

## Cours n°6 : Modifications post-traductionnelles, fonctions et stabilité des protéines

Lexique :

HDAC : histone désacétylase

HAT : histone acétyltransférase

HMT : histone méthyl transférase

Ex : exemple

*Les points importants à savoir sont donnés par les objectifs du cours situés à la fin de ce cours. Ce sont aussi les points susceptibles de tomber au partiel.*

*La professeure accepte de relire la ronéo.*

*C'est la première année qu'elle fait ce cours, en remplacement du professeur Cavé, mais beaucoup de ses diapositives sont reprises du cours de l'an passé.*

## SOMMAIRE :

### Introduction

#### **I. L'épissage alternatif**

- A. Généralités sur les protéines
- B. L'épissage alternatif

#### **II. Protéomique et spectrométrie de masse**

- A. Objectifs et principe de la protéomique
- B. Protocole de la spectrométrie de masse :

#### **III. Modifications post-traductionnelles**

- A. Définitions et exemples
- B. La phosphorylation
- C. L'acétylation et la méthylation des histones
- D. Acétylation de p53

#### **IV. Dégradation par le système ubiquitine-protéasome**

- A. Le système ubiquitine
- B. Le protéasome
- C. L'ubiquitination n'est pas toujours un signal de dégradation
- D. Les systèmes pseudo-ubiquitaires

#### **V. Dans un contexte médical**

### Conclusion

## Introduction :

La grande diversité des activités cellulaires par rapport au faible nombre de gènes qui codent pour les protéines peut s'expliquer par :

- **L'épissage alternatif** qui augmente la diversité du transcriptome
- Les nombreuses **modifications post-traductionnelles** qui augmentent les potentialités fonctionnelles des protéines.

Ce cours s'intéresse à la dynamique cellulaire et à la dynamique protéique dans la cellule. Il existe un certain nombre de dogmes en biologie et ils sont régulièrement revisités à la lumière des différentes techniques et des différentes connaissances.

### **I. L'épissage alternatif :**

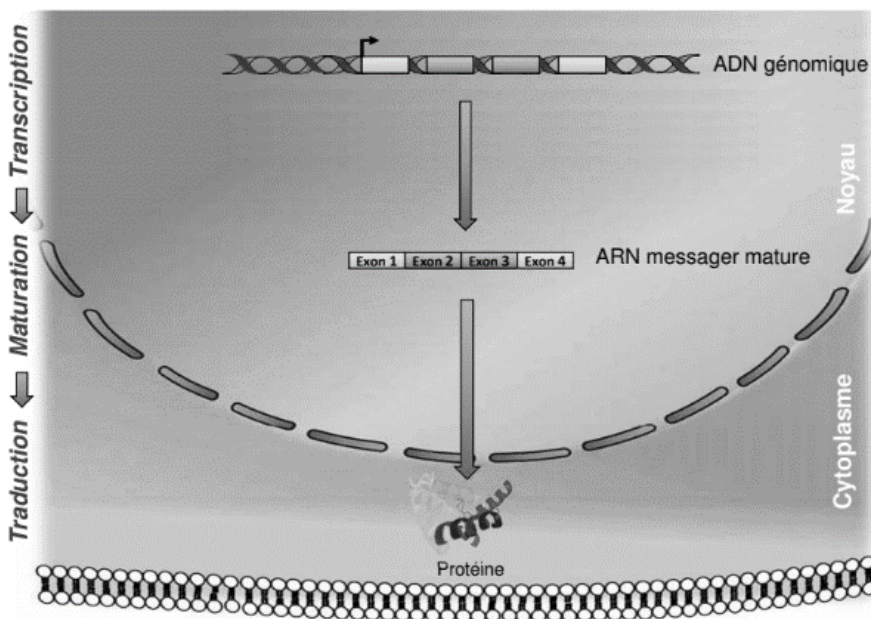
#### **A. Généralités sur les protéines :**

Les protéines sont les **effecteurs** des cellules, on ne se focalise donc plus seulement sur les acides nucléiques et le génome. Les protéines sont les acteurs réels de la cellule et ont la capacité d'endosser un certain nombre d'activités cellulaires.

La taille du génome humain est de l'ordre du **milliard** de paires de bases. La partie codante de ce génome est variable selon les données mais de l'ordre de **5%** du génome. Ce génome comprend environ **25 000 gènes** chez l'homme. La définition du gène étant : une séquence codante pour une protéine, cela voudrait dire qu'il y aurait au moins 25 000 protéines. Mais il y a beaucoup plus de protéines que cela.

Comment expliquer qu'à partir de 25 000 gènes, on puisse avoir plus de protéines ?

L'ADN génomique est composé de séquences codantes et non codantes (organisation modulaire). Les séquences non codantes comprennent les zones de régulation et les zones introniques (donc pas conservées après la transcription et la maturation du messenger)

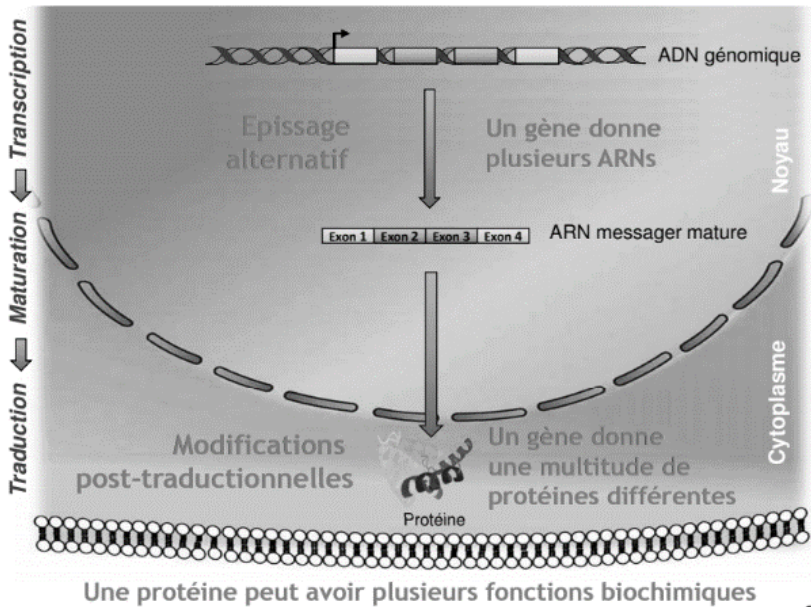


Le dogme classique, cru pendant un certain nombre d'années, dit que pour un gène, on obtient une protéine. On pensait donc que 25 000 gènes donnent 25 000 protéines.

Par exemple si on considère un gène avec 4 exons de couleurs différentes : l'ADN serait transcrit en un pré-messenger qui est mûri en un messenger. Ce messenger mature comprendrait donc les 4 exons fusionnés qui sont ici tous traduits en une protéine.

