

I. Le dogme de la biologie moléculaire

ADN → Transcription → ARN → Traduction → Protéines

- L'ADN (génomique) concerne le **génotype**.
- L'ARN (transcriptome) et les protéines (protéome) concernent le phénotype (=ce que l'on peut observer d'un organisme).
- Ce **phénotype** dépend du **génotype** et de l'**environnement**.

Il faut savoir qu'il y a une **amplification du signal** (un ADN → transcrit en beaucoup d'ARN → chacun traduit en beaucoup de protéines). Grâce à des techniques d'analyses à haut débit et de nanotechnologie, nous sommes aujourd'hui capables d'identifier et d'analyser entièrement le **génomique**, le **transcriptome** et la **protéome** (sciences des -omiques).

En fait, ce n'est pas que l'ADN qui va être transcrit puis traduit, c'est la **chromatine** (l'ADN + des protéines qui interagissent avec lui). C'est d'ailleurs cette même chromatine qui sera répliquée (lors de la mitose, l'organisation de la chromatine est conservée). L'information du **génotype** présent dans nos cellules n'est pas entièrement déterminée par la séquence d'ADN, et donc le **génotype**, mais également par **sa structure et son organisation** dans le noyau. A la notion de **génomique** se superpose une notion plus exacte d'**épigénomique**.

Epigénomique → la séquence de nos gènes dans un certain **environnement** qui correspond à la structure de la chromatine pris au sens large, c'est cet ensemble de l'ADN et des protéines qui interagissent avec lui pour le structurer que l'on transmettra aux cellules filles !

On a environ **20 000 gènes** dans nos cellules, mais ils ne sont **pas tous exprimés** en même temps (on dit qu'un gène s'exprime quand il est transcrit en ARN) : on exprime uniquement les gènes dont on a besoin (par exemple on n'a pas besoin de synthétiser d'hémoglobine dans une cellule de rétine).

- Quand un gène s'exprime, est transcrit → il est dit ON, celui-ci pourra être transcrit, subir un épissage, être traduit en protéine...
- Les autres qui ne s'expriment pas sont dits OFF → Ils n'aboutiront en aucun cas à la synthèse de protéines dans ces cellules.

Dans les différentes cellules de notre corps, on retrouvera le même génome, le même ADN, donc les mêmes gènes, cependant entre 2 cellules différentes on aura de grandes variations entre les gènes ON et les gènes OFF.

Dans nos cellules adultes, **la plupart des gènes sont OFF** : 10% seulement des gènes s'expriment dans les différents types cellulaires (les gènes qui s'expriment sont différents en fonction de l'organe considéré).

II. Programme transcriptionnel

Le programme transcriptionnel détermine qu'un gène soit ON ou OFF. Cela est permis par la cellule qui reçoit des **signaux externes** (hormones, molécules de signalisation, facteurs de croissance...), des **signaux physiques** (lumière, chaleur, signaux mécaniques...) ou encore des **signaux internes** (nombres de divisions, rythmes biologiques). Ces signaux seront reconnus par la cellule et interprétés.

→ Exemple d'expérience : expression d'un facteur de transcription myogénique chez un fibroblaste :

On prend un fibroblaste sur lequel on active un gène qui est normalement OFF (par exemple ici, un gène caractéristique des cellules musculaires), on peut transformer ce fibroblaste en cellule musculaire (c'est ce qu'on appelle une **transdifférenciation** et cela arrive aussi (rarement) de manière physiologique). Ces programmes transcriptionnels sont essentiels dans les phénomènes de différenciation.

→ Exemple d'expérience : les progéniteurs myéloïdes :

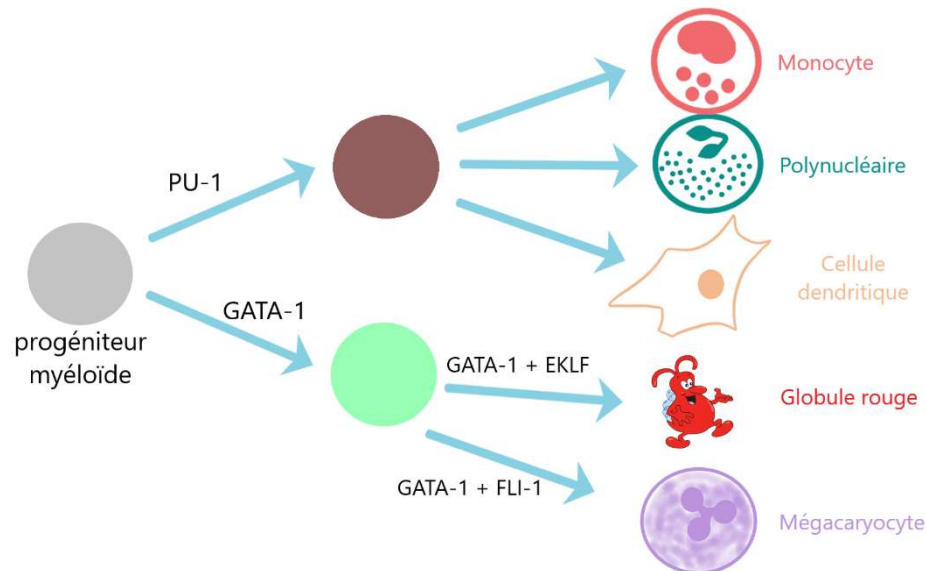
Ça se déroule dans la moelle osseuse tous les jours pour générer les lignées de cellules sanguines. A partir d'un progéniteur myéloïde, on a deux grands types de différenciation possibles : **la voie monocyttaire** et celle qui aboutit au **globule rouge**. Les deux voies débutent par la même cellule : la cellule progénitrice qui décide **selon les facteurs de transcription spécifique qu'elle reçoit**.

→ **PU1** entraîne la lignée vers la voie monocyttaire et en même temps **inhibe les gènes de la voie de différenciation des globules rouges**.

→ **GATA-1** entraîne la lignée vers la différenciation en globule rouge et en même temps inhibe les gènes de différenciation de la lignée monocyttaire.

Les cellules ayant exprimé **GATA-1** doivent alors **exprimer ensuite EKLf** pour devenir des Globules Rouges. Donc pour devenir des Globules Rouge, le précurseur doit avoir dans sa mémoire l'expression de **GATA-1**, et doit exprimer en plus **EKLf**. Donc pour obtenir un globule rouge, il faut que chacune des cellules intermédiaires ait eu une certaine histoire qui la conduise vers la différenciation en globule rouge.

Cellule progénitrice → exprime **GATA-1** → exprime **EKLf** → globule rouge



A- Mécanisme de régulation de l'expression des gènes

a) La compaction de l'ADN

La régulation va principalement venir de la conformation et de l'organisation structurale des gènes, un gène OFF aura une chromatine compactée / fermée (l'appareillage de transcription n'aura pas la place de se fixer pour agir) alors qu'un gène ON aura une chromatine ouverte (accessible à l'ARN polymérase pour la transcription). La structure de l'ADN persistera après la disparition du signal, la cellule gardera donc en mémoire l'expression des gènes (le signal entraîne la condensation de l'ADN, même lorsqu'on n'a plus de signal, la chromatine reste condensée, *c'est comme quand ta mère te dit de plier ton linge, elle c'est le signal, toi, t'es la cellule, tu plies ton linge, et ton linge c'est l'ADN et une fois que tu l'as plié il le reste*)

b) Le contrôle proximal de la transcription

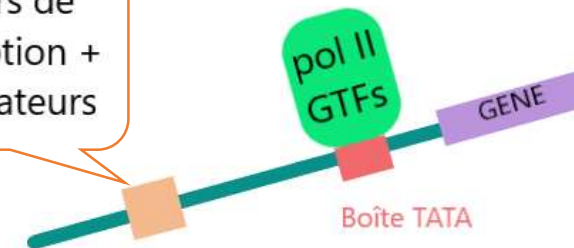
La transcription prend naissance au niveau du **promoteur** (=séquence d'ADN en amont du gène qui permet l'arrivée de l'ARN polymérase au niveau du gène). Ce promoteur est formé de la boîte TATA qui reconnaît le complexe d'initiation. Ce complexe d'initiation est composé l'ARN polymérase II (elle-même composée de 12 polypeptides) et des facteurs généraux de transcription (GTFs) [qui regroupent 23 polypeptides (TFIIA/B/C... pas à connaître par cœur)].

