



## **RONÉOS DFGSP2 2020 - 2021**

### **UE PL2.18 : BIOMOL 9-10**

**Date : 20/10/2020**

**Enseignant : DUFOURCQ Pascale**

**Plage horaire : 10h30-12h30**

**N°ISBN : 978-2-37366-088-3**

Ronéistes

MARMET Kelly – kelly.mrt.240@gmail.com

AGEL Nicolas: inicoagel@icloud.com

## **Production de protéines recombinantes à usage pharmaceutique : Les biomédicaments**

### **Plan du cours :**

#### **I. Production de protéines recombinantes à usage pharmaceutique : les biomédicaments**

##### **A- Introduction**

##### **B- Génie génétique et culture de cellules ou micro-organismes**

###### **1- Rappels : expression protéique**

#### **II – Génie génétique et culture de cellules ou micro-organismes**

##### **A- La boîte à outils**

###### **1- Les vecteurs**

###### **2- Choix de l'hôte**

###### **3- Insertion du vecteur dans l'hôte**

###### **4- Purification des protéines**

##### **B- Production de protéines recombinantes « biomédicaments »**

###### **1- Système procaryote: ex Insuline, hormone de croissance**

###### **2 - Système eucaryote**

\*\*\* a fait l'objet d'une question à l'examen

## I. Production de protéines recombinantes à usage pharmaceutique : les biomédicaments

### A- Introduction

Ces biomédicaments sont des **protéines recombinantes** : ce sont des médicaments produits à partir du génie génétique. Une protéine recombinante est produite au sein d'organismes vivants : procaryotes (bactéries), eucaryotes (levures ou cellules de mammifères). Cette cellule ou bactérie a été modifiée génétiquement, on a bien utilisé l'ADN recombinant. On aura soit une expression **homologue** (= rare) soit **hétérologue** (=la cellule qui produit cette protéine est différente de la cellule d'origine).

*Exemple : synthèse d'une hormone de croissance humaine dans la bactérie, donc on met un gène qui code pour une protéine humaine dans un système qui n'est pas une cellule humaine : **système hétérologue**.*

La majorité des médicaments sont issus de la synthèse chimique ou de l'extraction de plantes naturelles. Cependant, on peut essayer de produire des médicaments qui ont une structure protéique. Une protéine est extraite à partir de son environnement naturel (à partir du sang, urines...) mais dans ces tissus ou fluides provenant d'animaux ou d'humains, les **quantités** sont très **faibles** voire insuffisantes pour faire un médicament, il est parfois impossible de réaliser **l'extraction**. Les protéines peuvent être difficiles à **purifier** ou **isoler** par des techniques conventionnelles et puis on peut avoir un **risque biologique** : elles peuvent contenir des toxines, virus, prions.

*Exemple : dans les années 80, les hormones de croissance (prélevées de l'hypophyse de cadavre humain) étaient administrées aux enfants atteints de nanisme harmonieux. Jusqu'au jour où l'on s'est rendu compte que de nombreux enfants mourraient de la maladie à prion, dû à l'extraction de cette protéine contenant aussi le prion.*

De plus, une protéine est trop complexe pour être synthétisée par la chimie. Donc lorsqu'on parle de biomédicament en comparaison aux médicaments classiques, ils sont différents de par leur structure et leur production. Ils sont produits par biotechnologie et génie génétique. Ces médicaments doivent avoir une **quantité, qualité et biosécurité**. On peut aussi modifier les séquences pour essayer d'améliorer la pharmacocinétique. A partir de ce génie génétique, on peut créer des molécules complexes et de haut poids moléculaire, qui ne sont pas accessibles à la chimie : les protéines. On a la possibilité d'améliorer les capacités d'une protéine par le bio-engineering.

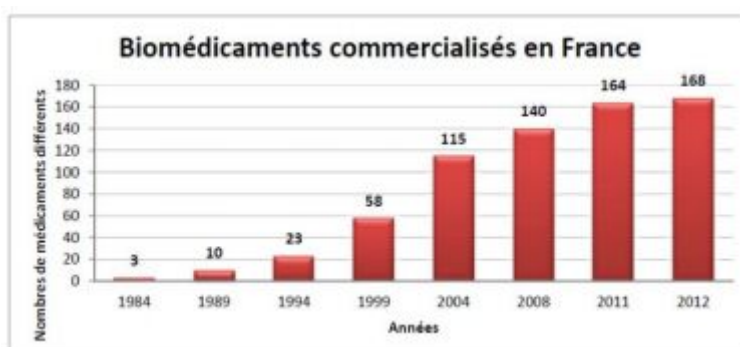
Premier médicament issu du génie génétique sur le marché : **insuline**  
Avant, l'insuline était extraite à partir de porc mais depuis les années 80, le premier biomédicament sorti est l'insuline synthétisée grâce aux progrès de la biotechnologie. On a pu isoler le gène de l'insuline humaine qu'on a pu cloner pour produire cette protéine recombinante.

Année agrément FDA	protéine
1982	Insuline (Genentech)
1985	Hormone de croissance (1988 en france)
1986	Tissue Plasminogen Activator (Genentech)
1989	Erythropoïétine
1994	Glucocérébroside
1998	Anti TNF-alpha

Autre révolution dans les années 2000, on a pu synthétiser des molécules encore plus complexes : **les anticorps**. On a toute une série d'anticorps qui sont produits par génie génétique. Aujourd'hui il y a à peu près **200 biomédicaments sur le marché**. 200 c'est peu par rapport à la synthèse chimique mais **80 % des nouveaux médicaments** sont des **biomédicaments**. Les biomédicaments sont utilisés pour des maladies très variées sur des catégories très larges.

## Différentes classes de Biomédicaments

Classe de protéines	Protéines recombinantes	Domaines thérapeutiques
<b>Hormones</b>	Insuline	Diabète
	Hormone de croissance	Déficit de croissance chez l'enfant
<b>Facteurs sanguins</b>	Facteur de coagulation (VII, VIII, IX)	Troubles de la coagulation
	Antithrombine III	
	Thrombolytiques, rtPA	
<b>Interférons</b>	IFN- $\alpha$ , $\beta$ , $\gamma$	Cancers
<b>Interleukines</b>	Anticorps anti IL-1	Maladies inflammatoires
	Anticorps anti IL-2	Cancers
<b>Facteurs de croissance</b>	EPO	Anémies
	G-CSF	neutropénies
<b>Cytokines</b>	Anti TNF- $\alpha$	Maladies inflammatoires
<b>Enzymes</b>	Imiglucerase	Maladies génétiques de surcharge lysosomales: Gaucher, Fabry, Hurler etc.
	Agalsidase	
	Laronidase	
	Alglucosidase alfa	
	Idursulfase	
	Galsulfase	
<b>Anticorps monoclonaux</b>		Cancer, maladies inflammatoires



On va voir 2 systèmes : production dans des systèmes procaryotes et eucaryotes, dans des systèmes cellulaires ou des micro-organismes (bactérie, levure et cellules de mammifères) et puis on finira par la production à partir d'animaux ou de plantes transgéniques.

## **B- Génie génétique et culture de cellules ou micro-organismes**

### **1-Rappels** : expression protéique

Pour la synthèse d'une protéine, il nous faut un gène, on active la transcription grâce aux facteurs de transcription qui se fixent sur le promoteur actif, il nous faut un ARNm stable puis une traduction et une maturation = **modifications post-traductionnelles**. On veut une protéine mature qui aura subi toutes ces étapes.

#### **Structures des protéines :**

- primaire : séquence en acides aminés
- secondaire : hélices alpha, feuillets beta...
- tertiaire : conformation 3D
- quaternaire : association de plusieurs peptides entre eux pour donner une protéine mature.

Ces modifications post-traductionnelles sont importantes pour avoir une protéine mature : ponts disulfures, glycosylation, phosphorylation...

