lac* exige un contrôle positif par une protéine activatrice du catabolisme (CAP) formant un complexe avec l'AMPc. Quand le niveau du glucose est faible, celui de l'AMPc est élevé. La répression par le glucose implique l'exclusion de l'inducteur, empêchant l'entrée du glucose dans la cellule, et le contrôle de la CAP par le taux de glucose.

L'opéron *trp* est contrôlé par le répresseur *trp*.

L'opéron *trp* est réprimé quand le tryptophane, fonctionnant comme corépresseur, s'unit au répresseur, modifie sa conformation de façon telle qu'il ne peut s'unir à l'ADN et bloque l'opéron. L'expression est ainsi empêchée en présence d'un excès de *trp*.

## 16.4 La régulation eucaryote

*Les facteurs de transcription peuvent être généraux ou spécifiques.*

Les facteurs généraux de transcription sont nécessaires à l'assemblage de l'appareillage de transcription et au recrutement de l'ARN polymérase II au niveau du promoteur. Les facteurs spécifiques fonctionnent en fonction du tissu ou du moment pour stimuler des taux plus élevés de transcription.

*Les promoteurs et les amplificateurs sont des sites de liaison pour les facteurs de transcription.*

Les facteurs généraux s'unissent au promoteur pour recruter l'ARN polymérase. Les facteurs spécifiques s'unissent aux amplificateurs, qui peuvent être éloignés du promoteur, mais sont rapprochés par des boucles de l'ADN.

*Les coactivateurs et les médiateurs unissent les facteurs de transcription à l'ARN polymérase II (figure 16.12).*

Certains facteurs de transcription, mais pas tous, ont besoin d'un médiateur. Le nombre de coactivateurs est faible parce qu'un même coactivateur peut servir à beaucoup de facteurs de transcription.

*Le complexe de transcription rassemble tout.*

## 16.5 Structure de la chromatine et expression génique

L'emballage de l'ADN dans les nucléosomes complique la transcription. Les cellules peuvent altérer la structure de la chromatine en modifiant l'ADN ou les histones. On considère que c'est la base de l'épigénétique.

*La méthylation de l'ADN, la modification des histones et l'ARN non codant peuvent affecter la structure de la chromatine.*

La méthylation des bases de l'ADN, principalement de la cytosine, est liée aux gènes qui ont été « mis hors circuit ». L'inactivation du chromosome X implique un ARN non codant et un complexe répresseur.

*Certains activateurs de la transcription modifient la structure de la chromatine.*

L'acétylation des histones entraîne l'inactivation des régions de la chromatine.

*Les complexes de remodelage de la chromatine modifient également la structure de la chromatine.*

Les complexes de remodelage de la chromatine contiennent des enzymes qui éliminent, repositionnent et transfèrent les nucléosomes.

## 16.6 Contrôle post-transcriptionnel des eucaryotes

*Les petits ARN contrôlent l'expression génique après la transcription.*

L'interférence par l'ARN est due à des siARN produits par scission d'ARN bicaténaire par la nucléase Dicer. Uni à une protéine Argonaute, le siARN forme un complexe (RISC) qui réduit l'ARN au silence. Le RISC peut scinder l'ARNm ou empêcher la traduction. Un autre type de petit ARN, le miARN, est formé par l'action de deux nucléases, Drosha et Dicer, à partir de structures d'ARN en tige et boucle. Il forme également un RISC capable de dégrader l'ARNm ou d'arrêter la traduction.

*Les petits ARN ont pu évoluer pour protéger le génome.*

Les ARN viraux sont dégradés et les transposons sont réduits au silence dans la lignée germinale par interférence par l'ARN. Ce mécanisme a une origine ancienne.

*Les petits ARN peuvent intervenir dans la formation de l'hétérochromatine.*

Dans la levure scissipare, la drosophile et les plantes, les voies d'interférence aboutissent à la formation de l'hétérochromatine.

*L'épissage alternatif peut produire de nombreuses protéines à partir d'un seul gène.*

En réponse à des facteurs spécifiques aux tissus, l'épissage alternatif du pré-ARNm d'un gène peut produire de multiples protéines.

*L'édition de l'ARN modifie l'ARNm après la transcription.*

*L'ARNm doit sortir du noyau pour la traduction.*

Les facteurs de traduction peuvent être modifiés pour contrôler l'initiation ; les protéines répresseurs de la traduction peuvent s'unir au début du transcrit, qui ne peut plus s'attacher au ribosome.

*La dégradation de l'ARNm est contrôlée.*

Le transcrit d'ARNm est relativement stable, mais il peut porter des cibles pour des enzymes qui le dégradent rapidement si c'est nécessaire pour la cellule.

## 16.7 Dégradation des protéines

*L'addition d'ubiquitine marque les protéines à détruire.*

Chez les eucaryotes, les protéines marquées pour être détruites ont reçu de l'ubiquitine comme marqueur.

*Le protéasome dégrade les protéines marquées par la polyubiquitine.*

Un organite cellulaire – un protéasome cylindrique – dégrade les protéines marquées à l'ubiquitine qui le traversent.



## Questions

### COMPRÉHENSION

- Chez les procaryotes, l'expression génique est généralement contrôlée à
  - l'épissage du pré-ARNm en ARNm mature.
  - l'initiation de la traduction.
  - l'initiation de la transcription.
  - tous ces stades.
- Les protéines de régulation interagissent avec l'ADN par
  - déroulement de l'hélice et modification du mode d'appariement des bases.
  - union à la colonne vertébrale sucre-phosphate de la double hélice.
  - déroulement de l'hélice et rupture des paires de bases.
  - union au sillon majeur de la double hélice et interaction avec les paires de bases.