#### **B. L'épissage alternatif :**

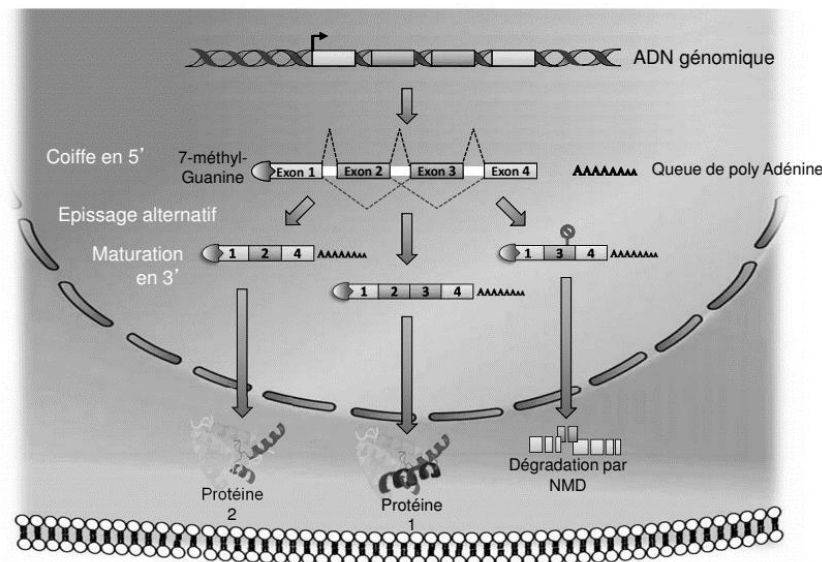
Or, on sait aujourd'hui que le processus est beaucoup plus complexe que cela et qu'on peut former en moyenne 4 protéines différentes à partir d'un seul gène. Cela est permis notamment par la diversité du transcriptome (ensemble des différents transcrits cellulaires) induite par **l'épissage alternatif**.



L'épissage alternatif correspond au fait qu'un unique gène peut donner plusieurs transcrits, plusieurs ARNm, puis une grande variabilité au niveau de la traduction. On peut donc obtenir plusieurs protéines potentielles, et celles-ci peuvent également subir des modifications post traductionnelles. **En partant de 25 000 on obtient donc un éventail considérable de fonctions finales => c'est l'augmentation des potentialités fonctionnelles de protéines.**

Pendant la transcription, il y a une copie conforme de tout l'ADN génomique en ARN complémentaire. Ces messagers sont soumis à des mécanismes de **régulation et de maturation**, notamment par les coiffes et les queues polyadénylées (il peut y avoir plusieurs zones de polyadénylation), qui vont contrôler l'expression de plusieurs transcrits et donc avoir une incidence en aval.

Mais dans un acide nucléique, comment différencier un intron d'un exon ? Il y a des sites consensus avec des sites donneurs et accepteurs d'épissage. Mais il y a énormément de dimères de nucléotides donc il est excessivement difficile d'être absolument sûr que ce di-nucléotide est un site donneur ou un site accepteur. Pour le savoir il y a nécessité de tester les dimères de nucléotides. Il y a plusieurs combinaisons de sites donneurs et de sites accepteurs et selon les paires qui sont formées on a des épissages différentiels.



Comme le montre ce schéma, il est possible d'enlever tous les introns et fusionner les 4 exons. On peut également avoir un saut d'exon, par exemple du 2 ou du 3.

On peut avoir des variations au niveau de la queue de polyadénylation, par exemple des signaux d'instabilité des messagers. C'est une modification post-transcriptionnelle, donc une modification de la structure du message. Cette modification permet de signaler à la cellule que ce message n'a pas de raison d'être et qu'il ne faut pas le traduire. Ce sont des messages de garantie qualité pour savoir si le message est correct bien polyadénylé et prêt à être traduit. Si ces conditions ne sont pas respectées, le message est dégradé.

Il existe des systèmes de **vérification de la qualité** de l'information génétique. On a par exemple un système qu'on appelle NMD qui permet la dégradation des ARN non fonctionnels (qui donneraient naissance à des protéines non efficaces voire délétères s'ils n'étaient pas détruits).

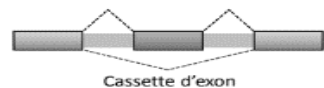
Par ailleurs on a une multiplicité des possibilités. Selon les gènes, on peut avoir un ou plusieurs transcrits. Par exemple CD44 peut avoir plusieurs transcrits, la neurexine environ 2 000, DSCAM a énormément de transcrits différents, épissés, polyadénylés en différents sites (niveau de complexité majeur).

Il arrive parfois qu'un même gène, à partir de variations d'épissage, donne des protéines avec des fonctions extrêmement différentes. Par exemple le gène CALCA peut donner deux protéines fonctionnellement très différentes :

- La calcitonine, hormone du métabolisme du phosphore et du calcium,
- La CGRP, vasodilatateur et médiateur de la douleur.

Autre exemple, le gène qui code pour Arf ou p16, deux acteurs majeurs dans la régulation du cycle cellulaire.

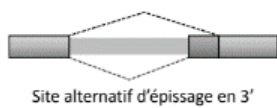
### Les 5 grands modes de l'épissage alternatif : (très important)



Saut d'exons : l'exon peut être conservé ou éliminé.

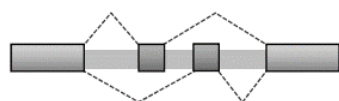


Site alternatif d'épissage en 5'



Site alternatif d'épissage en 3'

Variations d'épissage en 5' ou en 3' avec des sites alternatifs d'épissage.



Exons mutuellement exclusifs

Exons mutuellement exclusifs : les 4 exons ne seront jamais ensemble, le 2 et le 3 ici ne pourront jamais être conservés en même temps

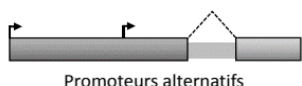


Rétention d'intron

Rétention d'introns : L'épissage n'aura pas lieu et on va conserver l'intron et les exons. L'intron va être traduit avec les exons et cela génère une nouvelle protéine.

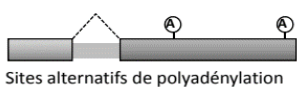
L'épissage alternatif permet :

- une modification de la stabilité des transcrits, ce qui peut avoir un impact majeur sur la protéine qui va être exprimée
- un changement de la localisation subcellulaire de la protéine codée (sélection ou non de l'exon qui code la localisation).
- des modifications des domaines d'interactions.
- des modifications des fonctions biologiques des protéines, notamment par des modifications de la forme et du repliement des protéines.



Promoteurs alternatifs

On pense souvent que la traduction commence au début du premier exon et se termine à la fin du dernier exon. La traduction dépend du point de départ et du point d'arrivée. Le ribosome cherche sa méthionine initiatrice, qui n'est pas nécessairement au début du premier exon, et s'arrête dès qu'il trouve un codon stop, qui n'est pas forcément à la fin du dernier exon.