Complexe d'initiation :
ARN polymérase II (ou polII) : 12 polypeptides
+ Facteurs généraux de transcription (GTFs) :
23 polypeptides formant les facteurs
TFIIA/B/D/E/F/G/H/J

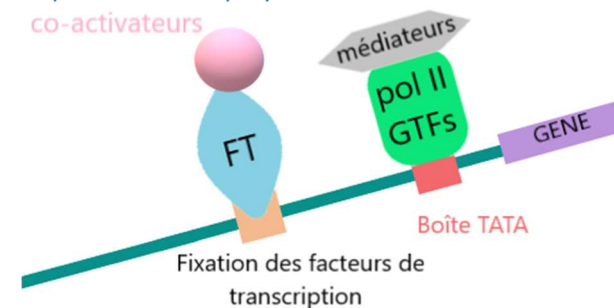


En mettant ce **complexe d'initiation** dans un tube à essai avec un **gène**, des chercheurs ont remarqué que la **polymérase** et les **facteurs** se **fixaient** mais qu'il n'y avait **pas de transcription**. En fait la **fixation n'est pas stable**. Ils ont alors trouvé qu'il existait d'autres protéines (= facteurs de transcription) qui se fixent ailleurs que sur le promoteur, et qui recrutent encore d'autres protéines = **les coactivateurs**.

Facteurs de transcription + co-activateurs



En mettant dans un tube à essai le **gène**, le **complexe d'initiation**, la **pol II** et les **GTFs** ainsi que les **coactivateurs** ils n'ont toujours **pas eu de transcription** (ou du moins très peu), donc il manque encore quelque-chose. Ils ont donc découvert un autre complexe = **le complexe médiateur**. Le complexe médiateur s'associe au complexe d'initiation (ARN polymérase + GTFs) et sert d'intermédiaire entre le facteur de transcription et l'ARN polymérase.



Le facteur de transcription stabilise l'ARN polymérase sur la boîte TATA par l'intermédiaire des médiateurs.

Les facteurs de transcription ont deux grands domaines :

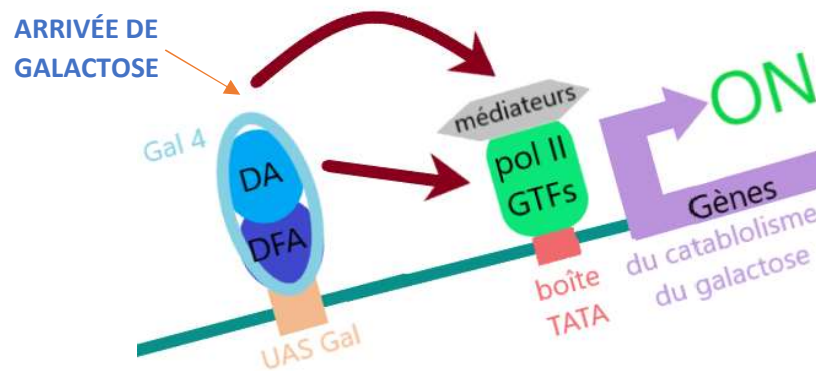
- Le **domaine de fixation** qui se fixe au promoteur et qui donne sa **spécificité** au facteur de transcription.
- Le **domaine d'activation** qui interagit avec les médiateurs pour **stabiliser** l'ARN polymérase.

Ces deux domaines appartiennent à **la même protéine** mais ils peuvent être séparés et intervertis artificiellement.

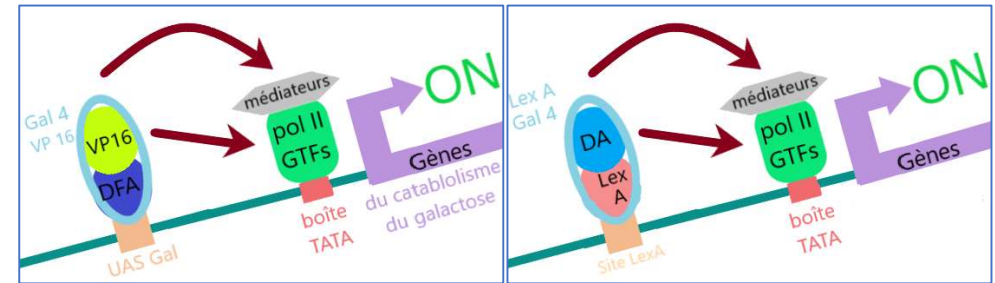
Exemple d'expérience : réaction de la levure en présence de galactose

Si l'on met du galactose dans le milieu de culture d'une levure, elle exprimera les gènes correspondant au catabolisme du galactose → le facteur de transcription **Gal4** se fixera sur le site de fixation (**UAS** chez la levure/**Eléments de réponse (ER)** chez nous) et va ainsi fabriquer des médiateurs...

Dans le cas de **Gal4**, le domaine de fixation est appelé **DFA** et le domaine d'activation est appelé **DA**.



Si l'on remplace **DFA** par un autre domaine de fixation (**LexA** = domaine de fixation d'origine bactérienne), on a **quand même transcription**. De même, si l'on remplace **DA** par un autre domaine d'activation (**VP16** = domaine d'activation d'origine virale) on a quand même transcription. Donc cette expérience montre que les **domaines d'activation et de reconnaissance d'un facteur de transcription sont séparables** et qu'il est **possible de les intervertir** (même avec des espèces différentes !).

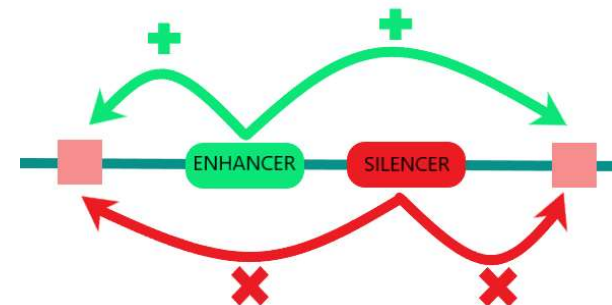


Les facteurs de transcription sont d'organisation modulaire

Les facteurs de transcription sont composés de plus petits modules interchangeables. Les contrôles proximaux sont des mécanismes très conservés avec des phénomènes universels (*un virus de l'herpès fonctionne chez la levure*).

c) Le contrôle distal de la transcription : les enhancers et les silencers

On a compris ce qui se passait au niveau des promoteurs, mais ce n'est qu'un niveau de régulation. Ce qui se passe au niveau des promoteurs ne constitue pas l'ensemble des éléments qui se déroulent : On retrouve aussi des **contrôles distaux** ! Chez les mammifères (et les métazoaires en général), on a des contrôles distaux (les bactéries ne comportent quant à elle que des contrôles proximaux).



Il y a d'autres séquences qui vont être importantes pour déterminer si le gène est ON ou OFF. Le contrôle distal (jusqu'à plusieurs kilobases) de la transcription se fait par deux catégories de séquence : les **enhancers** et les **silencers**.

Ils sont très différents des promoteurs, en effet :

- ➔ Ils ont **une position variable** par rapport au promoteur (en 5' ou 3').
- ➔ Ils peuvent être **très éloignés** (à des milliers / millions de paires de base).
- ➔ Ils agissent **indifféremment de l'orientation**.
- ➔ Ils agissent majoritairement en cis (sur un même chromosome), mais il y a des exceptions : certains **enhancers** peuvent agir en trans (sur un autre chromosome) mais c'est **très peu fréquent**.

Les enhancers activent, les silencers inhibent la transcription.

Ce sont des phénomènes de régulation qui agissent à distance et qui sont essentiels pour l'expression tissu-spécifique des gènes.