3. Chez *E. coli*, l'induction de l'opéron *lac* et la répression de l'opéron *trp* sont deux exemples de
  - a. contrôle négatif par un répresseur.
  - b. contrôle positif par un répresseur.
  - c. contrôle négatif par un activateur.
  - d. contrôle positif par un activateur.
4. L'opéron *lac* est contrôlé par deux protéines principales. Ces protéines
  - a. agissent toutes deux négativement.
  - b. agissent toutes deux positivement.
  - c. agissent de manière opposée, l'une négativement, l'autre positivement.
  - d. agissent au niveau de la traduction.
5. Chez les eucaryotes, l'union de l'ARN polymérase à un promoteur demande l'intervention de
  - a. facteurs spécifiques de transcription.
  - b. facteurs généraux de transcription.
  - c. protéines de répression.
  - d. protéines d'induction.
6. Chez les eucaryotes, la régulation de l'expression génique se situe
  - a. uniquement au niveau de la transcription.
  - b. uniquement au niveau de la traduction.
  - c. à l'initiation de la transcription ou après la transcription.
  - d. uniquement après la transcription.
7. Dans l'opéron *trp*, le répresseur s'unit à l'ADN
  - a. en l'absence de tryptophane.
  - b. en présence de tryptophane
  - c. en présence ou en l'absence de tryptophane.
  - d. seulement quand la cellule a besoin de tryptophane.
4. La régulation par les petits ARN et l'épissage alternatif sont semblables du fait que tous deux
  - a. agissent après la transcription.
  - b. agissent par l'intermédiaire de complexes ARN/protéines.
  - c. contrôlent l'équipement de transcription.
  - d. a et b sont corrects.
5. Les ARNm eucaryotes diffèrent des ARNm procaryotes parce qu'ils
  - a. contiennent généralement plusieurs gènes.
  - b. sont colinéaires par rapport aux gènes qui les codent.
  - c. ne sont pas colinéaires par rapport aux gènes qui les codent.
  - d. a et c sont corrects.
6. Dans le cycle cellulaire, la production des protéines de la cycline est en synchronisme avec le cycle. Cela suppose vraisemblablement
  - a. un contrôle de l'initiation de la transcription des gènes de la cycline et le marquage des protéines de la cycline par l'ubiquitine.
  - b. l'épissage alternatif des gènes de cycline pour donner des protéines différentes.
  - c. l'édition de l'ARN donnant les différentes protéines de cycline.
  - d. le couplage de la transcription et de la traduction.
7. Un mécanisme de contrôle chez *E. coli* non présenté dans ce chapitre implique un arrêt des ribosomes permettant la production d'un terminateur de la transcription dans l'ARNm. Chez la levure scissipare, un eucaryote, ce mécanisme
  - a. serait commun puisqu'il s'agit d'un unicellulaire.
  - b. ne serait pas commun, puisqu'il s'agit d'un unicellulaire.
  - c. n'existerait pas parce que la transcription se déroule dans le noyau et la traduction dans le cytoplasme.
  - d. n'existerait pas en raison de la possibilité de l'épissage alternatif.

## APPLICATION

1. Le répresseur de *lac*, le répresseur de *trp* et CAP sont tous
  - a. des régulateurs négatifs de la transcription.
  - b. des régulateurs positifs de la transcription.
  - c. des protéines allostériques s'unissant à l'ADN et à un effecteur.
  - d. des protéines capables de s'unir à l'ADN et à d'autres protéines.
2. Les facteurs spécifiques de transcription des eucaryotes interagissent avec des amplificateurs pouvant se trouver loin du promoteur. Ces facteurs de transcription
  - a. modifient alors la structure de l'ADN entre l'amplificateur et le promoteur.
  - b. n'interagissent pas avec le système de transcription.
  - c. peuvent interagir avec le système de transcription grâce à une boucle dans l'ADN.
  - d. peuvent interagir avec le système de transcription en éliminant l'ADN intermédiaire.
3. La répression de l'opéron *trp* et l'induction de l'opéron *lac* sont deux mécanismes qui
  - a. ne seraient possibles qu'avec une régulation positive.
  - b. permettent à la cellule de contrôler la quantité d'enzymes adaptée aux conditions environnementales.
  - c. ne seraient possibles qu'avec une régulation négative.
  - d. font que la cellule fabrique tout le temps les enzymes de ces deux opérons.

## RÉVISION

1. Vous avez isolé une série de mutants affectant la régulation de l'opéron *lac*. Ils sont tous constitutifs, c'est-à-dire que l'opéron *lac* s'exprime tout le temps. Vous disposez aussi des allèles mutant et de type sauvage pour chaque mutant dans toutes les combinaisons et sur des plasmides F' qui peuvent être introduits dans les cellules afin de les rendre diploïdes pour ces gènes. Comment utiliseriez-vous ces outils pour savoir quels mutants modifient les sites de liaison à l'ADN et ceux qui modifient les protéines qui s'unissent à l'ADN ?
2. On peut trouver des exemples de contrôle positif et négatif de la transcription dans la régulation de l'expression des opérons bactériens *lac* et *trp*. Utilisez ces deux systèmes d'opérons pour décrire la différence entre régulation positive et négative.
3. Quelles formes de contrôle de l'expression génique n'existent que chez les eucaryotes ? Les procaryotes pourraient-ils utiliser ces mécanismes ou proviennent-ils des différences entre ces types de cellules ?
4. Le nombre et le type de protéines trouvées dans une cellule peuvent dépendre des mutations géniques et de la régulation de l'expression génique. Donnez les différences entre ces deux processus.