Sites alternatifs de polyadénylation

Cela crée une variabilité avec l'existence de **promoteurs alternatifs** (situés sur le transcrit) qui permettent de commencer au début ou au milieu de l'exon. On peut avoir également des sites alternatifs de queue de polyadénylation.

L'épissage alternatif peut donc expliquer la diversité de transcription, et il intervient de façon très **fréquente**. A la fin des années 70-80, on pensait que l'épissage alternatif ne concernait que 1% des gènes humains et en 2010 on se rend compte que cela concerne **95%** des gènes.

Il a été établi une moyenne sur cette diversité fonctionnelle. On estime qu'il y a environ **100 000** protéines humaines, codées par des gènes comportant en moyenne **9 exons**.

## **II. Protéomique et spectrométrie de masse :**

On se rend compte que dans une situation physiopathologique, il est plus intéressant d'étudier le **protéome (ensemble des protéines)** plutôt que seulement la signature génétique du patient. Les techniques d'analyse du protéome sont différentes de celles du séquençage du génome puisqu'on ne parle pas des mêmes macromolécules (acides nucléiques ≠ acides aminés)



## A. Objectifs et principe de la protéomique :

La protéomique est l'étude à **grande échelle** sans a priori de l'ensemble des 100 000 protéines potentielles d'un organisme, d'un fluide biologique, d'un organe, d'une cellule ou d'un compartiment cellulaire. Actuellement, ces études peuvent se faire sur des tous **petits échantillons** et cela permet une analyse très précise. Les différents objectifs sont :

- Identifier et quantifier les protéines présentes dans un échantillon à un instant T :  
En physiopathologie, cette technique permet une analyse à la fois qualitative et quantitative.  
-On peut à la fois identifier c'est-à-dire détecter la présence de **tel ou tel peptide ou protéine** (aspect **qualitatif**).  
-Et selon les techniques on peut aussi les quantifier c'est-à-dire déterminer la quantité de chaque protéine (aspect **quantitatif**).  
Cela inclut également la description des modifications post-traductionnelles. Cette technique permet finalement de comparer les valeurs obtenues à celles d'une situation dite normale.
- Obtenir des données fonctionnelles :  
Ces protéines peuvent avoir des fonctions différentes et interagissent avec des partenaires différents, donc on peut affiner cette analyse en étudiant leurs **interactions** et leur localisation. On peut notamment **isoler** les différents compartiments de la cellule (par exemple pour vérifier qu'une protéine qui doit agir dans le noyau ne se trouve pas de façon anormale dans le cytoplasme).

La technique la plus utilisée en protéomique est la **spectrométrie de masse**. Cette technique a longtemps été réservée à des laboratoires spécialisés mais est désormais accessible dans des laboratoires diagnostics hospitaliers. Avant on avait besoin de beaucoup de matériel pour réaliser cette étude (par exemple un foie entier de souris), mais la sensibilité a beaucoup augmenté et maintenant on peut travailler sur des petits échantillons (par exemple une biopsie à l'aiguille d'un foie de patient).

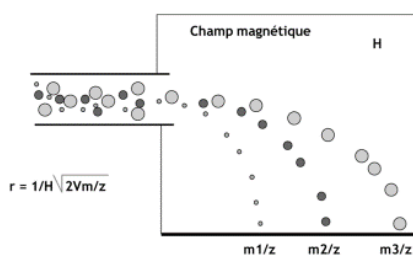
La spectrométrie de masse est une méthode d'analyse structurale permettant **d'identifier des molécules en fonction de la mesure précise de leur masse**.

Principe : déplacement de particules chargées sous l'influence de champs électriques et/ou magnétiques dans une enceinte soumise à un vide poussé pour éliminer les interactions avec les molécules de l'air. Le but est d'observer un déplacement strictement dépendant de la masse de la molécule.

## B. Protocole de la spectrométrie de masse :

*Pour rappel, les protéines peuvent être chargées positivement ou négativement en fonction de leur état d'ionisation. Ces charges sont portées par les **chaînes latérales des acides aminés, non impliquées dans des liaisons peptidiques.***

**Étape 1 : Séparation en phase gazeuse de molécules chargées (ions) en fonction de leur rapport masse/charge (m/z avec la masse exprimée en Da) :**

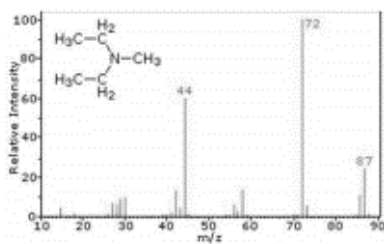


1. Obtention d'un extrait protéique issu d'un tissu ou d'une culture cellulaire.
2. Fragmentation de cet extrait pour obtenir des peptides (plus petits donc plus faciles à analyser).
3. Vaporisation et ionisation des peptides pour qu'ils puissent se mouvoir dans le champ.
4. Soumission à des champs électriques ou magnétiques.
5. La déviation change en fonction de la masse et de la charge des peptides.

La mesure des déviations des rayons par les particules dans le champ magnétique permet donc de déduire leur rapport masse/charge (m/z) et donc leur masse exprimée en Dalton (Da).

A l'issue de cette analyse on obtient une **empreinte moléculaire** de chaque composé du mélange.

## Étape 2 : reconstitution de l'identité des protéines à partir de leur spectre de masse :



La hauteur des pics est intéressante car elle permet l'analyse quantitative, mais c'est la distance entre les pics qui permet d'identifier les protéines. En se basant sur les spectres obtenus, on peut donc regarder si la masse ou la charge de certaines protéines a varié, et donc comparer une situation pathologique avec une situation **de référence**.

La spectrométrie de masse permet d'analyser un mélange d'échantillons complexes, pas une seule molécule mais un **mélange de molécules** contenant de très nombreuses protéines. Et quand on fait ça sur de nombreux échantillons, on se rend compte que les protéines sont **beaucoup plus complexes** et plus variées que ce que l'on pensait.

Ces analyses servent également en pharmacologie, toxicologie (hospitalière, médico-légale, environnementale), ainsi que dans les analyses de biochimie (par exemple en cas d'intoxication bactérienne, plutôt que de mettre un prélèvement en culture pendant 1 semaine, on peut rapidement **identifier le germe par spectrométrie de masse et avoir les résultats dans la journée**).

### III. Modifications post traductionnelles :

Les protéines selon leur configuration et leurs modifications post-traductionnelles jouent un rôle différent. Nous allons donc voir comment la régulation des modifications post traductionnelles peut changer les potentialités fonctionnelles des protéines.

#### A. Définitions et exemples :

Une modification post-traductionnelle est **l'ajout d'un radical** de taille et de nature biochimique variable à une protéine par modification covalente d'un acide aminé composant cette protéine, ayant pour conséquence **la modification de ses propriétés** (activité, stabilité, dégradation, localisation, capacités d'interaction). Ces modifications peuvent être labiles (se faire et se défaire) ou stables. Elles sont réversibles et multiples, c'est-à-dire que chaque protéine peut faire l'objet de plusieurs modifications.