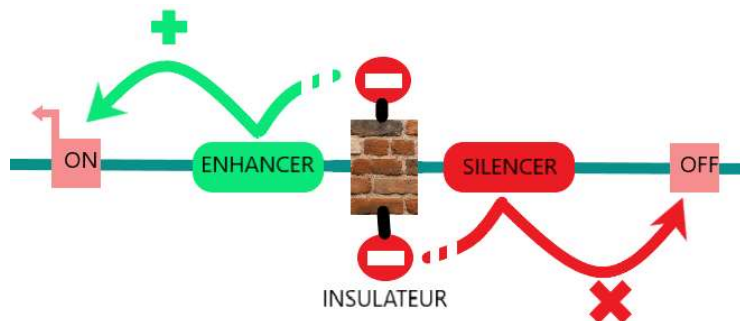
On retrouve presque **les mêmes facteurs de transcription** au niveau des **enhancers/silencers** qu'au niveau des promoteurs, certains facteurs sont même communs aux **silencers** et **enhancers** qui ont pourtant des conséquences très différentes. Ce n'est pas vraiment la fixation de ce facteur de transcription en particulier qui est important mais **la combinaison de différents facteurs** qui agissent. Ainsi, un même facteur de transcription va être activateur ou répresseur en fonction des autres éléments qui l'entourent.

L'action à distance de ces éléments induit une possible action sur le même gène en même temps des **enhancers** et des **silencers**, ce qui créerait une cacophonie dans la régulation. Ainsi, les chercheurs se sont demandé s'il n'y avait pas un autre élément qui intervenait pour segmenter cette régulation distale.

On a des **silencers** et des **enhancers** qui peuvent être activés par les mêmes facteurs de transcription, de plus des **enhancers** et des **silencers** peuvent s'exprimer en même temps il va donc falloir segmenter la régulation pour éviter de se perdre dans les ordres et les contre-ordres

d) Le contrôle distal de la transcription : les insulateurs

Insulateur ou élément frontière → élément qui empêche l'activation ou la répression, sans empêcher la fonctionnalité des **enhancers** et **silencers** !



Les insulateurs permettent de mettre un peu d'ordre dans tout ça. Ils permettent de donner un sens à l'action des **silencers** et des **enhancers**. Ces éléments frontières n'empêchent pas l'action des **enhancers** et des **silencers** dans notre génome, mais vont **les empêcher d'agir dans une direction**. L'insulateur (le mur) bloque l'influence des **enhancers** et des **silencers**. *Sur ce*

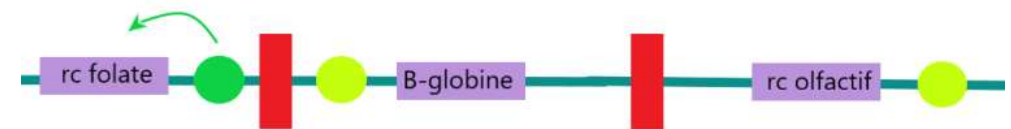
schéma l'insulateur du milieu empêche l'enhancer de gauche de favoriser la transcription plus à droite, et empêche le silencer de droite d'aller inhiber la transcription plus à gauche.

Ces insulateurs définissent des domaines de régulation spécifique dans notre génome

Exemple : l'érythropoïèse

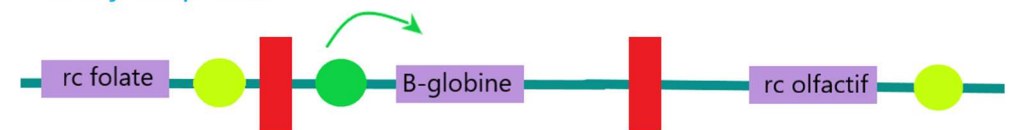
On a trois gènes, et on va voir ce qui se passe dans différentes cellules. On a côte à côte dans notre génome, le gène du récepteur folate, le gène β -globine (les deux impliqués dans l'érythropoïèse) et juste à côté, qui n'a rien à voir, un gène de la grande famille des récepteurs olfactifs (exprimé dans certaines cellules nerveuses). On a des **enhancers** tissu-spécifiques qui déterminent l'expression de ces gènes. Ils permettront d'activer le gène des récepteurs olfactifs dans les neurones et non dans les cellules de rates !

Début érythropoïèse



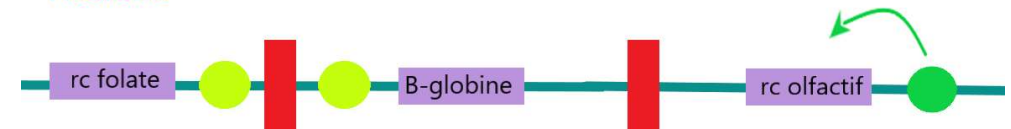
En début d'érythropoïèse : Le gène des récepteurs folate est ON, le gène de la β -globine n'est pas encore utile, et celui des récepteurs olfactifs est totalement inutiles : ils sont OFF.

Fin érythropoïèse



En fin d'érythropoïèse : on a plus besoin des récepteurs folate le gène est OFF, on a besoin de la β -globine il est ON, et le gène des récepteurs olfactifs est toujours OFF.

Neurone

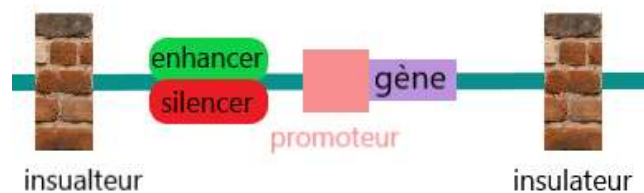


Dans les neurones : on ne fait pas d'érythropoïèse, les 2 premiers sont OFF, le gène des récepteurs olfactifs est ON

→ Le domaine du gène de la β -globine possède donc deux insulateurs (rectangles rouges sur le schéma) pour pouvoir réguler ces trois gènes convenablement.

Les **insulateurs définissent des domaines de régulation spécifique**, ici ils sont monogéniques mais pas nécessairement : on peut avoir des domaines encadrés par des insulateurs **composés de plusieurs gènes**.

Résumé des principaux éléments de régulations :



| Contrôle proximal | Contrôle distal | Régulation différentielle / distale |
|--|-----------------------------------|--|
| Promoteurs | Enhancer et silencer | Insulateur |
| En amont du gène | A distance Position variable | A distance |
| Agissent de manière unidirectionnelle (en amont) | Agissent dans les deux directions | Réduisent les enhancers et les silencers à une seule direction |

III. La structure de la chromatine

A- Compaction de l'ADN nucléaire

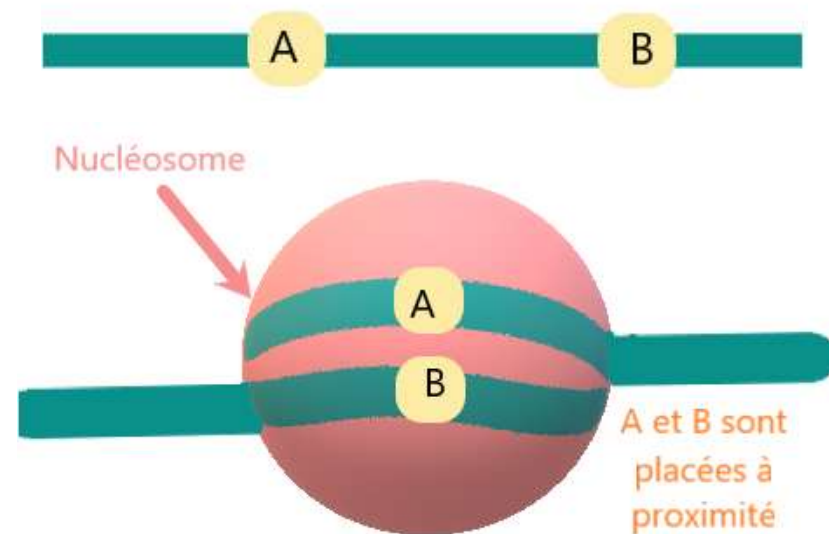
La **chromatine est le support des modifications** liées aux facteurs de transcription et à l'expression des gènes. En effet, elle constitue la façon dont l'ADN est organisé dans nos cellules : l'ADN n'est pas "nu". Chez l'Homme, le génome diploïde est composé de 6 milliards de bases (2 copies de 3 milliards de pb) alors que chez la levure, il y en a 200 fois moins. La taille totale de tout l'ADN dans chaque cellule fait 2 m qu'il faut faire rentrer dans un noyau inférieur à 10 microns (avec un volume de noyau de 400 à 1000 μm^3).