C'est important car cela permet **l'activité biologique** de notre protéine et toutes ces modifications peuvent intervenir dans la reconnaissance d'un récepteur, reconnaissance d'une adhésion cellulaire, modulation métabolique d'enzyme, modifications métaboliques, transport et adressage de protéines et peuvent avoir un rôle structural. Il faut faire attention à la fonctionnalité de la protéine quand on parle de biomédicament.

L'insuline a 2 chaînes (A et B) reliées par des ponts disulfures et un pont disulfure intra-chaîne (sur la A). Sans cette structure elle est inactive.

Autre possibilité, ce sont les glycosylations qui sont très importantes, de différents types : N ou O-glycosylations, qui permettent la stabilité, l'activité de la protéine et peuvent donner des réactions immunogènes. Attention, la glycosylation n'est pas la même dans toutes les cellules. Elle est pratiquement nulle ou très simple dans le système bactérien (transcription et traduction simultanées) mais très évoluée dans le système eucaryote

→ Système **procaryote** : ne fait pas ou peu de modifications

→ Système **eucaryote** : fait à peu près toutes les modifications (glycosylation, pont disulfure, repliement...).

Attention au choix de la cellule hôte.

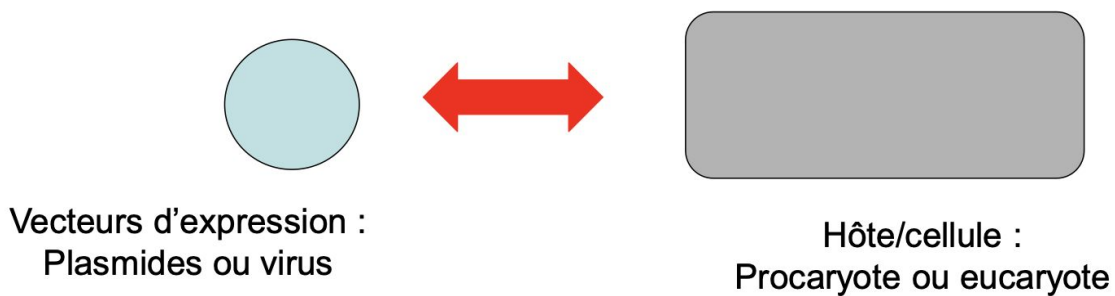
## II – Génie génétique et culture de cellules ou micro-organismes

### A- La boîte à outils pour produire des protéines recombinantes.

On a vu qu'il nous faut un vecteur **d'expression**, contenant l'ADN d'intérêt. Ce vecteur va être intégré dans « une usine » qui permet de produire cette protéine : c'est ce qu'on appelle la **cellule hôte** (bactérie, levure, etc). Ensuite on la cultive et on purifie la protéine d'intérêt que l'on a récupéré.

Pour cela, il va falloir choisir le vecteur d'expression, qui sera soit un plasmide soit une particule virale. Avec celui-ci, il faudra choisir la cellule hôte qui pourra produire la protéine d'intérêt. Le couple vecteur/cellule hôte est donc à définir.

Attention il y a des **compatibilités** qui doivent exister entre le vecteur et l'hôte: donc pour chaque couple vecteur hôte, on aura des avantages et des inconvénients. Ce choix est dépendant des **caractéristiques de la protéine** que l'on veut produire.



On ne choisira pas la même chose si la protéine est très simple, si elle est glycosylée ou si elle a un certain nombre de ponts disulfures.

D'autre part, on ne va pas choisir la même chose si on veut produire une petite ou une énorme protéine. Il faut donc faire attention aux modifications post traductionnelles et notamment aux glycosylations et la formation des ponts disulfures.

Enfin, il va falloir choisir la cellule hôte par rapport au fait que l'on peut produire des protéines toxique pour l'hôte ou le produit. Il faudra donc adapter les systèmes de production par rapport à ce que l'on veut produire.

## 1. Les vecteurs

Si on veut produire une protéine, il faudra utiliser un vecteur **d'expression** (qui est différent du vecteur de clonage, mais tous possèdent les mêmes séquences). Ils doivent contenir :

- Une **origine de répllication**
- Un **multisite de clonage** : c'est l'endroit où l'on va insérer la séquence que l'on veut introduire.
- Un **marqueur de sélection**: il donne une résistance à un antibiotique. Il sert à garder uniquement la bactérie qui a que le plasmide recombinant qui nous intéresse.

Cela va nous permettre de produire le vecteur que l'on va utiliser.

Le vecteur expression doit avoir plein de sites permettant une **transcription efficace** et une **traduction** de l'ARN. Il doit avoir toutes les informations nécessaires pour que la cellule puisse transcrire le gène en ARNm, qui sera par la suite traduit en protéine. Pour cela, il faut qu'il y ait:

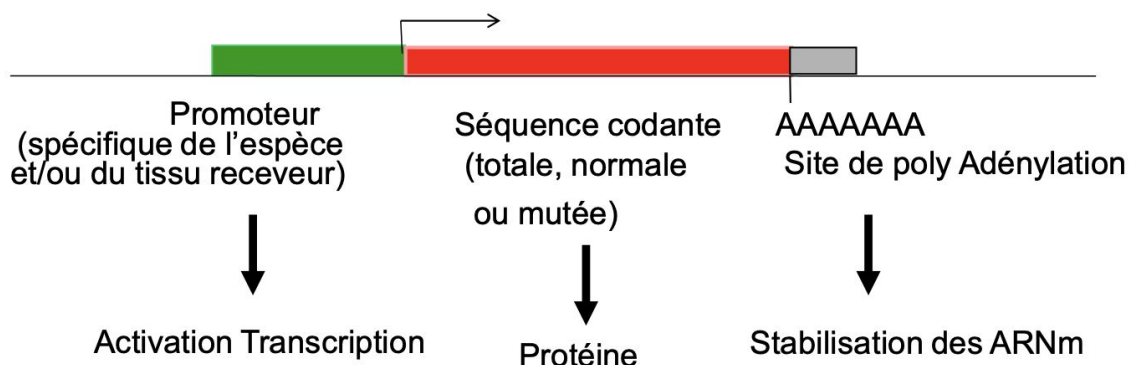
- Un promoteur **devant** la séquence. (*Vu en paces*)
- Une séquence **régulatrice**, plutôt **activatrice**
- Des séquences de **terminaison** de la transcription.

Enfin, pour que la **traduction** soit efficace, il faut :

- Ajouter une séquence de **polyA** (qui stabilisera et protégera l'ARNm).
- Une séquence **d'initiation** de la traduction avec des liaisons au ribosome.
- Une **région UTR** de longueur minimale

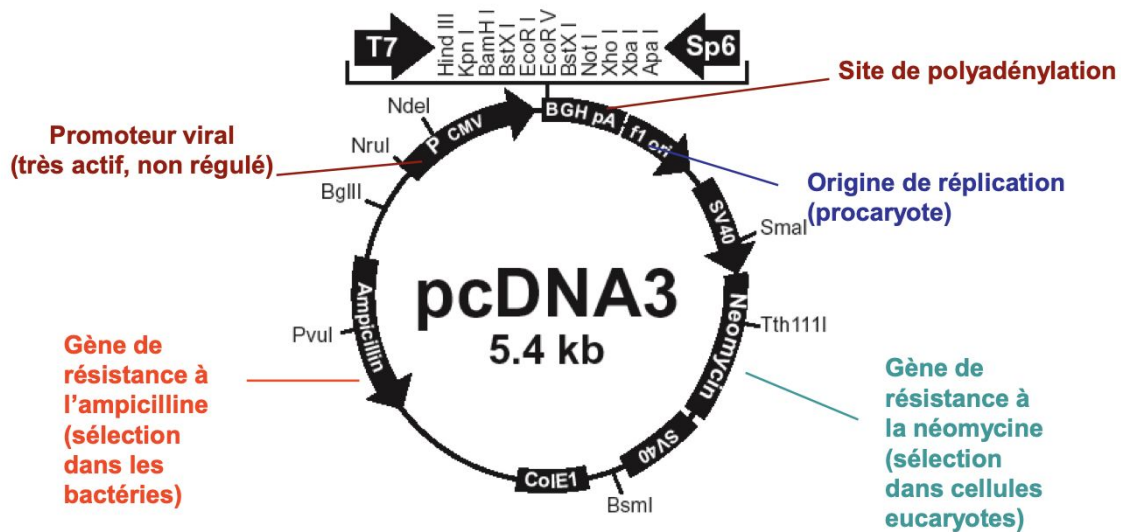
Il faut absolument tout ça pour que la cellule puisse produire la séquence d'ADN.