Exemples de modifications :

- Les modifications par changement de structure de la protéine :

Certaines modifications post traductionnelles sont irréversibles, comme le **clivage protéolytique** ou les **ponts disulfures**

- Les modifications par ajout de groupements chimiques :
  - **Phosphorylations** : ajout d'un phosphate sur Sérine, Thréonine ou Tyrosine,
  - **Acétylation** : -CH<sub>3</sub>-CO,
  - Hydroxylation : sur des prolines, par exemple hydroxyproline dans le collagène,
  - Méthylation : -CH<sub>3</sub>,
  - Glycosylation : modifications par des sucres,
  - Glucosaminylation : -O-GlcNac,
  - Farnésylation : modification covalente par des acides gras,
  - **Ubiquitination, SUMOylation** : modifications par d'autres peptides.

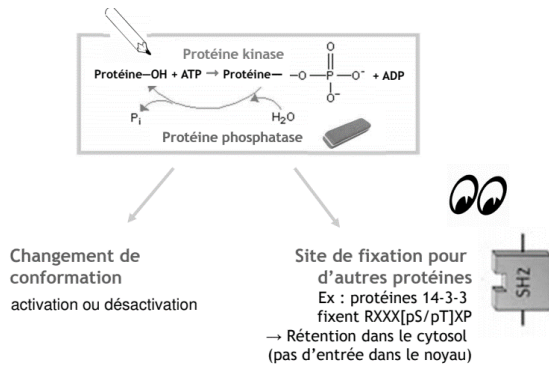
Ces modifications ont un rôle important dans la **régulation** et dans le **contrôle** de la fonction des protéines. Elles peuvent entraîner des **altérations** importantes sur les propriétés fonctionnelles des protéines (Par exemple la phosphorylation peut entraîner des modifications dans la voie de biosynthèse du glycogène). Elles peuvent aussi impacter leurs interactions avec leurs partenaires ou la capacité d'une protéine à interagir en multimères (*problématique car les protéines ne travaillent pas seules, mais toujours avec des copies d'elles-mêmes*).

⇒ On est donc passé de 25 000 gènes à 100 000 protéines potentielles, qui peuvent également subir un grand nombre de modifications post-traductionnelles. **Les modifications post-traductionnelles augmentent donc les potentialités fonctionnelles des protéines.**

Le schéma des modifications post-traductionnelles comporte 3 étapes :

- **Ecrire** : une enzyme effectue la modification post traductionnelle sur la protéine ciblée (ajout d'un phosphate, d'un ubiquitine ...).
- **Lire** : cette modification est détectée puis lue
- **Effacer** : une enzyme enlève la modification post traductionnelle ce qui permet un retour à l'état initial. La protéine peut alors subir de nouveau une autre modification post traductionnelle sans nécessiter une synthèse de novo.

## B. La phosphorylation :



Une protéine kinase effectue une phosphorylation : Elle appose un **PO<sub>4</sub><sup>-</sup>** qui provient de l'ATP sur une tyrosine, serine ou thréonine (cible possédant un OH).

Puis les protéines à domaines SH2 détectent si la protéine est phosphorylée ou non. Et enfin, une protéine phosphatase enlève la phosphorylation. On estime que **30%** des protéines sont régulées par phosphorylation.

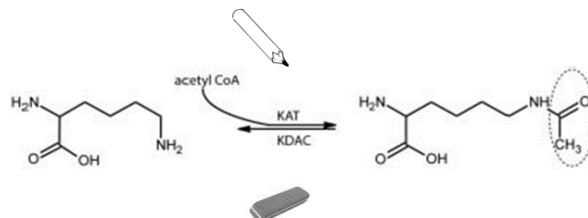
La phosphorylation impacte la **conformation** et la **fonction**, ainsi que les capacités d'interaction (par exemple la protéine 14-3-3 qui peut se fixer sur des domaines SH2, et dont la phosphorylation modifie les interactions et change la localisation cellulaire). Selon les protéines, la phosphorylation peut être activatrice ou inhibitrice : il n'y a pas de règles, chaque protéine doit être étudiée pour le déterminer.

Ce qu'on cherche à faire en médecine, c'est identifier les protéines et leur niveau de phosphorylation. Pour étudier la phosphorylation d'une seule protéine on peut utiliser l'**électrophorèse seule** (observation de la migration électrophorétique due au PO<sub>4</sub><sup>-</sup>), ou **suivie d'un Western Blot** (en utilisant des anticorps spécifiques aux sites de phosphorylation).

## C. L'acétylation et la méthylation des histones :

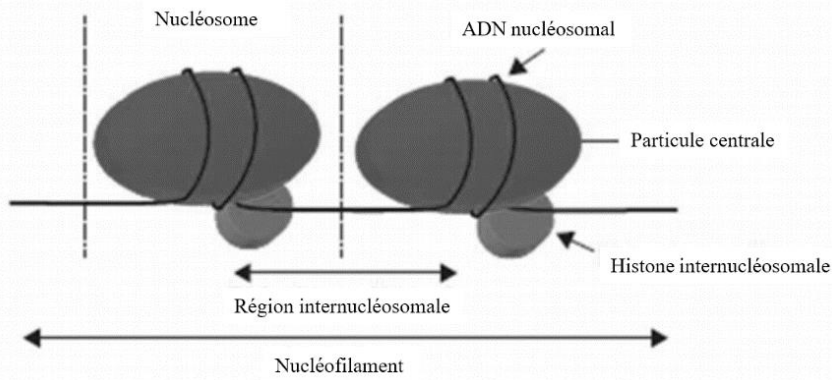
### 1. L'acétylation des histones :

L'acétylation est un mécanisme dynamique et réversible. Elle correspond à l'ajout d'un groupement acétyl -CH<sub>3</sub>-CO par une acétyltransférase (réaction inverse catalysée par une désacétylase). Le donneur du groupement acétyl est l'**acétyl-CoA**. Cet acétyl est pris en charge par des enzymes d'acétylation et est transféré sur la chaîne latérale d'une lysine ou d'un acide aminé en N-Ter. *A noter que l'acétylation favorise un **état actif** de la transcription.*



La lysine est un acide aminé basique chargé positivement au pH physiologique. L'acétylation neutralise cette charge positive et modifie la taille de la chaîne latérale des lysines, ce qui entraîne un changement de conformation et une modification de l'interaction avec les molécules cibles.

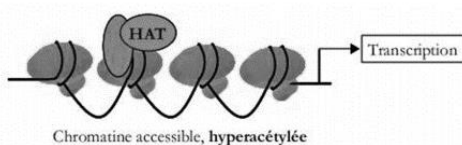




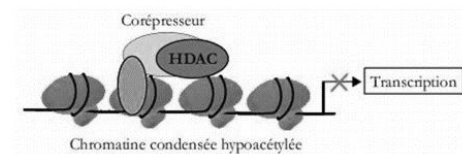
L'acétylation a dans un premier lieu été identifié au niveau des **histones**.