Il faut donc que l'ADN **se compacte** plusieurs milliers de fois pour entrer dans un noyau eucaryote. Ce **phénomène de compaction est indispensable** pour pouvoir contenir l'ensemble du génome dans le noyau. C'est comme si on voulait faire rentrer un fil fin de 40 km dans une balle de tennis.

| | Génome diploïde | Taille totale | Diamètre du noyau | Volume du noyau |
|--------|----------------------|----------------------|-------------------|----------------------------|
| Homme | 5.8×10^9 pb | 2m | 10 μm | 400 – 1000 μm^3 |
| Levure | 2.7×10^7 pb | 9×10^{-2} m | 2 μm | 4.2 μm^3 |

Par exemple, si on prend une molécule d'ADN avec deux nucléotides A et B séparés de 80 pb dans une structure linéaire, une fois dans le noyau, ils vont être placés à proximité : cela est dû au premier niveau de compaction de l'ADN : le **nucléosome**. Un nucléosome va être entouré d'environ 150 pb, et grâce à ça nos séquences A et B peuvent se retrouver très proches.

2 Séquences A et B séparées de 80 pb



→ L'ADN va s'entourer autour de protéines **histones** pour former la structure du nucléosome.
 → Le nucléosome est le **premier niveau de compaction** de l'ADN.
 → Il est formé de la molécule d'ADN et de 8 molécules d'histones.
 → Le nucléosome est un octamère correspondant à l'élément de base de la chromatine.

Exemple d'expérience : Etude de la chromatine in-vitro

On prend un ADN circulaire "nu" (ça n'existe pas en vrai, normalement on a toujours des protéines accrochées) que l'on met dans un tube à essai. On met cet ADN nu en présence d'extraits d'œufs de Xénope (les œufs de Xénope contiennent tout ce qu'il faut pour faire de la chromatine et de manière concentrée). On obtient alors un ADN très compacté, sous forme de chromatine s'opposant à l'expression des gènes.

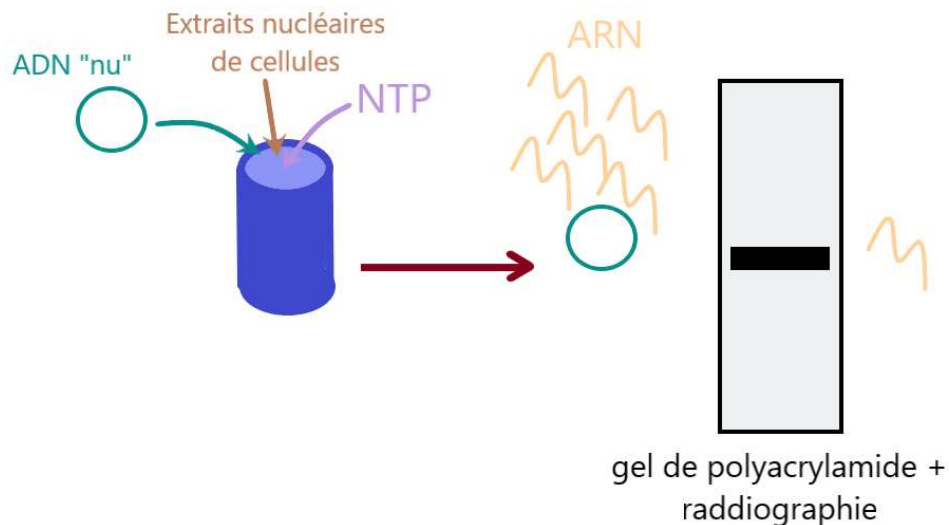
ADN circulaire nu + extraits d'œufs de xénope → ADN très compacté

On fait une transcription in-vitro en présence d'ADN nu, de NTP radioactifs et d'extraits nucléaires de cellules là où il y a transcription. Lorsqu'on fait une autoradiographie, on voit que l'ARN est visible sur le gel de polyacrylamide, il y a donc transcription.

En revanche, si l'on remplace l'ADN nu par de l'ADN obtenu à partir des œufs de Xénope (très compacté), il ne se passe rien il n'y a donc pas transcription, et donc **la structure de la chromatine est un inhibiteur de la transcription.**

ADN nu → **Transcription**

ADN des œufs de Xénope (compacté) → **Pas de transcription**

B- Structure nucléosomale

Un nucléosome est formé de :

→ 4 paires d'histones (un octamère) 2 x H2A, 2 x H2B, 2 x H3, 2 x H4



→ Cet octamère est entouré de 146 pb d'ADN qui en font 2 fois le tour

→ 1 nucléosome fait 108 kDa

-Les histones sont des protéines basiques, très riches en AA **chargés positivement**, ce qui importe dans l'interaction avec l'ADN chargé négativement (les opposés s'attirent)

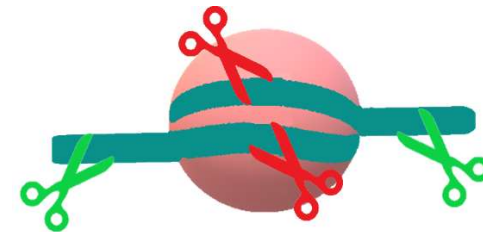
-Les histones ont été très bien conservées au cours de l'évolution.

-Il y en a environ **60 millions par noyau** de cellule **eucaryote** pour permettre de condensé tous l'ADN dans le noyau.

-Les eubactéries n'ont pas de nucléosomes !

On étudie ces structures grâce à une nucléase particulière, la **nucléase micrococcalle**. Cette nucléase a des propriétés de coupures spéciales :

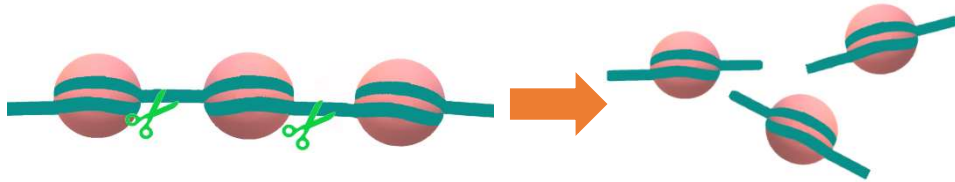
→ Elle est **incapable de couper au sein du nucléosome**, elle est obligée de couper à l'extérieur des nucléosomes, elle coupe donc **entre** les nucléosomes. *Et en fonction du degré de coupe, on pourra déduire la proportion d'ADN sous forme de nucléosome.*



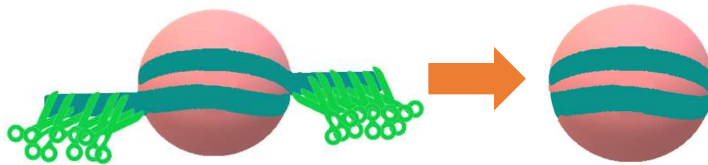
→ Elle nous permet en digérant l'ADN linker (= ADN internucléosomal) de le séparer des particules cœur (= nucléosome entouré des 146 pb) qui seront, elles, préservées.

On peut utiliser cette enzyme de 2 manières :

-Partielle : la nucléase va couper en moyenne 1 fois entre 2 nucléosomes donnant des fragments d'environ 200 pb, donc les nucléosomes sont séparés de 200 pb (la taille de l'ADN du nucléosome (146 pb) + taille de l'ADN linker (200-146=54pb))



-Totale : La nucléase coupera partout où elle le pourra, soit partout sauf sur les nucléosomes → On obtiendra donc une particule cœur de 146 pb (que l'ADN qui entoure les nucléosomes)



→ On peut ensuite **dissocier l'ADN des histones** en utilisant du **sel**, on obtiendra alors les 8 histones et le morceau d'ADN de 146 pb séparément.



→ Pour séparer les **différentes histones**, on devra utiliser de l'**urée** ou un **pH acide** (<5).



Tous ces phénomènes sont spontanés, c'est une propriété intrinsèque de l'ADN, cependant on retrouve des protéines facilitant ces réactions (**pas des enzymes** !)

Les protéines qui permettent les assemblages corrects de ces octamères sont des **protéines chaperonnes**, on en retrouve des spécifiques pour chaque histone. Et cet assemblage se fera dans un certain ordre :

1. Un premier chaperon est associé à un hétérodimère H3/H4 qu'il forme.
2. Ensuite, toujours à l'aide d'une protéine chaperonne, on aura l'assemblage de H2A et H2B.

