

## Chapitre III. Bioprocédés et production de protéines thérapeutiques 1

### Généralités sur l'ingénierie moléculaire

L'ingénierie protéique a pour but de modifier de façon plus ou moins importante la structure des protéines afin d'optimiser leurs propriétés fonctionnelles.

Deux techniques étaient principalement utilisées en vue de l'optimisation d'une protéine d'intérêt :

#### 1. La mutagenèse dirigée (*rational design en anglais*)

La conception rationnelle de protéines ou « Protein Design » est une approche qui nécessite une parfaite connaissance de la structure 3D de la protéine, et surtout du rôle précis de chaque acide aminé. Les mutants sont choisis après modélisation, réalisés, exprimés puis testés.

Dans l'ingénierie rationnelle les mutants sont créés par mutagenèse dirigée

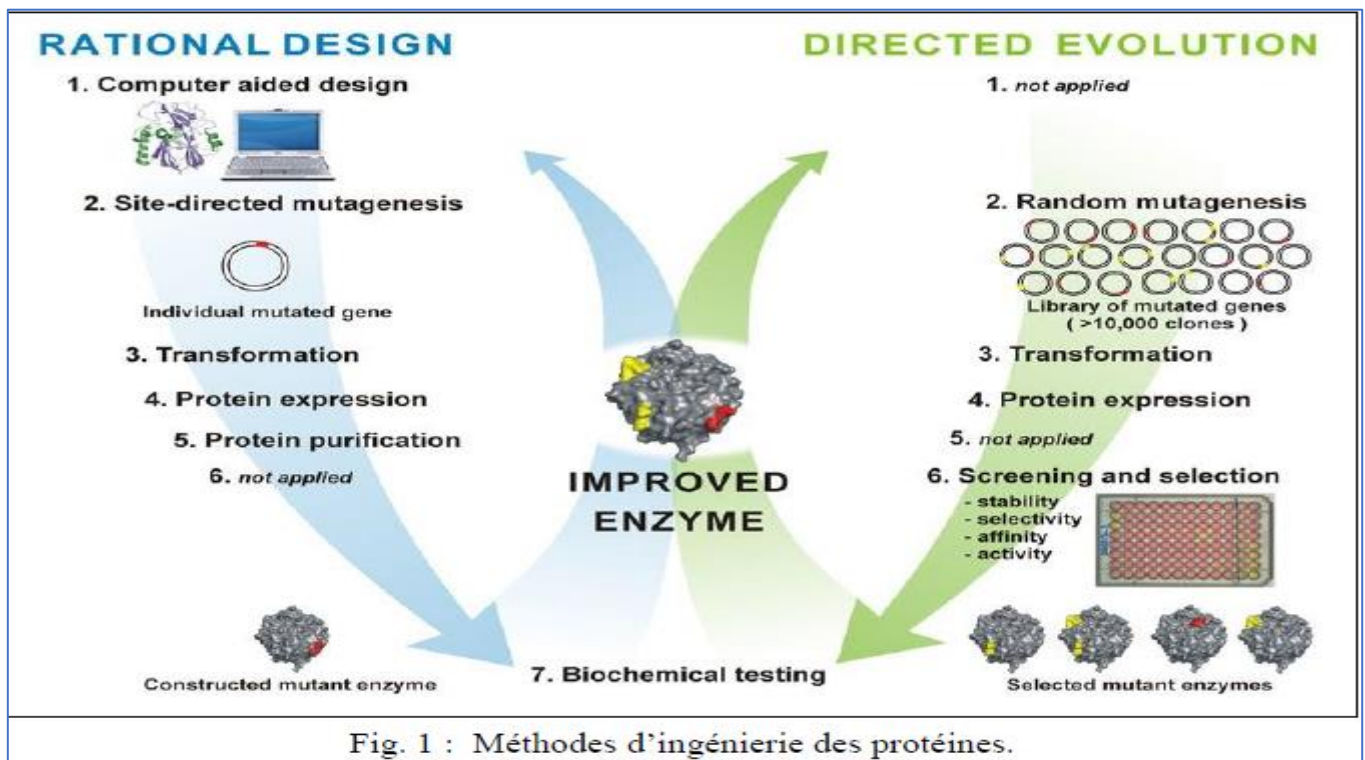
#### 2. Evolution dirigée

L'évolution dirigée est un ensemble de technologies permettant d'améliorer une protéine ou un acide nucléique, en reproduisant artificiellement le processus naturel de l'évolution mais en cherchant à l'orienter dans une direction choisie. Les approches d'évolution dirigée nécessitent deux étapes essentielles.

Les paramètres à améliorer sont variés : Par exemple, on peut obtenir une enzyme qui a une thermo-stabilité accrue ; optimiser une protéine thérapeutique; obtenir un peptide qui se lie avec une forte affinité à un ligand donné; créer *in vitro* un anticorps contre à peu près n'importe quel ligand, pour une utilisation en diagnostic; obtenir une enzyme à la fois résistante à la chaleur et à ou élargir la gamme de pH dans laquelle l'enzyme est efficace tout en augmentant son activité.