L'histone est une protéine qui organise la chromatine. Les modifications post-traductionnelles des histones sont majeures pour expliquer la structure de la chromatine et donc l'expression des gènes. Un nucléosome est un enroulement de 147 nucléotides autour d'un octamère de 2\*4 histones (H2A H2B H3 H4).

La chromatine possède un dilemme : elle doit être suffisamment « rangée » pour rentrer dans le noyau mais pas trop pour permettre la transcription. L'acétylation permet de résoudre ce dilemme en **régulant la compaction** et donc l'expression des gènes (transcrit ou non). Cette régulation se passe au niveau des chaînes latérales de lysine (très présentes dans les histones), qui sont responsables des interactions électrostatiques avec l'ADN (chargé négativement). Le **degré d'interaction** entre l'ADN et le noyau de l'histone contrôle la transcription :



- L'**hyperacétylation** des histones par des **HAT** (histones acétyltransférases) permet de relâcher la pression électrostatique de l'histone sur l'ADN qui se décompacte et devient accessible en neutralisant la charge (+), ce qui **favorise** la transcription.



- La **désacétylation** des histones par les **HDAC** (histone désacétylase) retire le groupement acétyl des lysines qui redeviennent chargées (+). Cela augmente l'interaction avec la chromatine et la compaction, ce qui entraîne une **répression** de la transcription.

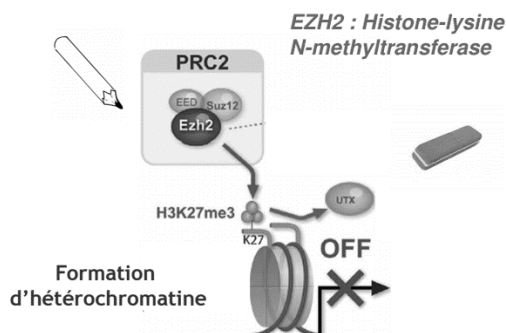
=> Les modifications post-traductionnelles des histones impactent fortement la signalisation et donc **l'expression intracellulaire des gènes**.

Ici la professeure a rajouté pour notre culture : c'est la spectrométrie de masse qui a permis de caractériser l'acétylation de certaines histones

## 2. La méthylation des histones :

Les histones peuvent également être méthylées par des histones méthyl transférase (HMT) :

- 2) Les enzymes **PRMT** méthylent les **arginines**.
- 3) Les protéines à domaines **SET** méthyles les **lysines** => problème car on a deux signaux (acétylation et méthylation) possibles sur un même acide aminé. Ce sont les conditions cellulaires qui déterminent le signal à transmettre, **la communication** entre les modifications post-traductionnelles est donc très importante.



Le complexe répresseur **Polycomb (PRC2)** est un complexe de contrôle de la compaction de la chromatine. PRC2 contient EZH2 qui est une HMT à domaine SET et qui est responsable de la méthylation des histones. La tri méthylation de la lysine n°23 sur l'histone H3 entraîne une compaction de la chromatine et donc une répression de la transcription (contrairement à l'acétylation, la méthylation favorise donc un état **répresseur** de la chromatine : l'**hétérochromatine**, ce qui défavorise la transcription.)

La **déméthylation** est effectuée des **déméthylases**.

Toutes ces modifications vont être spécifiquement reconnues par différents domaines :

- Les **bromodomains** qui reconnaissent spécifiquement les lysines **acétylées** et qui sont donc **présentes** dans des co-activateurs transcriptionnels.
- Les **chromodomains** qui se lient aux lysines méthylées et sont donc **présentes** dans des protéines induisant la formation de l'hétérochromatine.

Ces enzymes de régulation peuvent subir des mutations héréditaires ou somatiques qui entraînent certaines pathologies chez l'Homme (cancers).

Modification	Regulator	Disease caused by		
		Hereditary mutation	Somatic mutation	Environmental change
DNA methylation	DNMT1	HSNE1	Cancer	Cancer, schizophrenia
H3K27 methylation	EZH2	Weaver syndrome	Cancer	Cancer
H3K4 methylation	MLL2	Kabuki syndrome	Cancer, CHD	Cancer
Histone acetylation	CBP, P300	Rubinstein-Taybi syndrome	Cancer	Cancer, aging, neurodegeneration
Chromatin remodeling	CHD7	CHARGE syndrome	ASD, CHD, cancer	Unknown



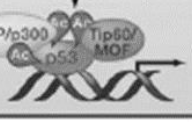
### 3. Le code des histones :

Dans un nucléosome, les **4 histones** (H2A, H2B, H3, H4) peuvent subir une multitude d'actions (elles peuvent être acétylées, méthylées, partenaires de facteurs à bromodomaine ou à chromodomaine). On a donc à la base 4 gènes qui codent pour 4 protéines, mais en fonction de ce que subit chaque histone, on aboutit à une **situation spécifique**. Le « code des histones » regroupe les différentes modifications post traductionnelles que peuvent subir les histones et permet de comprendre quels protéines sont modifiées dans ces modifications post traductionnelles : il va y avoir **une lecture intégrée et combinatoire** de chaque modification post traductionnelle.

#### D. Acétylation de p53 :

p53 est une protéine très instable, cible de modifications post-traductionnelles, et un facteur de transcription qui régule la **mort cellulaire**. On n'a pas besoin de p53 dans une cellule saine, donc elle est dégradée physiologiquement en moins de 20 min par Mdm2. Par contre en cas de stress (cassure d'ADN, stress mitogénique), p53 est **stabilisée et** activée par acétylation. Cette acétylation a lieu sur des lysines mais celles -ci sont également le lieu d'autres modifications : il a donc une **compétition** pour les sites. Le fait d'acétyler une lysine empêche donc son ubiquitination.

*Le nom des protéines fixées sur l'ADN via p53 dans la première colonne du tableau n'est pas à savoir.*

⚡ DNA damage & other types of stress	Transcriptional targets	Biological consequence	p53 modification requirement
<b>A</b> 	<b>MDM2</b>	Feedback and cell survival	Acetylation is not required
<b>B</b> 	<b>P21</b>	Growth arrest DNA repair	Partial acetylation is sufficient
<b>C</b> 	<b>BAX</b>	Irreversible apoptosis	Specific acetylation is required

Plusieurs lysines peuvent être acétylées donc il existe un gradient d'acétylation. Et il existe un « **code p53** » qui influe sur les effets de p53 selon les niveaux d'acétylation :

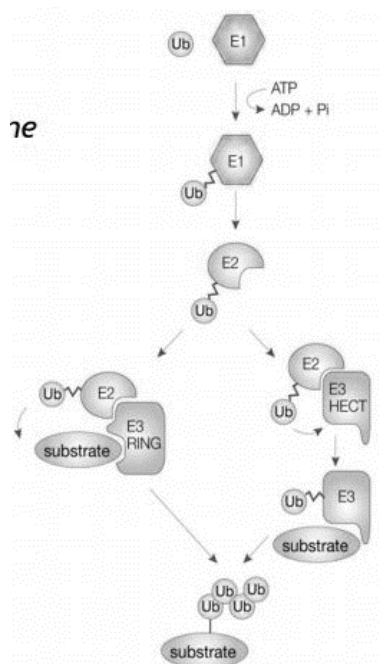
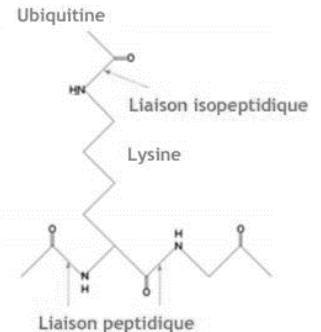
- Pas d'acétylation : dégradation rapide par Mdm2 (feedback)
- Mono ou double acétylation : bloque son ubiquitination + active p21 qui entraîne l'arrêt du cycle cellulaire, ce qui permet de réparer les dommages de l'ADN dus au stress
- Triple acétylation : active BAX qui induit l'apoptose.