C- La diversité des nucléosomes

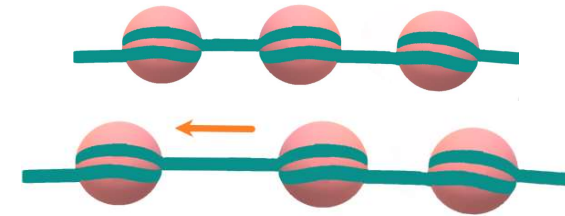
On a longtemps cru que tous les nucléosomes étaient identiques : ce n'est pas le cas ! Il existe 3 grandes façons de modifier des nucléosomes :

a. Le complexe de remodelage

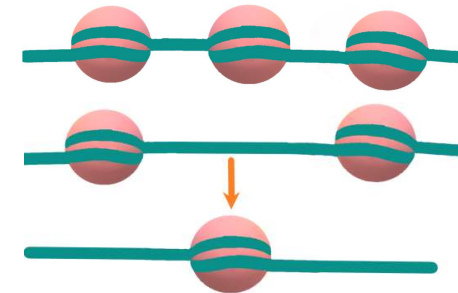
On peut modifier un nucléosome en changeant sa position. Il existe donc des complexes protéiques (=complexe de remodelage) qui vont, en fonction de la localisation dans le génome, choisir de déplacer tel ou tel nucléosome.

La position des nucléosomes peut être modifiée de différentes façons :

→ Déplacement en **CIS**, le long du brin d'ADN



→ Déplacement en **TRANS**, le nucléosome quitte le brin d'ADN pour aller sur une autre molécule d'ADN / apport d'un nucléosome depuis l'extérieur



Cela permet une diversité de positions des nucléosomes.

b. Les variants d'histones

A la place de certaines histones, on peut les remplacer par des versions un peu différentes qui sont les variants d'histone. Ainsi on a une forme principale et toute une série de variants. Chaque variant a une **fonction particulière** par rapport à certains domaines de chromatine, ces variants (comme les autres protéines) peuvent se modifier en post-traductionnel.

H4 est la seule histone qui ne possède pas de variants car elle est codée par un seul gène.

| Variants de H2A | Variants de H3 |
|-----------------------------------|---|
| H2AX H2A-Z MacroH2A ... etc | H3.1 H3.3 CenpA : Variant de H3 spécialisé dans les centromères. Donc les nucléosomes des centromères ont l'histone CenpA à la place de l'histone H3. |

c. Les modifications post-traductionnelles des histones

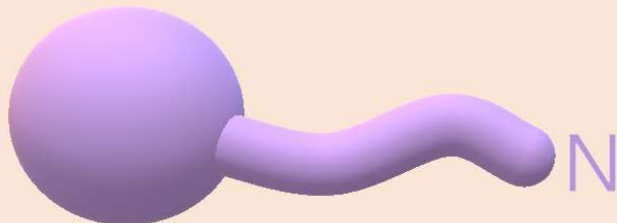
Une protéine, après avoir été synthétisée par le ribosome peut subir des modifications. C'est le cas pour les histones où des modifications spécifiques peuvent avoir lieu après la traduction grâce à des enzymes spécialisées. Il existe pleins de types, et on en découvre chaque semaine. Parmi les plus importantes il existe :

- ➔ Acétylation
- ➔ Phosphorylation
- ➔ Méthylation
- ➔ Ubiquitylation

Ces modifications ont un sens fonctionnel particulier pour la cellule

Chaque histone est formée de 2 parties principales :

- ➔ Une **partie globulaire centrale** qui a une structure en hélice α (elle permettra l'interaction entre les différentes molécules d'histones)
- ➔ Une **partie N-term la queue**. Cette queue traverse les nucléosomes entre les deux tours d'ADN.



La queue est donc dirigée vers l'extérieur. Ces queues sont très sensibles aux protéases, elles sont la cible de nombreuses modifications post-traductionnelles et sont très riches en AA basiques chargés positivement (arginine, lysine, etc). **Ce sont les seules structures exposées du nucléosome.**

1. Acétylation (Lysine)

C'est une des modifications les plus classiques, elle concerne les **lysines**.

➔ Les enzymes qui catalysent cette **acétylation** s'appellent des **HAT = Histone Acetyl Transferase** (exemple : p300, CBP, PCAF, GCNS... Le cofacteur de ces réactions est l'Acétyl-CoA).

➔ Ces réactions sont **réversibles** grâce à des **Histone DéAcetylases (HDAC)** qui catalysent la réaction inverse, la **désacétylation**. Elles sont capables d'enlever cette acétylation et de ramener la lysine à son état d'origine (exemple : Il en existe 2 catégories HDAC : I, II, IV et HDAC sirtuines qui jouent un rôle dans le phénomène de vieillissement en utilisant le NAD comme coenzyme).

On s'aperçoit qu'il existe un **lien entre la structure de la chromatine** et le **métabolisme énergétique** car l'Acétyl-CoA et le NAD sont des produits du métabolisme. Donc **quand on modifie le métabolisme, on modifie la chromatine.**

2. Méthylation (Lysine et Arginine)

C'est une autre des modifications les plus classiques des histones.

➔ Les **HMTases = Histone MéthylTransférase** qui catalysent la **méthylation** (exemple : domaine SET = Su(var)3-9 + Enhancer-of-zeste + Trithorax...).

➔ Il y a aussi les **HDM = Histone DéMéthylase** (LSD1, domaine JMJC) qui **déméthylent** la lysine.

➔ On peut passer d'aucune modification à 3 méthylations. Chacune des 3 méthylations nécessite **sa propre HMTases** spécifique. Donc 1 HMTase pour chaque méthylation. Elles sont aussi **spécifiques de la position de la lysine**

Ne pas confondre méthylation des lysines et méthylation de l'ADN

Il existe une très grande diversité d'histone. Cela nous amène à la notion de **code Histone**.

➔ En effet, ce code histone se rajoute code génétique qui correspond à l'ADN. Il est modifiable facilement.

➔ Ces codes histones vont être différents d'une cellule à l'autre.

• En général :

➔ Les **serines** sont modifiables par **phosphorylation**.

➔ Les **lysines** sont modifiables par **méthylation** voir par **acétylation** (parfois les deux en même temps)

➔ Les **arginines** elles sont modifiables par **méthylation**.

3. La technique CHIP

La technique CHIP (ImmunoPrécipitation de Chromatine) nous permet d'étudier des modifications post-traductionnels des histones. C'est la technique la plus simple pour étudier l'état chromatinien des cellules à un moment spécifique.

Etape 1 → Fixation des cellules au formaldéhyde :

On va créer des pontages, des liaisons covalentes au sein de la molécule d'ADN, on cherche à attacher les histones à l'ADN. C'est ce qu'on appelle le cross-linking (les cellules sont encore vivantes)

→ On va utiliser le formaldéhyde car c'est une fixation qui permet de lier, de rajouter des liaisons covalentes au sein de la molécule.



Etape 2 → Fragmentation par ultrasons = sonication :

On va d'abord lyser les cellules.

Puis on récupère la chromatine qu'on va segmenter par utilisation d'ultra-son : On va utiliser des ondes ultra soniques qui vont être capable de casser l'ADN et en particulier de casser entre les histones de façon aléatoire

Etape 3 → Ajout d'Ac spécifiques de la modification recherchée :

On va utiliser des anticorps spécifiques du type de modification que l'on recherche. Par exemple, si l'on recherche des acétylations on va utiliser des Ac anti acetyl. Les Ac vont alors se fixer sur les nucléosomes qui portent cette modification.

Etape 4 → Purification des complexes immuns :

Précipitation des complexes immuns (Ac + chromatine), on se débarrasse de tout ce qui ne possède pas la modif recherchée. On obtient donc un panel de séquences comportant au moins une histone avec la modif, on récupère en fait grâce aux Ac tous les complexes qui contiennent les modifications recherchées.

Etape 5 → Réversion du pontage et purification de l'ADN

Réversion du pontage et purification de l'ADN en se débarrassant des histones. Ça va nous permettre de détruire les liaisons covalentes qu'on avait créé précédemment (à l'étape 1 avec le formaldéhyde) et du coup de séparer l'ADN des histones.

Donc tout d'abord, on va grâce au formaldéhyde coller l'ADN aux histones, puis on va couper l'ADN en petits morceaux avec les ultrasons. Ensuite, on va reconnaître les modifications qu'on veut étudier avec des Ac spécifiques, pour pouvoir les faire précipiter et donc finalement ne récupérer que les parties de l'ADN qui sont accrochées à ces histones modifiées. Finalement on va séparer l'ADN des histones en cassant les liaisons covalentes. On obtiendra donc les histones séparées des parties d'ADN associées, la prochaine étape sera donc de savoir quels sont les gènes régulés par ces histones modifiées.