A partir de ce vecteur on va avoir une séquence codante d'intérêt qui sera placée **devant** un promoteur et on aura un site polyA qui va permettre la **stabilisation**.



**Vecteur d'expression** ≠ Vecteur de clonage

• Exemple de vecteur d'expression eucaryote: Vecteur plasmidique



Le site clonage est constitué de séquences **palindromiques** pouvant être reconnues par des enzymes de restriction. Le promoteur est placé **devant** le site de clonage (CMV promoteur fort, actif chez eucaryote). Enfin, on a un site de polyadénylation qui se trouve **après** le site de clonage et un gène de résistance à l'ampicilline.

Sur ce plasmide on a tous les éléments pour faire du clonage, de la traduction, de la transcription et de la production.

On peut également utiliser un deuxième type de vecteurs, que l'on nomme les **vecteurs viraux**. (Vu en PACES).

Des particules virales sont utilisées comme vecteurs d'ADN ou d'ARN. Ce sont des transporteurs de notre ADN d'intérêt, ils vont l'amener à la cellule hôte.

Cependant, il y a quelques inconvénients : ils coûtent chers, il faut travailler dans des laboratoires protégés (type P2) , etc.

Selon les particules virales que l'on va avoir, on aura soit des virus à ADN ou ARN. On retrouve les :

- **Adénovirus** : Virus non enveloppé à ADN, s'intègre dans la cellule même si elle n'est pas en division, mais ne s'intègre pas au génome.
- **Rétrovirus** : Virus enveloppé à ARN, peut s'intégrer au génome et il a besoin que la cellule soit en division pour s'intégrer.
- **AAV**: Virus associé à adénovirus
- **Lentivirus** : Rétrovirus complexe, pas besoin de réplication

## 2. Choix de l'hôte

On retrouve des eucaryotes ou des procaryotes.

Chez les **procaryotes**, on retrouve la bactérie **Escherichia coli**, des levures, CHO (Chinese Hamster Ovary), des cellules de mammifère.

Concernant les protéines recombinantes thérapeutiques, on utilise majoritairement la bactérie **Escherichia coli** : grâce à cette bactérie, on produit le plus de protéines recombinantes. L'avantage de cette bactérie est que cela coûte **moins cher** qu'une culture de cellules de mammifères par exemple. Tous ces systèmes ont des avantages et des inconvénients.

Escherichia coli, CHO et levures constituent la majorité de la production de protéines recombinantes.

On tend de plus en plus à utiliser les cellules de mammifères voir humaines car on a des protéines compliquées à synthétisées, notamment par leur glycosylation, etc. Certaines vont avoir besoin de modifications post traductionnelles pour être actives. Il faudra donc absolument travailler dans des systèmes de cellules mammifères ou humaines.

### Systèmes d'expression utilisés pour produire les protéines recombinantes thérapeutiques

Systèmes d'expression	% utilisation
<i>Escherichia coli</i>	39%
Autres bactéries	1%
Levures ( <i>Pichia pastoris</i> )	15%
CHO (Chinese Hamster Ovary)	35%
Autres systèmes mammifères	10%

Peu de progrès sur les systèmes d'expression en 20 ans  
E. Coli + Levure + CHO = 89%

De plus en plus de médicaments sont fabriqués à partir de cellules humaines.

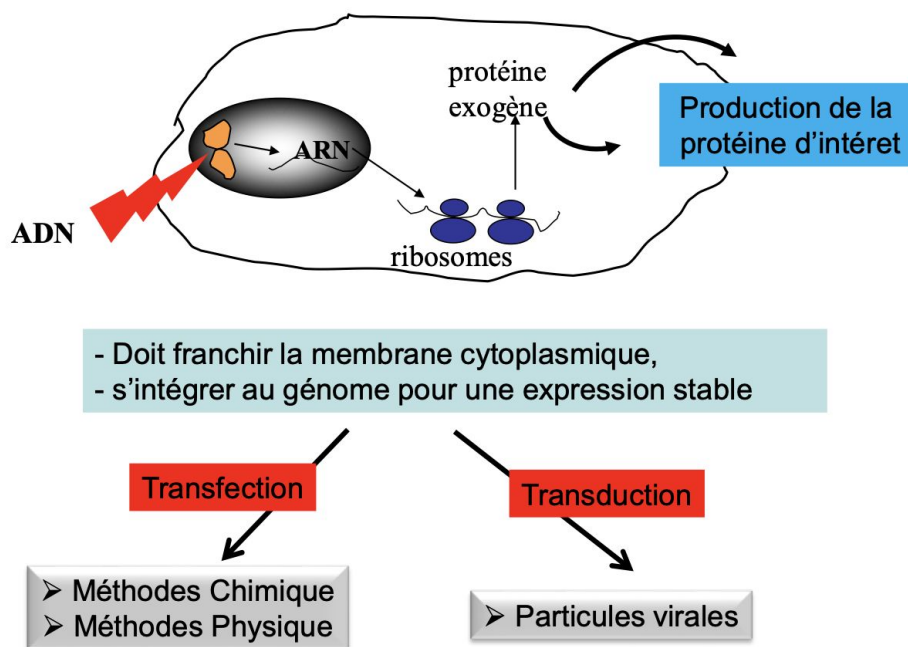
## 3. Insertion du vecteur dans l'hôte

On veut insérer la séquence du gène ou de l'ADN recombinant que l'on veut utiliser dans la cellule hôte, en passant par différentes étapes.

Quand on va utiliser un plasmide, on va parler de transfection . Si on utilise des particules virales, on va parler de transduction .

Pour cela, il va falloir que l'ADN recombinant puisse franchir la membrane plasmique et s'intégrer au génome de la cellule hôte pour une expression stable .





*Comment l'ADN recombinant peut-il pénétrer dans une cellule ?*

### 1ère méthode : la Transfection

- **Méthodes physiques** : par choc électrique ou thermique, techniques très utilisées.

On peut utiliser des **impulsions électriques** dans le cas de l'**électroporation** chez les bactéries, mais aussi chez les cellules en suspension. Cela va créer des **pores**, trous temporaires dans la membrane plasmique de la cellule. Grâce à cette technique, l'ADN recombinant passe à travers les pores de la membrane cellulaire sans altérer cette dernière.

On peut aussi faire un **choc thermique**, méthode qui marche bien chez la bactérie. Pour produire des animaux transgéniques, on peut directement **micro-injecter** l'ADN dans le noyau de la cellule.

Chez les plantes, on utilise un système de **micro-projection**. On applique sur une feuille des particules de tungstène chargées avec de l'ADN, qui peut donc pénétrer les cellules.

- **Méthodes chimiques** : On utilise des liposomes pour faire de la **lipofection**.

Le principe est d'encapsuler l'ADN (chargé négativement) dans des liposomes, que l'on met ensuite en contact avec la cellule hôte : c'est ce qu'on appelle la **transfection**.

On réalise un complexe lipide/ADN, contenant la séquence d'ADN qui nous intéresse, et on le met en contact avec la cellule hôte. Après endocytose de ce complexe,

l'ADN sera relargué dans la cellule: cet ADN restera dans le cytoplasme ou rentrera dans le noyau.

### **2ème méthode : la transduction**

Les vecteurs sont des particules virales : ce sont des transporteurs ayant la capacité de se fixer sur la membrane de la cellule grâce à des récepteurs **spécifiques**. Le virus va être internalisé par **endocytose** et va amener le transgène au niveau du noyau, puis s'intègre ou non dans le génome.