#### IV. Dégradation des protéines par le système ubiquitine-protéasome :

##### A. Le système ubiquitine :

Les protéines peuvent être également la cible de **peptides**. Cette modification cible des résidus particuliers notamment les lysines au niveau de nombreuses protéines cellulaires. L'ubiquitine est un peptide de 76 acides aminés qui est **très conservé** au sein des espèces, cela signifie que c'est une signalisation importante dans le développement et dans l'émergence d'espèces stables dans le temps biologique. C'est un petit peptide **ubiquitaire** qui contient notamment 7 lysines, **K dans le code à une lettre**, (*information importante dans la suite de ce cours*). Il va falloir réussir à le **coller** à des protéines cibles.

L'ubiquitination est la création d'une liaison particulière **isopeptidique transversale réversible** avec le NH<sub>2</sub> latéral d'une lysine par le COOH de la glycine en C-terminal de l'ubiquitine. La liaison isopeptidique est une liaison amide, semblable à une liaison peptidique sauf qu'on a l'implication d'une **chaîne latérale**.



L'ubiquitine est un signal d'activation pour certaines protéines et la **polyubiquitination** est un signal de destruction. La fixation de l'ubiquitine est régulée car sinon il y aurait une dégradation constitutive d'un grand nombre de protéines.

La fixation de l'ubiquitine fait intervenir plusieurs étapes car chacune est un point de contrôle et de régulation. Il y a une action de 3 familles d'enzymes :

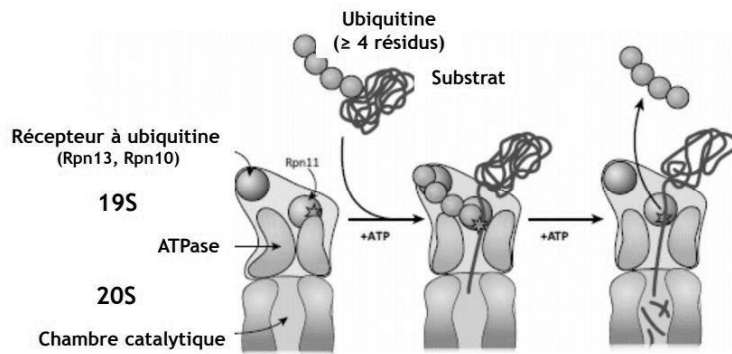
- Les enzymes d'**activation E1** ont pour rôle d'activer l'ubiquitine et de se lier avec de manière active (en consommant de l'énergie sous forme d'ATP). Mais E1 est incapable de transférer l'ubiquitine sur une protéine cible.
- E1 transfère donc cette ubiquitine à **E2** qui est une enzyme de **conjugaison**. Mais E2 est incapable de reconnaître le substrat de fixation de l'ubiquitine.
- **E3** (Ubiquitine **ligase**), peut reconnaître le substrat avec deux grandes classes d'enzymes, (par exemple : les E3 RING qui peuvent reconnaître un substrat et E3 avec HECT peut en reconnaître un autre) ce qui permet d'avoir une grande spécificité. E3 **transfère** donc l'ubiquitine sur la protéine cible.

Dans la cellule, le nombre d'enzymes qui régulent ces étapes varie et cela crée un gain de diversité. On a, chez les mammifères,

- 1 seule enzyme E1
- 40 enzymes E2
- 650 à 750 enzymes E3 ce qui entraîne une **spécificité** dans le choix du substrat.

##### B. Le protéasome :

*Pour rappel, le protéasome est un complexe macromoléculaire permettant la dégradation spécifique des protéines intracellulaires en petits peptides. Il est formé de 2 sous-unités régulatrices 19S et d'une sous-unité catalytique 20S.*



Le protéasome possède 2 parties. La première partie du protéasome va reconnaître l'adressage, les protéines polyubiquitinées. Elle possède un récepteur à l'ubiquitine à sa périphérie pour un contrôle des entrées. Ensuite moyennant la consommation d'énergie, il va y avoir tout d'abord clivage de la chaîne d'ubiquitine, pour la relarguer et éviter ainsi sa dégradation par le protéasome. Elle va être recyclée pour éviter de devoir la resynthétiser.

La protéine à dégrader continue sa progression dans la deuxième partie du protéasome au sein de l'unité 20S, la chambre catalytique, qui va cliver la protéine tous les 5 à 10 acides aminés. On a ainsi une sélection des peptides à cliver.

Le couplage de l'ubiquitination à la protéolyse confère à la cellule eucaryote la capacité d'éliminer **sélectivement** certaines protéines selon les **conditions physiologiques**.

## 1. Rôles importants du système ubiquitine- protéasome :

- 1) **Contrôle qualité des protéines :** Elimination des protéines malformées (repliées, mutées) ou endommagées (modification enzymatiques), protéines mal localisées : association du ribosome, des protéines chaperonnes et du protéasome. C'est un contrôle qualité quasiment co-translationnel. Dégradation de 4g protéines /kg / jour.
- 2) **Mécanisme de dégradation rapide des protéines :** Renouvellement des composés cellulaires après utilisation (turnover protéique), demi-vie variable des protéines. C'est important car cela donne des acides aminés, les briques de construction des protéines à resynthétiser. Cela permet l'adaptation rapide de la cellule à de nouvelles conditions.
- 3) **Production d'acides aminés :** selon les besoins énergétiques on peut libérer les acides aminés pour pouvoir utiliser d'autres sources d'énergie (substrat de néoglucogenèse), notamment en cas d'atrophie musculaire post immobilisation.
- 4) **Fonctionnement du système immunitaire et son adaptation :** il a besoin de produire des antigènes toujours plus variables d'un jour à l'autre. Ce sont des antigènes protéiques et il faut des acides aminés pour les produire.
- 5) **Régulation de certaines fonctions cellulaires par protéolyse.** Ex : Rôle de la dégradation de la cycline B dans le contrôle de la progression du cycle cellulaire quand on en a plus besoin. Ex : Régulation de p53 par Mdm2.