On sait maintenant que l'ADN récupéré est l'ADN qui correspond à celui autour ou à proximité des histones possédant la modification recherchée. Il existe plusieurs techniques, mais toutes ont en commun le fait de quantifier le degré d'enrichissement de certaines régions particulières du génome.

On va alors utiliser plusieurs techniques afin d'étudier ces parties d'ADN :

- **L'hybridation**
- **La PCR** (détaillée en UE11 ou génétique) : technique d'amplification utilisée couramment pour identifier des séquences.
- **Séquençage**

Ces méthodes vont nous permettre de **déterminer l'enrichissement relatif** d'une séquence donnée par rapport :

- **A une séquence contrôle**
- **A l'ADN de départ (=input)**

Si on fait une PCR comparative entre ce qu'on avait au début et ce qu'on obtient après la précipitation et que la **composition est différente** → on va pouvoir déterminer **l'enrichissement de l'ADN**.

→ En effet, cet enrichissement est en fait une marque qui nous indique quelle quantité de la protéine que vous visiez avec votre anticorps se trouve sur cette séquence d'ADN.

Cet enrichissement F se calcule : On compare alors la quantité relative de la modification recherchée présente dans **l'ADN de départ**, appelé input (=Rc), avec celle présente dans l'immunoprécipitat (=ADN récupéré après précipitation = RIP). ($F = \text{Rip} / \text{Rc}$ avec $\text{RIP} = \text{Q fragment PCR} / \text{Q ADN réaction}$ et $\text{Rc} = \text{Q fragment PCR} / \text{Q ADN réaction}$)

- Si $\text{RIP} = \text{Rc}$ → aucun enrichissement du fragment qui nous intéresse n'est particulièrement associé à cette modification post-traductionnelle
- Si $\text{Rip} \gg \text{Rc}$ → enrichissement de l'ADN correspondant à une association de l'ADN à ces histones comportant la modification post-traductionnelle voulu

Si $RIP = R_c$, ça signifie qu'il y a proportionnellement autant de la modification d'histone étudiée dans l'ADN en général que dans la portion d'ADN étudiée. Alors que si $RIP \gg R_c$, on aura beaucoup plus de cette modification au niveau de la portion étudiée que dans le reste de l'ADN ! Elle sera donc très enrichie.

→ Pour une même extrémité N-term on peut avoir de nombreuses combinaisons de ces modifications mais **toutes les extrémités ne sont pas modifiées de la même façon**.

→ Ces modifications post-traductionnelles + l'utilisation des différents variants d'histones constituent un code qui se superpose au code génétique = **code épigénétique**. Il donne des informations supplémentaires à la cellule sur son programme d'expression, pour transcrire les bons gènes au bon moment.

On utilise cette technique au niveau des gènes vus p4 (gène folate, de la B-globine et du Rc olfactif). On s'aperçoit qu'au début de l'érythropoïèse, on a des histones 3 acétylées au niveau du gène folate, et qu'en fin d'érythropoïèse les histones 3 acétylées ne sont plus au niveau de ce gène mais au niveau de celui de la B-globine. On en déduit que l'acétylation de H3 est en général associée à une transcription active du gène → Il y a donc une relation entre les modifications post traductionnelles et le niveau de transcription des gènes.

4. Traduction fonctionnelle du code histone

Au niveau de l'acétylation :

→ Si la chromatine est **hyperacétylée**, on a pratiquement toujours une **transcription active**

→ Si la chromatine est **hypoacétylée**, on a pratiquement toujours une **transcription inactive**

Au niveau de la méthylation :

→ Si la chromatine est **méthylée en H3 K4** (lysine 4 de l'histone H3), la **transcription est active** ;

→ Si la chromatine est **méthylée en H3 K9** (lysine 9 de l'histone H3), la **transcription est inactive**.

Lors des acétylations, on neutralise la charge positive de la queue, cette dernière ne peut plus interagir avec l'ADN => Il n'y a plus de compaction = configuration ouverte (euchromatine). La méthylation ne modifie pas la charge. Ce n'est donc pas tellement le type de modification qui importe, mais plutôt sa localisation.

Ces modifications d'histones vont être lues par des protéines spécifiques :

→ Les **lysines acétylées** sont reconnues par des protéines qu'on appelle des **protéines à bromodomains**.

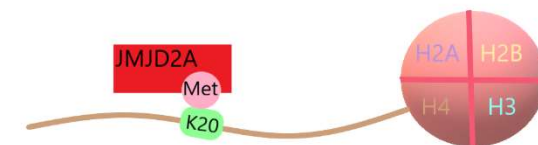
→ Une **méthylation en H3 K9 ou K27** est reconnue par des **protéines à chromodomaine**.

→ Une **phosphorylation en H3 S10** est reconnue par des **protéines à domaine 14-3-3**. Cette modification facilite l'acétylation et l'**activation de l'expression des gènes** en réponse au stress.

→ Une **méthylation de H4 K20** est reconnue par des **protéines au domaine Tudor**, ce qui a un rôle dans la **réparation de l'ADN**. Il y a 2 protéines à domaine Tudor qui peuvent se fixer sur H4 K20 méthylée et ce, de manière exclusive, c'est-à-dire que si l'une se fixe, l'autre ne peut pas : **JMJD2A** et **53BP1**. Ces protéines déterminent la réparation des cassures de l'ADN.

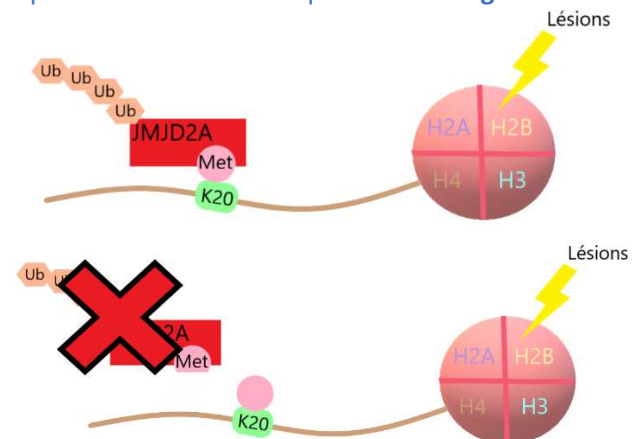
2 situations possibles :

→ **Pas de dommage à l'ADN** : Toutes les H4 méthylées en position K20 sont fixées par la protéine **JMJD2A** empêchant donc la fixation de **53BP1** impliquée dans la réparation.

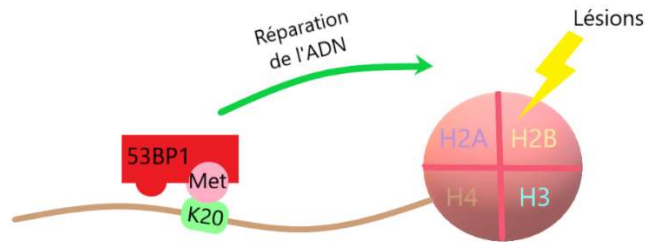


→ **Cassure accidentelle de l'ADN** au niveau d'un nucléosome

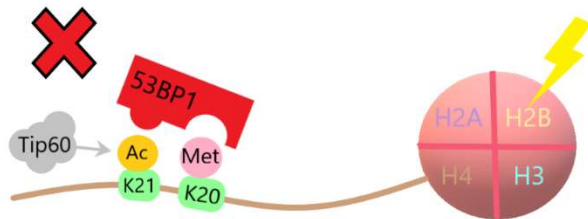
→ Poly ubiquitination de **JMJD2A** qui est donc **dégradée**



→ **53BP1** peut alors être **recrutée** ce qui déclenche la **réparation de la cassure** d'ADN



→ **MAIS** la fixation de **53BP1** peut être inhibée : par une acétylation de H4 K16 grâce à une HAT nommée **Tip 60**.



Toutes les modifications d'histones sont lues par d'autres protéines, lectrices du code et ayant une affinité particulière pour chacune des modifications : **les readers**. **Chaque lecteur va être un peu spécialisé.**

→ Les « writers » sont celles responsables des modifications (HAT, HMT, HD...)