Chacune de ces techniques ont leurs avantages et inconvénients.

#### **• Dans le cas de l'ADN nu ou encapsulé dans liposome :**

→ **Avantage** : facile à faire, pas cher, pas besoin de travailler dans des laboratoires sécurisés car on ne retrouve pas tous les problèmes que l'on peut rencontrer avec les virus.

→ **Inconvénient** : le taux de transfection est faible. Cela veut dire que peu de cellules auront intégré le plasmide, et s'il n'est pas intégré, la cellule ne peut pas produire de protéine. On va devoir faire des lignées stables qui auront intégré le gène d'intérêt dans le génome. Cela demande du temps et de l'argent.

On va donc avoir des systèmes permettant de sélectionner la cellule qui aura correctement intégré le plasmide dans son génome, grâce à l'utilisation d'un gène de résistance à un antibiotique.

**• Les particules virales** : Ce sont des particules qui ne sont pas pathogènes (on a enlevé toutes les séquences pathogènes).

→ **Avantage** : Fort taux de transfection : la majorité des cellules auront intégré le gène d'intérêt.

→ **Inconvénient** : on doit travailler dans des laboratoires sécurisés (par exemple P2).  
> Cela augmente donc le coût de production

### **4- Purification des protéines : culture, extraction, purification et vérification**

#### **- Culture :**

Une fois que l'on a sélectionné la cellule ayant intégré le gène d'intérêt, il faut cultiver, produire, amplifier cette cellule : on va faire cela grâce à des **bio-réacteurs**.

Au début on travaille sur des plaques de cultures cellulaires (on parle de mL de milieu de culture). Ensuite on produit de grands volumes (10-150 litres pour la mise au point et jusqu'à 2 000 - 20 000 L de capacité)

C'est à partir de ces cultures de bio réacteurs qu'on va pouvoir extraire la protéine qui nous intéresse et la purifier. En fonction des cellules (procaryotes, eucaryotes) on aura des bio réacteurs plus ou moins importants / compliqués.

### **- Extraction**

Il va falloir extraire la protéine à partir des cultures présentes dans les systèmes de production.

On peut avoir protéine qui est : **sécrétée ou intra-cytoplasmique.**

Si elle est sécrétée, on peut récupérer le surnageant et purifier la protéine. Cependant, cela ne marche pas très bien : on aura une protéine qui reste dans le cytoplasme, que ce soit au niveau des bactéries mais aussi dans les cellules de mammifères.

Si la protéine est à l'intérieur d'une bactérie, il va falloir lyser sa paroi par **sonication**. Ensuite on centrifuge afin de séparer la protéine d'intérêt et les débris membranaires. On travaille à 4°C pour **inhiber les protéases** qui peuvent cliver la protéine d'intérêt.

Il existe un autre problème lorsque l'on parle de bactérie: il y a des corps d'inclusion à l'intérieur de la bactérie qui rendent insoluble (*mais la prof n'a pas développé ce point en cours*)

### **- Purification :**

Une fois la protéine récupérée, on a encore beaucoup de particules : il va falloir la purifier.

Pourquoi ? Car on veut isoler la protéine intégrée et éliminer toutes les traces de cellules, de virus, de milieu culture, de produits dégradation.

Comment ? On utilise des technique de **chromatographie** :

- Purification par filtration sur gel
- Chromatographie d'affinité
- Chromatographie échangeuse ion

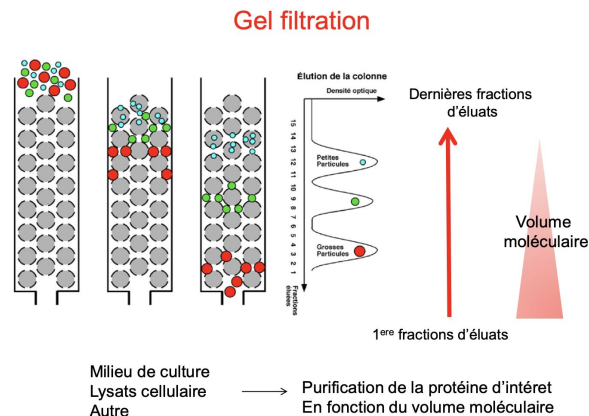
- **Gel filtration** : (*on le fera en TP*)

Cette technique repose sur la purification de la protéine d'intérêt en fonction du **volume moléculaire**.

Ce n'est pas une technique spécifique : on n'arrive pas à vraiment purifier la protéine, mais on enlève pas mal de particules.

### **Principe :**

On ajoute le milieu de culture qui contient la protéine d'intérêt en haut d'une colonne. On laisse la chromatographie se faire. Dans la colonne se trouvent des grilles avec des billes qui vont retenir les petites particules. Les grosses particules, quant à elles, vont être exclues du gel. Donc les premières particules à sortir sont les particules de grosse masse/volume moléculaire. On a ensuite les particules moyennes puis en dernier les petites particules, retenues par les billes de la colonne.



- **Chromatographie échangeuse de cations / anions :**

Pour cette technique, on va utiliser des colonnes chargées :

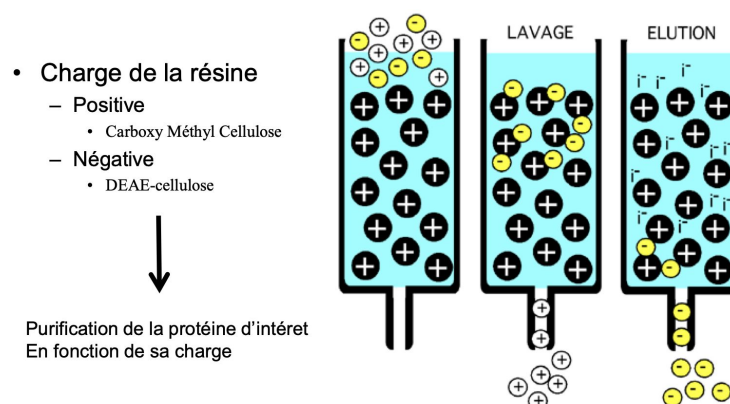
- Positivement : carboxy méthyl cellulose
- Négativement : DEAE-cellulose

Ce n'est pas une technique spécifique car beaucoup d'autres protéines peuvent être chargées négativement par exemple. Mais elle permet tout de même d'enlever beaucoup d'impuretés.

**Principe :**

On ajoute dans la colonne le mélange que l'on veut purifier, et toutes les molécules chargées négativement vont être retenues par la colonne chargée positivement. Toutes les charges positives vont être exclues de la colonne.

Pour récupérer toutes les molécules accrochées à la colonne, on va devoir faire une élution. Pour cela on va utiliser un tampon qui va contenir un ion chargé négativement (donc de la même charge que notre protéine d'intérêt). Les concentrations de cet ion vont augmenter et il va entrer en compétition avec les molécules chargées négativement, et va permettre de détacher les charges négatives de notre colonne chargée positivement. On va donc pouvoir récupérer notre protéine d'intérêt.