Pour p53, son niveau d'ubiquitination varie selon le stress, comme on l'a vu pour l'acétylation. *Attention sur le diaporama il est écrit 1 min mais la professeure a répété plusieurs fois qu'il s'agissait de quelques minutes à quelques heures après le stress et l'a corrigé pour ce schéma à l'oral.*

Comme vu précédemment, en absence de stress cellulaire sa dégradation constitutive est permise par son partenaire Mdm2, qui est l'équivalent de l'E3 ligase. ~~Mdm2 est quasi-spécifique de p53, il agit sur d'autres protéines mais c'est moins important.~~ Si il y a du stress, il va y avoir acétylation, Mdm2 n'est plus au contact de p53 et n'induit plus l'ubiquitination. p53 va donc être stabilisée et va pouvoir induire la réponse transcriptionnelle.

## 2. Inhibiteurs et activateurs du protéasome :

Ces rôles cités précédemment sont importants à savoir car ils sont importants en thérapie, diagnostic, ou en compréhension des mécanismes physiopathologiques. Avec le protéasome, il y a une question d'équilibre qui entre en jeu. On a des situations où le protéasome va être très riche en protéines à dégrader et est ce qu'il est capable de les dégrader ?

Quand on regarde la charge du protéasome : cette charge augmente quand il y a plus de synthèse protéique, quand il y a des protéines mal repliées, quand les protéines chaperonnes ne fonctionnent pas bien, quand on a des situations d'inflammation.



A l'inverse, il y a des facteurs qui déchargent le protéasome, quand on a une augmentation de la déubiquitination.

Il existe des facteurs favorisant la capacité du protéasome : son assemblage, les différentes sous unités qui doivent être fonctionnelles et l'utilisation d'activateurs.

A l'inverse, il y a des facteurs intrinsèques **endogènes** qui impactent le protéasome : le vieillissement, la sénescence et l'utilisation d'inhibiteurs du protéasome.

L'accumulation de protéines à dégrader entraîne une surcharge de mauvais acteurs, de mauvais opérateurs et cela va créer un signal fort pour l'**apoptose**. Les inhibiteurs du protéasome ont été testé dans des tumeurs (car elles ont besoin de beaucoup de protéines pour grandir assez vite même si elles sont mal repliées).

Si on fait pencher la balance par un inhibiteur du protéasome, le protéasome est chargé de protéines à dégrader mais il n'est **pas actif**. Cela provoque donc une boucle de prolifération des protéines et l'absence de dégradation des protéines anormales génère un stress majeur de la cellule (un stress du réticulum endoplasmique qui va, avec différentes autres actions comme la suppression de NF-κB et la stabilisation de p53 et de protéines pro-apoptotiques, mettre en place une signalisation pro apoptotique pour déstabiliser et éliminer les cellules tumorales).

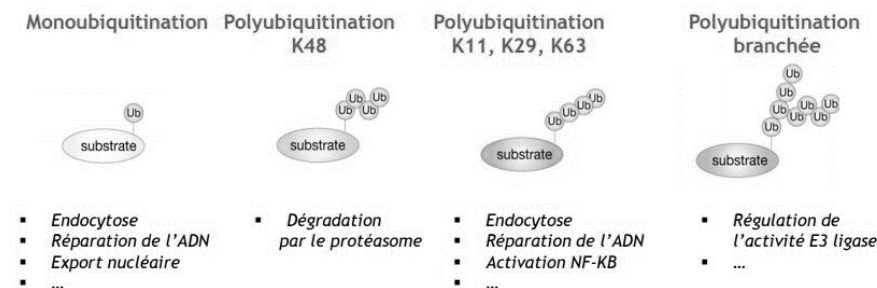
Ce sont des thérapies utilisées de façon majeure, souvent couplées à d'autres chimiothérapies, notamment dans le traitement du myélome multiple. Cette pathologie implique la prolifération d'un plasmocyte anormal dans la moelle osseuse qui sécrète des immunoglobulines (anticorps) anormales.

### C. L'ubiquitination n'est pas toujours un signal de dégradation :

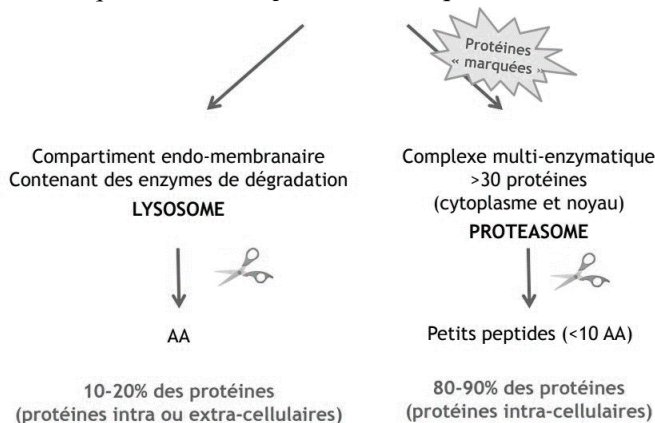
Les protéines peuvent être ubiquitinylées de différentes manières :

- 1) Monoubiquitinylées.
- 2) **Multi-ubiquitinylées** sur plusieurs lysines.
- 3) Polyubiquitinylées en réseau/en arborescence : selon les chaînes d'ubiquitine le signal sera différent.

Le vrai signal d'adressage au protéasome c'est la **polyubiquitination en réseau** : plusieurs ubiquitines ramifiées sur une seule lysine. Ceci est possible grâce à la présence de lysine dans l'ubiquitine (liaison des ubiquitines entre elles par leurs lysines respectives). **Les arborescences K48 sont des signaux de dégradation par le protéasome, et d'autres arborescences peuvent induire d'autres mécanismes.**



Sur cette diapo, la professeure n'a pas parlé de la dégradation des protéines non marquées et est peu entrée en détail sur l'ubiquitination des protéines marquées.



La dégradation des protéines « non marquées » se fait dans le **lysosome**, un compartiment endomembranaire qui contient des enzymes de dégradation. Les protéines sont découpées en acides aminés. Cela concerne **10 à 20%** des protéines intra- ou extracellulaires.

Les protéines marquées par polyubiquitination sont ensuite envoyées au protéasome qui les découpe en fragments de 10 à 20 acides aminés. Cela concerne les complexes multienzymatiques (supérieurs à 30 protéines), du cytoplasme et du noyau soit **80 à 90%** des protéines uniquement **intracellulaires**.



Pour pouvoir caractériser ces différents types d'ubiquitination, on a fait des tests in vitro sur des cultures cellulaires : on a utilisé des inhibiteurs du protéasome afin que les protéines ubiquitinylées ne soient pas dégradées et on a observé que les formes qui s'accumulent le plus rapidement sont celles qui impliquent les lysines 11, 29, 48, 63 et d'autres ne sont pas modifiées. Les ubiquitines qui s'accumulent sont celles qui ont le **plus fort impact** dans la dégradation.