→ Les « erasers » sont celles responsables des suppression (HDAC...)

IV. La fibre nucléosomale

A- Niveaux de compaction

On observe **2** niveaux de compaction de l'ADN :

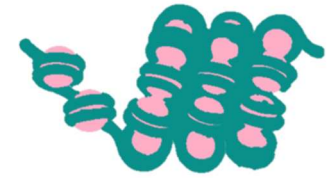
PREMIER NIVEAU :

- On a une fibre de **11 nm** (diamètre du nucléosome)
- C'est une succession de nucléosomes associés par un ADN de liaison (**linker**)
- Comparable à un collier de perle
- **Ça correspond à la configuration OUVERTE de l'ADN**



SECOND NIVEAU :

- On a une fibre de **30 nm** (un peu moins de 3 nucléosomes) → c'est un solénoïde
- En fibre de nucléosomes, cette forme est moins connue
- **Ça correspond à la configuration FERMÉE de l'ADN**



Plus on compacte de l'ADN plus on rend sa transcription difficile

→ La transition de la fibre de 11 nm (niveau 1) en fibre de 30 nm (niveau 2) se fait grâce à l'**histone H1** (qui n'appartient **jamais** au nucléosome formé de H2A, H2B, H3, H4) elle se fixe à l'**extérieur du nucléosome, à la base de la particule de 11nm.**

- H1 se fixe sur 1 seul nucléosome qui est séparé d'autres par de l'ADN linker afin de former des **fibres nucléosomales**.

B- Remodelage de la fibre nucléosomale

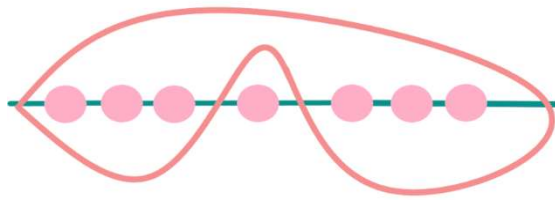
C'est souvent en remodelant ces fibres nucléosomales qu'on peut accéder aux promoteurs et fixer les protéines régulatrices (ex : *facteurs de transcription*). Et ce grâce aux **facteurs de remodelage** (FR), de grosses machines très **consommatrices d'énergie**, qui rendent l'ADN accessible. Ils vont être capable de **créer localement des zones sans nucléosomes** (en les déplaçant grâce à de l'ATP).

Il existe **différents types** de remodelage ATP-dépendants :

| Déstructuration | Mobilité en cis du nucléosome | Échange en trans de 2 nucléosomes |
|--|--|---|
| | | |
| On changera la structure du nucléosome, sans changer sa position , permettant par exemple d'exposer le site de fixation à un facteur de transcription) | On bougera en cis les nucléosomes (donc sur le même brin d'ADN), exposant ainsi une zone « nue » d'ADN où les facteurs de transcription pourront se fixer. | On échangera 2 nucléosomes entre 2 régions différentes de l'ADN, modification en trans. |

Il existe **3** grandes familles d'ATPases de remodelage :

1. SWI/SNF



- Gros complexe de 2MDa
- Il contient 11 polypeptides, avec une unité ATPase : **SWI2/SNF2**
- Il augmente l'accessibilité des zones de fixation aux facteurs

de transcription sans modifier la position du nucléosome

- On va agir sur la **structure** de l'ADN qui entoure le nucléosome
 ➔ *C'est comparable au fonctionnement d'un vaisseau extra-terrestre qui atterrit sur 7 nucléosomes et qui concentre son action sur le nucléosome central (genre comme les Ets qui survolent une famille mais ne kidnappe qu'un seul membre)*

2. ISWI/NURF

- Complexe plus petit de 500 kDa.
- Il contient 4 polypeptides
- Il est associé à l'activation de la transcription, augmentant la mobilité des nucléosomes (déplacements ou échanges = cis ou trans)
- On a différents types de complexes de ce genre :
 ➔ ACF = ISWI + ACF1
 ➔ CHIRAC = ASWI + ACF1 + p14/p16 = Jacques + Bernadette
(Ces exemples osent totalement)

3. Mi2 ou NURD

- NURD = Nucleosome Remodeling Deacetylase
- Il va avoir (comme son nom l'indique) un rôle de **remodelage** et un rôle de **désacétylation** (donc souvent associé à une répression de la transcription)

V. Boucles et domaines

La fibre chromatinienne s'organise en superstructures ➔ des organisations en **boucles et domaines**

A- Etude expérimentale

Pour étudier ces structures, on va utiliser une endonucléase pancréatique : la **DNaseI**, elle diffère beaucoup de la nucléase micrococcalle (qui permet d'étudier la structure nucléosomale), notamment parce qu'elle **peut couper sur le nucléosome**.

A propos de la DNase 1 :

- ➔ Elle peut couper l'ADN qui est sur les nucléosomes
- ➔ Elle coupe **très régulièrement** l'ADN entourant le nucléosome (tous les 10 pdb) car le petit sillon de l'ADN est exposé à la surface du nucléosome à chaque tour d'hélice (chaque tour d'hélice est espacé de 10 pdb).



- ➔ L'activité de la DNase 1 sera fortement augmentée lorsque l'ADN n'est pas associé au nucléosome, au niveau de l'ADN linker (vu qu'elle peut couper dès que le petit sillon n'est pas collé au nucléosome, elle peut donc couper tout le long de l'ADN linker). **Plus la chromatine est compactée, moins bien la DNase1 pourra couper.**

On va donc utiliser cette nucléase pour sonder le niveau de compaction de la chromatine

- ⇒ La chromatine des gènes transcrits (actifs) est plus sensible à la DNase 1 que celle des gènes non transcrits, car non condensée.

Exemple : le gène globine

- On va prendre le gène globine dans des érythroblastes de poulet.
- Ces gènes vont normalement synthétiser la protéine globine.
- Ce gène sera actif dans certaines cellules et inactif dans d'autres qui n'ont pas besoin de globine.

➔ On va faire un traitement à la DNase1

On a 2 solutions possibles :

- Soit le gène globine n'est pas exprimé : l'ADN étudié sera donc peu digéré par la DNase 1. En effet : pas d'expression ➔ compaction ➔ peu digéré par la DNase 1.
- Soit le gène globine est exprimé : En fonction de la dose de DNase 1, on n'a plus rien à la fin car la chromatine, en configuration ouverte laisse tout l'ADN accessible à la DNase 1 qui le digèrera.

La DNase 1 permet de savoir quels sont les gènes transcrits, car ils vont être plus sensibles à son action.

Récap :

- Un gène s'exprime ➔ il est totalement digéré par la DNase 1
- Un gène ne s'exprime pas ➔ il est très peu digéré par la DNase 1

A force d'expériences, on a fini par découvrir qu'au sein des domaines sensibles à la DNase1, on a des zones encore plus sensibles (hypersensibles HS). Ces éléments HS correspondent à des éléments de régulation souvent localisés en 5' des gènes (ex : le promoteur, certains FT spéciaux).

B- Les sites hypersensibles

Ces zones sont hypersensibles parce qu'elles ont encore moins de nucléosome que l'ADN au niveau des zones sensibles. Il y a **aucun** nucléosome dans ces domaines HS.

Ces sites se trouvent surtout au niveau des **promoteurs** des gènes et des sites de démarrage de transcription mais aussi au niveau des **enhancers** et des **insulateurs**.

Ces zones HS correspondent globalement à des zones de fixation des acteurs de la transcription.

Cependant on a aussi découvert en augmentant encore la résolution, au nucléotide près, qu'au sein de ces zones HS, on a des petites zones complètement insensibles à la DNase1 → ces zones correspondent **au lieu précis** du site de fixation des facteurs de transcription.

Dans les zones sensibles à la DNase1, on a des zones HS, et dans ces zones HS, on a des zones insensibles

C- Niveau d'activité des gènes

En fonction de la résistance à la DNase1, on peut définir le niveau de compaction de la chromatine (euchromatine ou hétérochromatine) et donc le niveau d'expression des gènes. On définit **3 états** pour les gènes :

→ Gène inactif (pas exprimé) = Gène dans un état fermé.