- **Chromatographie d'affinité :**

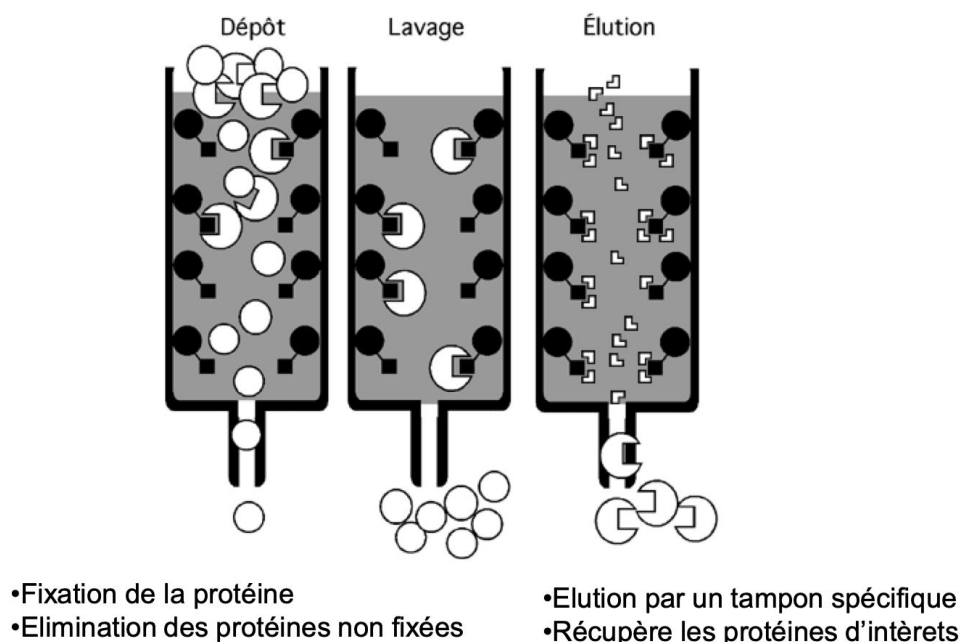
C'est une technique **spécifique** : elle est basée sur la reconnaissance d'un ligand à une molécule, qui est une fixation spécifique.

**Principe :**

On a une colonne qui contient, sur ses parois ou lignes, des petits carrés. Ils permettent la liaison spécifique avec notre molécule d'intérêt.

On dépose notre mélange en haut d'une colonne. Notre molécule d'intérêt va se fixer spécifiquement à la colonne et toutes les autres molécules ne vont pas être retenues.

On va ensuite faire une élution afin de décrocher la molécule d'intérêt fixée sur la colonne. Pour cela, on va utiliser un tampon spécifique contenant une molécule entrant en compétition avec la liaison. Enfin, on récupère spécifiquement notre molécule d'intérêt.



Pour faciliter la purification, on peut fabriquer une **protéine de fusion** : c'est une protéine d'intérêt à laquelle on a ajouté des séquences **spécifiques** que l'on choisit, que l'on nomme **étiquette** ou **tag**. C'est cette étiquette qui sera reconnue dans la colonne à un endroit particulier. C'est une protéine qui est créée, qui n'existe pas dans la vraie vie.

Les protéines de fusion doivent avoir une liaison réversible avec le ligand de la chromatographie

Cette protéine de fusion sera produite par le choix du vecteur que l'on aura fait **avant** de lancer la purification.

### Principe :

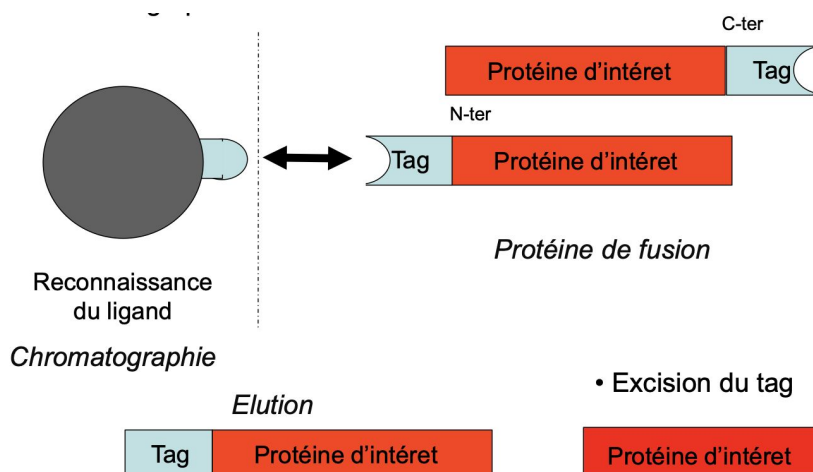
L'étiquette (ou tag) ne peut être mise que dans la partie **C terminale** ou **N terminale**. Si elle s'insère au milieu ou bien même dans la protéine, cela va changer sa configuration, et donc son activité. Or ce n'est pas le but.

Cette protéine de fusion va pouvoir reconnaître **spécifiquement un ligand** qui sera sur une colonne de chromatographie. Le but est qu'il y ait reconnaissance.

Ensuite on va éluer spécifiquement la protéine qui nous intéresse avec un composé qui rentre en **compétition**.

On va utiliser un substrat qui va venir en compétition pour dissocier la liaison, on récupère la protéine et enfin on utilise une enzyme qui enlèvera le tag car il ne nous intéresse pas. Il doit **absolument être enlevé** car cela peut modifier l'activité de la protéine, peut induire des effets secondaires. Le système immunitaire peut reconnaître le tag comme étant étranger et induire une réaction immunitaire.

Il faut faire attention lorsque l'on fait ce type de manipulation à bien utiliser l'enzyme à la fin pour enlever le tag.



Il va donc falloir **un vecteur d'expression** pour exprimer notre protéine de fusion. Ce sont des vecteurs qui ont été adaptés et il existe tout un tas d'étiquettes : histidine, GST... permettant une liaison réversible avec un ligand qui se trouve sur la chromatographie.

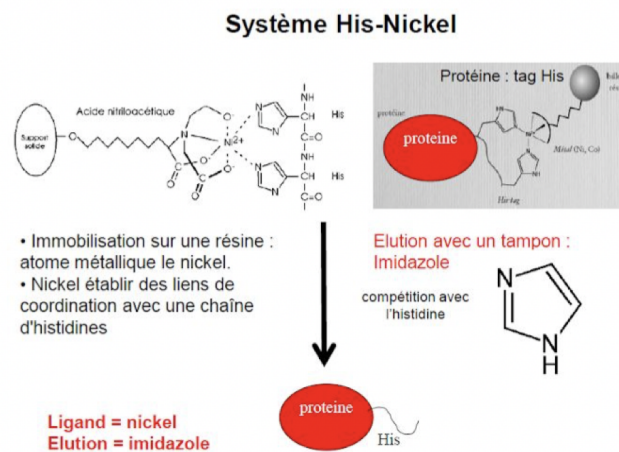
On produit une protéine qui doit être activée ! Il existe différentes chromatographies et différents tag qui vont pouvoir fixer réversiblement différents ligands sur la colonne de chromatographie.

Les deux systèmes les plus utilisés sont :

- Le tag **histidine** (acide aminé), c'est à dire une séquence d'histidine ajoutée à la protéine et pouvant **chélater des métaux**.
- Le tag **glutathion S transférase(GST)** pouvant lier le **glutathion**.

## 1er cas : le système His-Nickel

Premièrement on va faire une protéine de fusion qui aura un tag HIS, en position carboxy-terminale ou N-terminale. La **colonne** contient un support qui va permettre de lier le **nickel**. Ce nickel, quand on va faire passer la protéine d'intérêt sur la colonne, va venir se **complexer avec des HIS**. On aura donc une **immobilisation** de notre **protéine** grâce à cette liaison nickel-HIS, notre protéine sera donc fixée. Maintenant pour pouvoir **détacher** la protéine d'intérêt, on va rajouter un **compétiteur** de la liaison nickel-HIS, il va être dans le tampon d'élution et ce sera de **l'imidazole**. L'imidazole ressemble à l'HIS, il va donc entrer en compétition avec le tag HIS. On va faire des concentrations croissantes de l'imidazole et à un moment à la concentration idéale on aura une compétition entre nickel-HIS et nickel-imidazole. On va donc détacher la protéine, que l'on va retrouver dans notre éluat, elle sera taguée HIS. Ensuite il faudra enlever ce tag.