Multiplicité des signaux : en faisant varier le niveau d'ubiquitination (mono, poly ou multi) et les lysines impliquées, on obtient des situations différentes. *Comme pour les histones, on a une situation à paramètres multiples.* Il existe donc un **code ubiquitine**.

#### D. Les systèmes pseudo-ubiquitine :

Le système ubiquitine a été dupliqué un très grand nombre de fois : on connaît actuellement des dizaines de peptides du même type appelées **pseudo-ubiquitines** (« ubiquitine-like »). Les systèmes pseudo-ubiquitine fonctionnent à peu près de la même manière que les ubiquitines. Le système « ubiquitine-like » le plus connu et le plus important est le système **SUMO** (**S**mall **U**biquitin **M**Odifier), qui possède de nombreuses cibles et de nombreuses fonctions. Le peptide SUMO est encore **mieux conservé** au sein des espèces que l'ubiquitine.

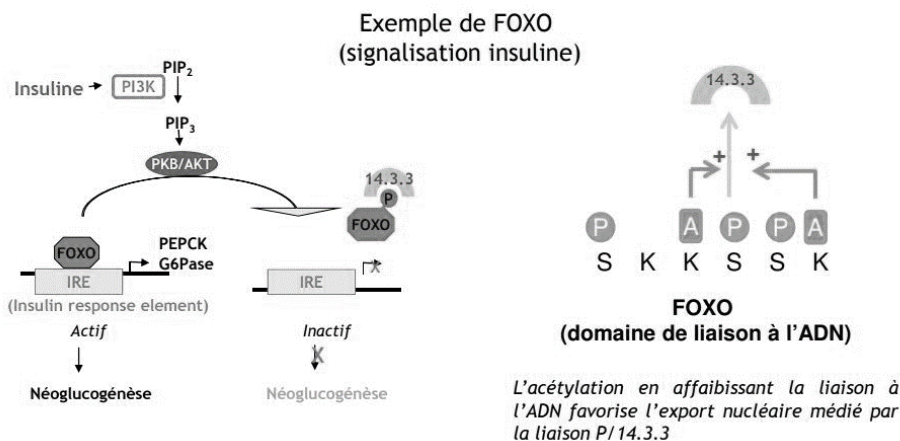
#### Ressemblances avec l'ubiquitine :

- SUMOylation de lysines,
- Contrôle de la SUMOylation par 3 enzymes,
- Mono ou multi-SUMOylation,
- La SUMOylation peut recruter de l'ubiquitine ce qui montre une **communication** entre les modifications post-traductionnelles.

Toutes ces modifications ciblent souvent les mêmes acides aminés et les modifications sont souvent exclusives donc on a des **régulations croisées**.

Il existe des **interactions** entre les modifications post-traductionnelles et elles peuvent avoir des effets :

- Antagonistes : méthylation/acétylation ou méthylation/phosphorylation **exclusives** sur des histones.
- Coopératifs : exemple de FOXO => acétylation sur une lysine et phosphorylation sur la Sérine adjacente. Cela facilite l'interaction entre FOXO et son partenaire 14-3-3 (l'acétylation favorise la phosphorylation qui favorise l'interaction). Il y a donc 2 signaux qui vont dans le même sens mais en ciblant 2 acides aminés différents. *D'après le schéma, FOXO est un intervenant dans la signalisation de l'insuline. Lorsqu'il est phosphorylé, il se lie avec son partenaire 14-3-3 et ne peut pas se lier au gène IRE. Ce gène ne peut donc pas être transcrit et dans ce cas il n'y aura pas de néoglucogénèse. Dans le cas contraire, FOXO se lie au gène IRE et entraîne sa transcription, rendant possible la néoglucogénèse.*
- Exclusifs (pas antagonistes ni coopératifs) : exemple de P53 et Mdm2 => quand P53 est acétylé, Mdm2 ne peut pas interagir avec pour l'ubiquitiner.



## V. Dans un contexte médical :

Dans cet exemple, on a un génome qui donne différents transcrits pour des protéines, dans une situation normale. Si on provoque une exposition à un carcinogène notamment, cela entraîne une modification de la cellule et progression de la tumeur avec des altérations au niveau de certains exons et de certains gènes. Au fur et à mesure de la progression dans une pathologie on peut accumuler différentes altérations qui vont forcément avoir des impacts sur nombreux acteurs protéiques.

Ces altérations progressives sont autant de cibles thérapeutiques. On peut donc essayer différentes thérapies pour réactiver certains exons, réguler l'expression d'un gène qui n'aurait pas dû être dérégulé, et aussi cibler certaines protéines grâce à des anticorps.

Toutes ces connaissances sur la diversité des protéines vont nous donner des informations sur des situations pas nécessairement tumorales, mais sur des situations physiopathologiques plus larges.

### Exemple : étude américaine centrée sur la protéomique différentielle en fonction du vieillissement.

A un âge donné est ce que certaines personnes vieillissent mieux que d'autres ? On parle principalement de déclin cognitif : pourquoi certaines ont toutes leurs capacités à 75 ans et d'autres pas ?

C'est très problématique actuellement avec le vieillissement de la population. Cette étude a corrélié la **protéomique** (pas la génétique) à la situation cognitive du patient au moment du prélèvement. Principalement sur des autopsies car cette étude a nécessité le prélèvement de tissu cérébral. Ils ont fait une analyse comparative, ils ont placé les 3900 différentes protéines qu'ils avaient identifiées comme différentiellement **exprimées** dans différents clusters pour identifier des grandes familles, ~~par exemple des familles qui sont impliqués dans les giantomes, dans différents endroits. Ils ont identifié de façon corrélée (et non causalité), que les personnes qui avaient un déclin cognitif plus important~~ avaient plutôt des protéines de l'inflammation et des protéines de l'apoptose. A l'inverse, une diversité protéique synaptiques et une diversité des protéines mitochondriales qui peuvent être associées par corrélation à une meilleure évolution cognitive.

**Ces données identifient l'importance de la diversité des protéines exprimées dans la dynamique cellulaire et la fonctionnalité tissulaire.**

## **Conclusion :**

Le protéome est une entité dynamique et complexe :

- Un même gène peut donner naissance à plusieurs protéines par épissage alternatif : le protéome contient donc un nombre beaucoup plus important de protéines que le génome ne contient de gènes
- De nombreuses modifications post-traductionnelles augmentent les potentialités fonctionnelles des protéines
- Au sein de chaque cellule, le contenu de protéines se modifie en permanence en fonction des conditions intra ou extra cellulaires
- Explique la grande diversité des activités cellulaires, surprenante au regard du faible nombre de gènes

A la fin de ce cours vous devez connaître :

- Les 5 modes d'épissage alternatif
- La définition d'une modification post-traductionnelle
- La « carte de visite » des principales modifications post-traductionnelles
- Le principe, mécanisme moléculaire et différentes conséquences fonctionnelles de l'ubiquitination

=>Vous devrez avoir compris : La notion de modification post-traductionnelle et son intérêt dans l'augmentation des potentialités fonctionnelles des protéines.