→ Gène compétent = Gène dans un état intermédiaire

→ Gène actif (exprimé) = Gène dans un état ouvert

1. Gène inactif

On a des marques qui répriment la transcription :

→ Méthylation de H3K9

C'est un état fermé donc résistant à la DNase1 → chromatine dense + domaines insensibles à la DNase1.

Ça correspond à la chromatine sous forme de fibre de 30 nm par exemple

2. Gène compétent

C'est un gène non-exprimé, mais prêt à l'être si besoin :

→ Acétylé en H3/H4, mais **non méthylé en H3K4**.

Il possède des domaines sensibles à la DNase1, mais **sans zones HS**.

3. Gène actif

On a :

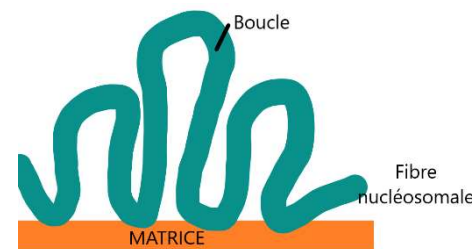
→ Acétylation en H3/H4

→ Méthylation de H3K4

On aura donc une sensibilité à la DNase1 avec aussi des zones HS.

Cet état correspond à la fibre de 11 nm, gène actif en cours de transcription avec des marques d'activités des gènes.

D- Co-régulation



Dans un même domaine, les gènes peuvent être **corégulés**, ils s'exprimeront alors en même temps. C'est ce qu'on appelle les **domaines de co-régulation**. On peut grâce à différentes techniques cartographier ces domaines.

Les fibres nucléosomales sont organisées

en **boucles** attachées à la matrice nucléaire.

Ces domaines de co-régulation correspondent finalement à ces **boucles** : les gènes corégulés appartiennent à la même boucle

E- La matrice nucléaire

C'est en quelques sortes le cytosquelette du noyau, peu connu il n'est pas aussi bien caractérisé que le cytosquelette. Les fibres de nucléosomes s'attacheront à cette matrice à des endroits particuliers.

La matrice cellulaire c'est ce qu'il reste du matériel nucléaire qui a :

→ Subi une extraction des nucléosomes et les autres protéines avec des détergents,

→ Subi une digestion à la DNase1

→ Ou été exposé à des concentrations salines très importantes.

Elle est caractérisée par le fait qu'elle est **insoluble** lorsqu'on lyse le noyau (= c'est un composant **insoluble** du noyau).

On a établi certains des **constituants** de cette matrice :

→ Complexes nucléoprotéiques

→ Un certain nombre de protéines (dont la lamina nucléaire)

→ Un réseau fibreux du nucléosquelette formé **d'actine** (microfilaments) et de protéines de filaments intermédiaires comme les **lamines A et C** (ces

deux lamines sont plus solubles que la lamine B qui est enchâssé dans la membrane nucléaire interne)

- La protéine **NuMa spécifique du noyau** (+++), c'est une protéine de type filament intermédiaire.

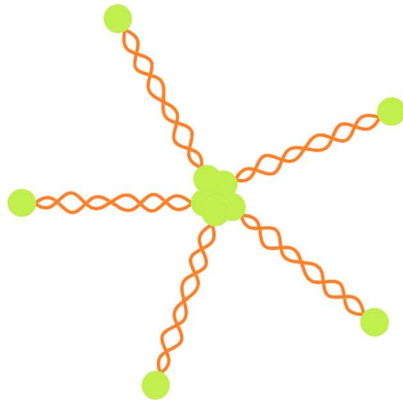
L'organisation de la matrice :

Cette matrice, grâce à son aspect fibreux peut aussi être appelée « **scaffold** » (échafaudage), ou bien nucléosquelette.

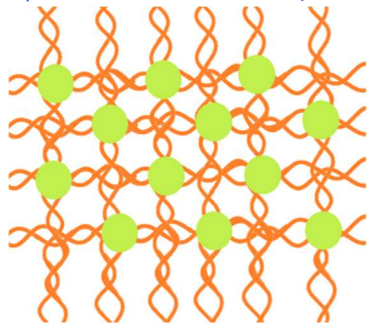
- On a un monomère NuMa, une protéine **allongée** qui se présente avec une structure fibreuse centrale et des extrémités globulaires.



- 5 monomères s'associent en pentamères.



- Les pentamères s'assembleront ensuite pour former un réseau fibreux (quadrillage) participant à la matrice qui interagira avec l'ADN (on va former un filet sur lequel s'accrochera l'ADN).



Cette matrice a à la fois un rôle **structurel** et **fonctionnel** (rôle dans l'expression des gènes)

Expérience : Inhibition de NuMa

On va utiliser des cellules épithéliales mammaires : on a une lame basale artificielle sur laquelle on va faire reposer les cellules en suspension. On pourra observer, grâce à des techniques d'immunofluorescence, 2 constituants cellulaires, les lamines A/C d'une part et d'autre part la protéine NuMa.

→ On a d'abord une monocouche polarisée de ces cellules épithéliales, avec des jonctions serrées.

- On verra que dans cette configuration-ci, la protéine **NuMa est présente de manière diffuse dans le noyau** (=partout dans le noyau) excepté dans le nucléole.

→ On va alors faire varier les conditions de l'expérience afin d'induire la formation de glandes avec formation d'acini. On reproduira donc la structure de la glande mammaire.

- On observe ensuite nos cellules au microscope à fluorescence et on remarque que par rapport au début de l'expérience, la distribution **des lamines A et C** dans les cellules **n'a pas été modifiée** par la formation d'acini.
- En revanche, la distribution de la protéine **NuMa** et donc du nucléosquelette **est modifiée au cours de ce phénomène de différenciation**. En effet depuis la différenciation, NuMa est localisé dans des **foyers précis du noyau**.

→ On va alors inhiber la protéine NuMa en utilisant des Ac Anti-NuMa.

- On observe alors qu'en l'absence de NuMa, l'épithélium quitte sa conformation glandulaire pour revenir à une monocouche polarisée → **il se dédifférencie**.

Conclusion :

La protéine NuMa est importante dans le processus de différenciation en intervenant dans la formation de la matrice nucléaire.

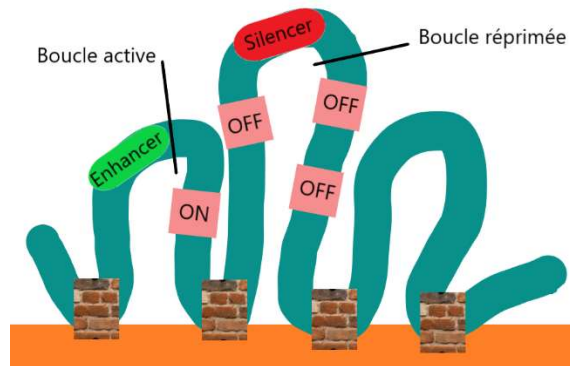
- La protéine NuMa intervient dans la formation de la matrice nucléaire.
- La matrice nucléaire est importante pour la différenciation des cellules.

On s'est aperçu que les **zones de fixation** des boucles de fibre chromatinienne sur la matrice nucléaire étaient des **régions insulatriques**. Ce sont des éléments frontières qui **séparent physiquement les boucles**.

- Les **domaines transcriptionnels** (gènes actifs ou réprimés) correspondent aux **boucles limitées par les insulateurs**.

Cette matrice est **importante pour la différenciation des cellules**.

C'est en s'attachant à la matrice que **les insulateurs** vont définir des **domaines** dans lesquels des éléments de régulation à distance comme les enhancers ou les silencers vont pouvoir agir et vont donc limiter leur action en **séparant physiquement les boucles**. (Quand dans une boucle, une région d'ADN comprise entre 2 insulateurs fixés à la matrice nucléaire, on retrouve 3 gènes et que cette boucle est sous l'influence d'un silencer, par conséquent les 3 gènes ne s'exprimeront pas). On peut donc avoir une boucle activée d'un côté d'un insulateur et de l'autre côté de celui-ci une boucle inactivée.



Voilà la fiche sur la première partie du noyau, je vous sors bientôt une fiche avec la suite et une autre avec les deux parties en une ! Vous avez pour l'instant la première ronéo sur le noyau et les 20 premières pages de la seconde à peu près ! J'espère que vous l'aimerez bien !

(Et cœur sur les pouilleux bien sûr)