## 2eme cas : le système GST-Glutathion

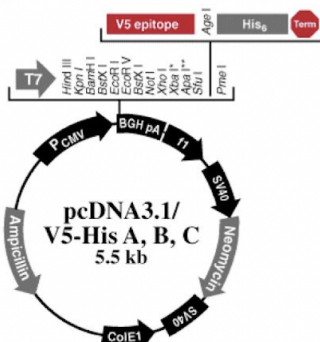
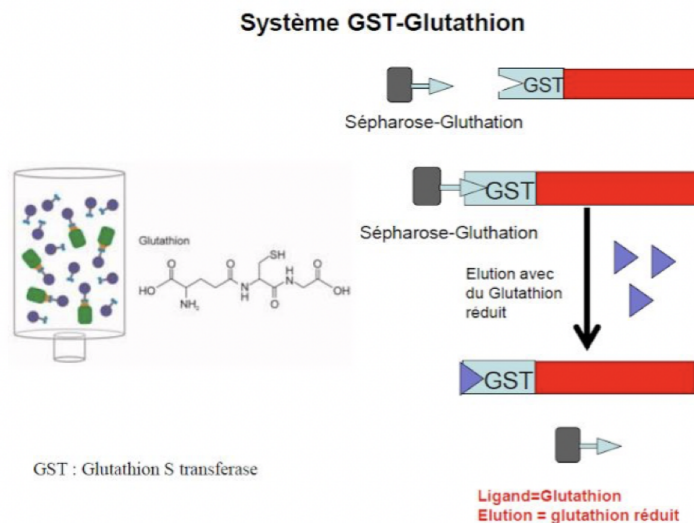
On a une protéine de fusion qui contient du GST. Cette GST va reconnaître le glutathion de la colonne qui est lié à de la sépharose sur la colonne. Donc le **GST** reconnaît la **Sépharose-Glutathion** et crée une liaison.

On va **éluer** en remettant dans la colonne un **excès de glutathion réduit** qui va venir entrer en compétition avec GST.

Notre protéine va venir se fixer sur la sépharose-glutathion, on va ensuite lavé puis en ajoutant le glutathion réduit on va faire sortir notre protéine. Encore une fois la dernière chose à faire c'est de couper avec une enzyme pour enlever ce tag.

On a donc plusieurs systèmes qui existent. Cela veut dire que si on veut produire une protéine de fusion il faut avoir penser avant la production. Il faut y avoir pensé quand on a

fait notre clonage, car c'est le vecteur que l'on va utiliser qui va amener la séquence que l'on veut introduire (HIS ou GST).



Ci-contre on a un exemple de plasmide qui est un plasmide utilisé pour faire 2 types de tag (tag HIS et tag V5). Ça veut dire que ici, si on met le gène qui va coder pour notre protéine au niveau du site de clonage on aura un ajout de 6 HIS qui sera en carboxy-terminale ici. Les cellules vont produire ce qu'il y a sur la séquence c'est à dire le gène qui va coder pour notre protéine plus la séquence de 6 HIS.

La dernière chose qu'il va falloir faire c'est vérifier si on a bien notre protéine.

### - Concentration

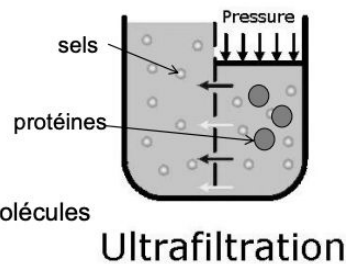
Cela consiste à concentrer le produit final afin d'en faire un médicament. Pour concentrer la protéine, on va utiliser une membrane qui va retenir les grosses molécules et laisser passer les petites molécules.



- **Concentration : dialyse**

- **Membrane de dialyse :**

- Laisse passer les petites molécules (sels, eau)
    - Ne laisse pas passer les grosses molécules (protéine)

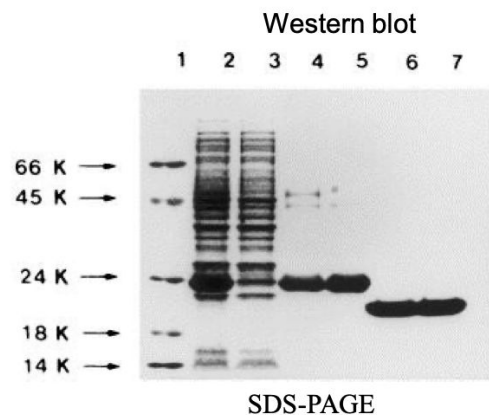


### Vérification de la protéine

Si on veut savoir si la protéine est présente dans la solution on peut faire un **Western Blot**. Cela consiste à prendre un **gel d'acrylamide** qui contient du **SDS** (donc il est dans des conditions dénaturantes). Sur ce gel on va faire **migrer** notre mélange de protéine ou la solution que on veut tester. Ce type de migration est une électrophorèse, donc les **protéines** vont migrer uniquement par rapport à leur **poids moléculaire** (PM), car on est dans des conditions de SDS et que ce dernier va charger négativement toutes les protéines. On a donc des protéines de haut PM en haut du gel et celle de bas PM en bas du gel.

Une fois qu'on a fait migrer les protéines, on va pouvoir mettre en évidence ces dernières, tout simplement en faisant une coloration (colorants viennent se fixer sur les protéines ).

Sur toutes ces bandes on a plusieurs protéines qui correspondent toutes à peu près à la taille où elles migrent. (ex : trait épais à 24K, on a plusieurs protéines)



*Comment savoir si dans un de ces mélanges (traits) on a notre protéine ?*

Après avoir fait migrer les protéines, on fait un **transfert** sur une **membrane**, et on l'**incube** avec un **anticorps** qui va reconnaître spécifiquement la protéine. C'est ce qu'on appelle un Western Blot.

Ainsi, si sur la membrane où on a incubé l'anticorps, on a une seule bande qui correspond au PM attendu, on peut dire que la molécule est bien présente et qu'elle fait le bon PM.

Mais, attention, ce n'est pas parce que on a produit une protéine de bon PM qu'elle sera active. La dernière chose qui faut faire est donc de vérifier l'activité de notre protéine. On le fait en regardant, si elle est capable de se fixer à un récepteur et on étudie cette capacité, soit c'est une enzyme et je met un substrat et je regarde son activité,...

Résumé : Quand on produit une protéine, après la purification il faut vérifier que on a la protéine dans notre mélange, qu'elle est pure, qu'elle est au bon PM, que la structure est bonne et qu'elle à l'activité attendue.

## Contrôle de qualité

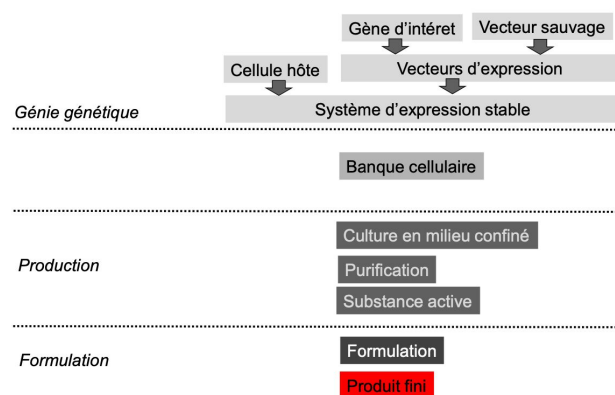
Pour obtenir, ainsi que les commercialiser, des biomédicaments, il va falloir faire tous les contrôles qualités nécessaires pour la production d'un médicament. On va devoir respecter les bonnes pratiques de fabrication (GMP), les protocoles d'opérations normalisés, la validation des procédés et de la documentation. Tout doit être fait pour commercialiser un médicament.

ANALYSE DE LA PROTÉINE	
CARACTÉRISTIQUE	TECHNIQUES ANALYTIQUES
IDENTITÉ	<ul style="list-style-type: none"> <li>Western-blot / ELISA</li> <li>Spectrométrie de masse (MS) : carte peptidique</li> </ul>
PURETÉ	<ul style="list-style-type: none"> <li>ELISA</li> <li>PCR quantitative en temps réel</li> </ul>
VARIANTS DE MASSE	<ul style="list-style-type: none"> <li>Electrophorèse (SDS-PAGE, Capillary)</li> <li>Chromatographie HPLC (SEC +/- Light Scattering)</li> </ul>
VARIANTS DE CHARGE	<ul style="list-style-type: none"> <li>Isoélectrofocalisation (IEF ou (i)C-IEF)</li> <li>Chromatographie HPLC (IEC)</li> </ul>
VARIANTS DE GLYCOSYLATION	<ul style="list-style-type: none"> <li>Chromatographie HPLC (RPC)</li> <li>Spectrométrie de masse (MS)</li> </ul>
ACTIVITÉ BIOLOGIQUE	<ul style="list-style-type: none"> <li>ELISA (binding assay)</li> <li>Test cellulaire (bio-assay)</li> </ul>
QUANTITÉ	<ul style="list-style-type: none"> <li>Spectrophotométrie UV 280nm</li> </ul>
QUALITÉ MICROBIOLOGIQUE	<ul style="list-style-type: none"> <li>Dénombrement de la charge microbienne</li> <li>Dosage d'endotoxines</li> </ul>

Le tableau ci-dessus (a ne pas savoir) présente l'ensemble des caractéristiques à vérifier lorsque l'on a produit notre protéine.

Si on veut comme médicament une protéine recombinante, il aura fallu avant, intégrer la séquence qui va permettre de coder pour cette protéine dans un vecteur, il va falloir cloner, puis intégrer ce vecteur dans un système d'expression (transfection ou infection). A partir de ça on va pouvoir sélectionner des banques cellulaires qui vont avoir le gène qui code pour notre protéine, il va ensuite falloir le cultiver, le purifier et obtenir la substance active, cela correspond aux étapes de production de la protéine. Et enfin, on va faire de la formulation.

## Procédés de fabrication



△ **Attention** △ : On ne peut donner une protéine recombinante par **voie orale** car ce système contient des protéases et donc, les protéines recombinantes seraient détruites avant même d'arriver à la cellule cible.

## **B- Production de protéines recombinantes « biomédicaments »**

### **1. Le système procaryote**

L'exemple du cours, la bactérie **E.coli**. Il s'agit du système le plus utilisé pour produire des protéines recombinantes. Avec ce système on peut former de **l'insuline, des hormones de croissance, des interférons, de l'interleukine-2 et d'autres médicaments**.

#### **Les avantages d'E.coli :**

- C'est le système le **plus simple** pour produire des protéines recombinantes :

- La culture est **rapide** : On obtient beaucoup de molécules (des centaines et des milliers). En effet, les bactéries ont un temps d'incubation très court, 30 min. De ce fait on peut en produire beaucoup en très peu de temps.
- **Pas cher** car on utilise des milieux de culture simple.
- **Non pathogène pour l'Homme**.

- Le système génétique est très bien connu. De ce fait, on arrive à bien "maîtriser". De plus, on a la possibilité d'utiliser des souches très spécifiques grâce aux biotechnologies.

- Il y a énormément de vecteurs d'expressions qui sont disponibles. On constate donc **une facilité de transfection**.

- L'expression des protéines est **cytoplasmique** et elles peuvent être retenues dans des structures particulières, les **corps d'inclusion**, les protéines seront insolubles dans ces structures. Les protéines peuvent aussi **être sécrétées** dans l'espace **périplasmique**, mais cependant, jamais en extra-cellulaire. Il faudra donc casser la membrane pour récupérer la protéine.

- Les protéines recombinantes **sont faciles à purifier** : il suffit de casser la membrane de la bactérie.

- La culture en milieu liquide dans des fermenteurs permet de travailler avec des grandes quantités.

- **E.coli** possède en plus un bon rendement, c'est à dire la quantité de protéine produite par la bactérie. La protéine recombinante peut représenter 25% des protéines totales. Ceci est un bon rendement.

#### Les inconvénients d'E.coli :

- La bactérie **est incapable de faire des modifications post-traductionnelle**. De ce fait, si une protéine est glycosylée ou phosphorylée, E.coli ne pourra pas être utilisée pour la produire.

- Difficulté pour produire des grosses protéines. On est limité à de petites capacités. En effet, cela reste possible **entre 5 et 30 kDa, mais devient impossible au delà de 50 kDa**.

- Il peut exister une **toxicité** de la protéine recombinante pour la bactérie.

- Pour certaines bactéries, les protéines recombinantes vont être en excès dans les corps d'inclusion. Cela pourrait indiquer un mauvais repliement de la protéine finale (structure tertiaire n'est pas bonne) et une insolubilité de celle-ci.

- On peut trouver la présence **d'endotoxines bactériennes**.

#### Exemple de l'insuline :

C'est le 1er biomédicament synthétisé, il est utilisé pour traiter les diabétiques insulino-dépendants. C'est la **1ère protéine recombinante** commercialisée en France. Elle provient du **génie génétique**.

L'insuline, avant l'existence du génie génétique, était extraite à partir de tissus de porc. Cependant, cela induisait des allergies ainsi que la fabrication d'anticorps anti-insuline dans le corps des patients injectés.

Au niveau de sa structure, on observe 2 chaînes polypeptidiques(A et B) qui sont reliées par 2 ponts disulfures et 1 pont intrachaîne.

De nos jours, en France, l'insuline provient du génie génétique mais, les protéines recombinantes **coûtent très chers**, les pays pauvres et ceux en voie de développement n'ont pas la capacité de produire, du fait de leurs moindre pouvoir d'achat.

L'insuline comporte des ponts disulfures alors que nous venons de voir que le système procaryote (E.Coli) **était incapable de faire des modifications post-traductionnelles.**

Comment faire?

Il y a 2 façon de produire de l'insuline :

- On utilise le gène entier de l'insuline et elle sera synthétisée sous forme de pré-proinsuline, qui sera maturée en insuline.

*Rappel : dans la cellule humaine on a l'insuline sur un stade de pré-proinsuline qui est ensuite clivée en proinsuline et puis grâce à des clivages on a formation de l'insuline qui sera mature.*

**OU**

- On produit **la chaîne A et la chaîne B dans des bactéries séparées** puis, par un traitement chimique on va reconstituer l'insuline active, car **on fait ces ponts disulfures ex vivo.**

En effet, on peut utiliser **E.coli**. On exprime le vecteur bactérien contenant le gène de la chaîne A qui est transfecté dans une bactérie à laquelle on va faire subir la sélection expliquée précédemment.

Ensuite, on cultive ces bactéries. On a donc la **chaîne A** dans cette bactérie. **On suit le même processus**, en parallèle, pour la **chaîne B**. Par un traitement, on va pouvoir **refaire les ponts disulfures** entre les chaînes **A et B**. En bidouillant les bactéries on est capable de leur faire faire des choses que normalement elle ne serait pas capable de faire.

Le **génie génétique** nous a permis de modifier les séquences qui codent pour l'insuline pour **modifier la pharmacocinétique** de cette protéine. Cela veut dire que en modifiant des séquences, on va pouvoir parler d'insuline à action rapide et à action lente. Cela est due à la possibilité de changer les acides aminés dans cette molécule d'insuline. On a donc une molécule légèrement différente de celle qui est endogène, on va l'améliorer en changeant la pharmacocinétique.

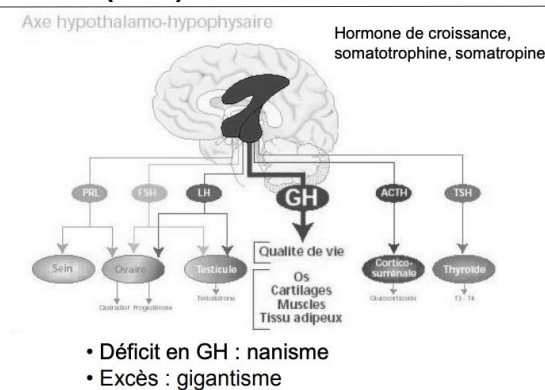
**Exemple de l'hormone de croissance :**

Avant qu'elle soit fabriquée par **génie génétique**, on extrayait cette hormone à partir **d'hypophyse d'animaux, mais pas de très grande efficacité...**

Plus tard on a extrait l'hormone de croissance à partir de l'hypophyse de cadavres. Cependant, certains enfants ont développé la maladie de **Creutzfeldt-Jakob** à cause de prions présents dans la préparation de ces hormones extraites de cadavres contaminés.

C'est une hormone **synthétisée par l'hypophyse** et qui est très importante pour la **croissance**. Elle est déficiente chez des patients atteints de nanisme et elle est en excès dans le gigantisme.

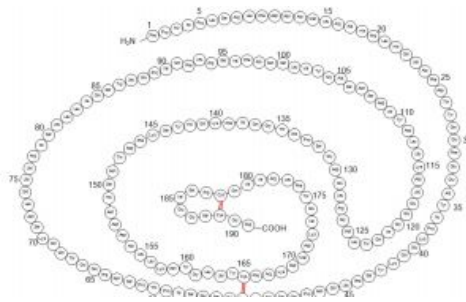
## Hormone de croissance : Growth hormone (GH)



En France, plus de **120 décès d'enfants** ont été recensés après injection de cette hormone de croissance.

Depuis 1988, cette hormone de croissance est synthétisée en France grâce au génie génétique. Le problème a donc disparu. Ceci est possible grâce à E.coli. On retrouve des ponts disulfures dans l'hormone grâce à la possibilité d'avoir des souches bactériennes modifiées capables de faire des ponts disulfures grâce à des manipulations du génome de cette bactérie.

Cette hormone n'est **pas glycosylée**. Elle possède une structure **assez simple** et possède des **ponts S-S**.



*Cela est vrai France, pas dans tous les pays. Comme dit précédemment, il y a encore pleins de pays pas assez **avancés ou riche pour utiliser la biotechnologie**.*

Nous pouvons la produire aussi à partir de **E.coli**. On va faire le clonage, on va apporter le vecteur avec la séquence qui va bien pour cette hormone de croissance. Puis on va tout mettre dans la bactérie, on la cultive et on obtient notre protéine purifiée.

Elle possède aussi 2 ponts S-S. On peut donc aussi faire ces ponts à l'aide de système bactérien car on a des souches que l'on a modifié et qui sont donc capable de faire ces ponts. Alors que normalement elles en sont incapables.

## 2 - Système Eucaryote

Eucaryote inférieurs: Les levures et certains champignons. Séquences très bien connues par l'Homme, le rendement est bon, le poids moléculaire accepté est entre 50-100 kDa, bonne culture en bioréacteur,... La modification post-traductionnelle est possible (glycosylation simple, acylation).

### Avantages levures

- Matériel génétique simple et séquencé
- Non toxique pour l'homme
- Bon vecteur d'expression
- Protéine : 50-100 kDA
- Bon rendement de production
- Bonne culture en bioréacteur (simple, innocuité)
- Capable de modification post-traductionnelles simples :
  - Glycosylation simple
  - Acylation

### Inconvénients des levures

- Nécessité de casser la paroi pour récupérer la protéine
- possibilité de sécrétion mais rendement + faible
- Problèmes de modifications post-traductionnelles de type eucaryote sup.

Les inconvénients des levures sont :

- La protéine n'est, en général, pas sécrétée, et le rendement est plus faible. Il va donc falloir isoler notre protéine en cassant la membrane. Et puis même si on peut faire de la glycosylation, certaines protéines demande des modifications complexes et on ne pourra donc pas utiliser cette cellule.

Eucaryotes supérieurs: C'est donc les cellules mammifères.

Beaucoup de cultures eucaryotes se font avec des cellules de mammifère : des cellules adhérentes (CHO) qui sont des lignées de cellules qui sont immortelles. On peut donc en cultiver, ça va très vite.

Attention, ce ne sont pas des bactéries. Le temps de culture, n'est pas le même que pour les bactérie (=30 min). Il se compte en heures, environ 12H, pour les cellules mammifères.

Certaines entreprises développent ces systèmes en utilisant des cellules humaine pour avoir la même composition que dans notre corps, mais cela coûte plus chère.

Nous connaissons tous l'infarctus, cette pathologie qui a pour conséquence une obstruction des voies sanguines car il y a une plaque athérome, cela forme un thrombus. On cherchera donc à faire une fibrinolyse, c'est à dire détruire le caillot/thrombus.

Le tPA (activateurs tissulaire du plasminogène) est produit par le système eucaryote. On va le donner en urgence au personne qui font un infarctus. Le tPA permet la lyse du thrombus. Ce médicament est le premier à être mit sur le marché provenant du système eucaryote.



Actuellement, les système eucaryotes permettent aussi la production d'EPO, HC, de facteurs de coagulation et d'anticorps monoclonaux.

### Avantages cellules eucaryotes supérieurs : ex CHO

- Bonne culture en bioréacteur
- Synthèse de protéines complexes et de haut PM >100 kDA
- protéines maturées par les modifications post-traductionnelles proches de l'homme
- Possibilité d'obtenir des clones stables
- Pdt protéines solubles

### Inconvénients cellules eucaryotes : ex CHO

- Croissance lente/bactérie (24h)
- Rendements faibles
- Système fragile et culture coûteuse
- Instabilité génétique (perte du vecteur)
- Sélection de la lignée prend beaucoup de temps

Un des **gros avantage, super important** de ce système est le fait que les molécules soient directement **sécrétées dans le milieu de culture, dans le surnageant**. Donc pas besoin d'éclater la membrane, à contrario de E.coli ou des levures.

Il suffit juste de récupérer le surnageant et de purifier la protéine. De plus, comme la membrane reste intact, **la cellule ne meurt pas, du moins, moins vite**, et donc on peut continuer de l'utiliser pour qu'elle produise notre protéine.

L'inconvénient est que la croissance est lente par rapport à des bactéries et le rendement est faible. Le système est fragile et beaucoup plus coûteux que pour les bactéries. On peut avoir une instabilité génétique (perte de vecteur) et la sélection de lignée cellulaire qui va exprimer notre protéine de façon stable demande du temps.

Les **anticorps monoclonaux** sont produits par **génie génétique** notamment par ce système eucaryote.

Pour rappel, un anticorps est constitué de **4 chaînes polypeptidiques**, qui sont associées pour former une structure quaternaire très compliquée. Pour comparaison, la structure de l'anticorps est beaucoup plus complexe que celle de l'aspirine.

Avec de tels méthodes, aujourd'hui il nous est possible de fabriquer des molécules beaucoup plus complexes, tel que **l'anticorps Anti-VEGF**. Il y a des anticorps bloquant les facteurs de croissance tel que VEGF, qui à la base va induire l'angiogenèse.

On ne peut que le produire par **génie génétique**, par système eucaryote. Donc le but est que l'anticorps anti-VEGF se fixe sur des récepteurs qui induisent l'angiogenèse, afin qu'il

**inhibe** la **croissance** de **nouveaux vaisseaux** sanguins dans une **tumeur**. On vient piéger ce VEGF qui ne va plus pouvoir activer les récepteurs sur les cellules et donc on aura pas d'angiogenèse. Ce médicament a été utilisé pour inhiber la vascularisation/croissance d'une tumeur **ou bien on l'utilise aussi contre la DMLA** (Dégénérescence Maculaire Liée à l'Âge).

La **DMLA** a plusieurs formes, dont la forme humide qui est due à cet excès de VEGF et cette croissance de vaisseaux au niveau de l'oeil, ça tire sur la rétine et provoque des cécités. Chez ces patients on peut faire une injection directement dans l'oeil pour inhiber le VEGF.

### Anticorps anti-VEGF : Le Bevacizumab (Avastin™)

- Production
  - CHO
- Anticorps anti-VEGF
  - VEGF : facteur croissance essentiel pour la formation et développement des vaisseaux sanguins
  - Fixation sur récepteurs : Flt1/KDR
- Mode d'action :
  - Bloque la voie de signalisation du VEGF
  - Inhibe la formation des vaisseaux
- Utilisation dans certains cancers (colorectal métastatique)
  - Bloque l'angiogénèse tumorale
  - Tumeur ne se développe plus