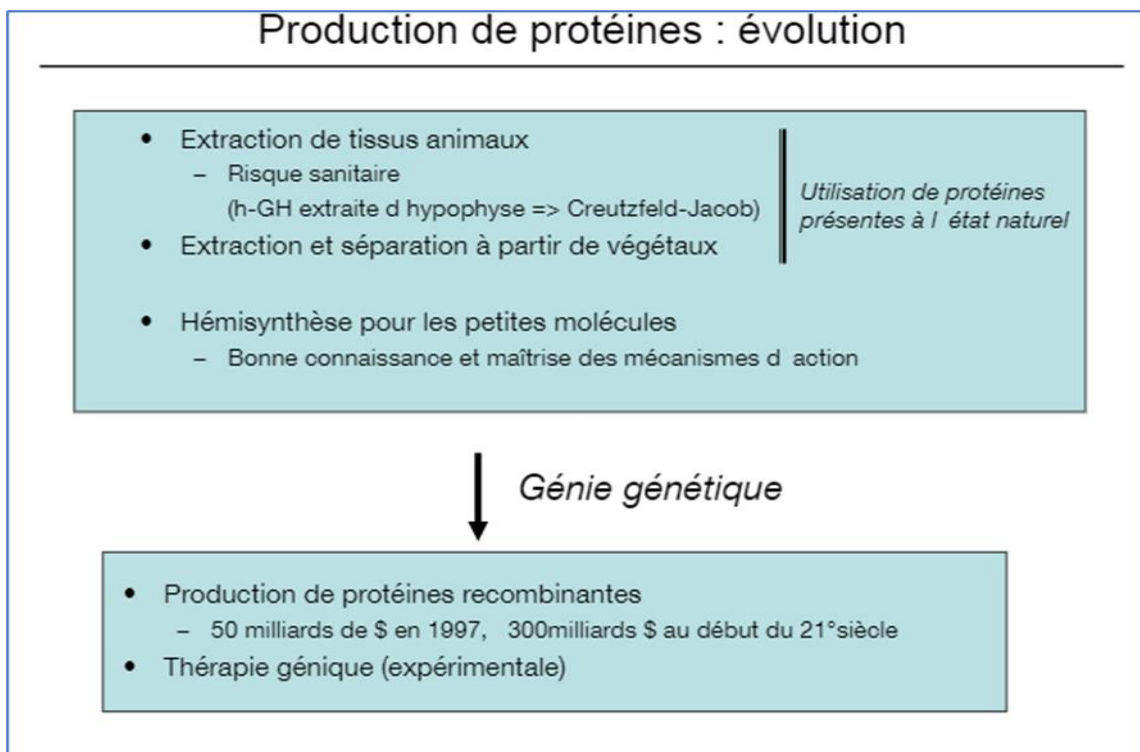
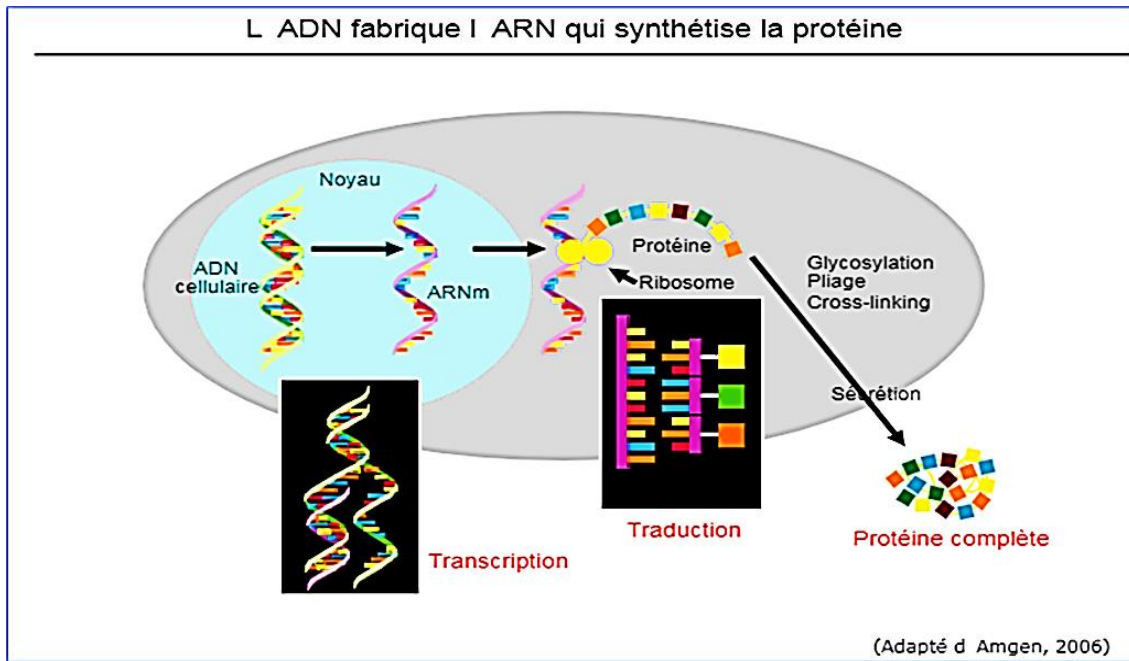
Enfin les enzymes permettent aujourd'hui de remplacer certains procédés chimiques lourds et polluants par des procédés beaucoup plus respectueux de l'environnement (on parle de chimie verte).

Parmi les protéines thérapeutiques qu'il serait intéressant d'améliorer, on peut notamment citer les hormones, les cytokines, les interférons, les vaccins, et les anticorps.

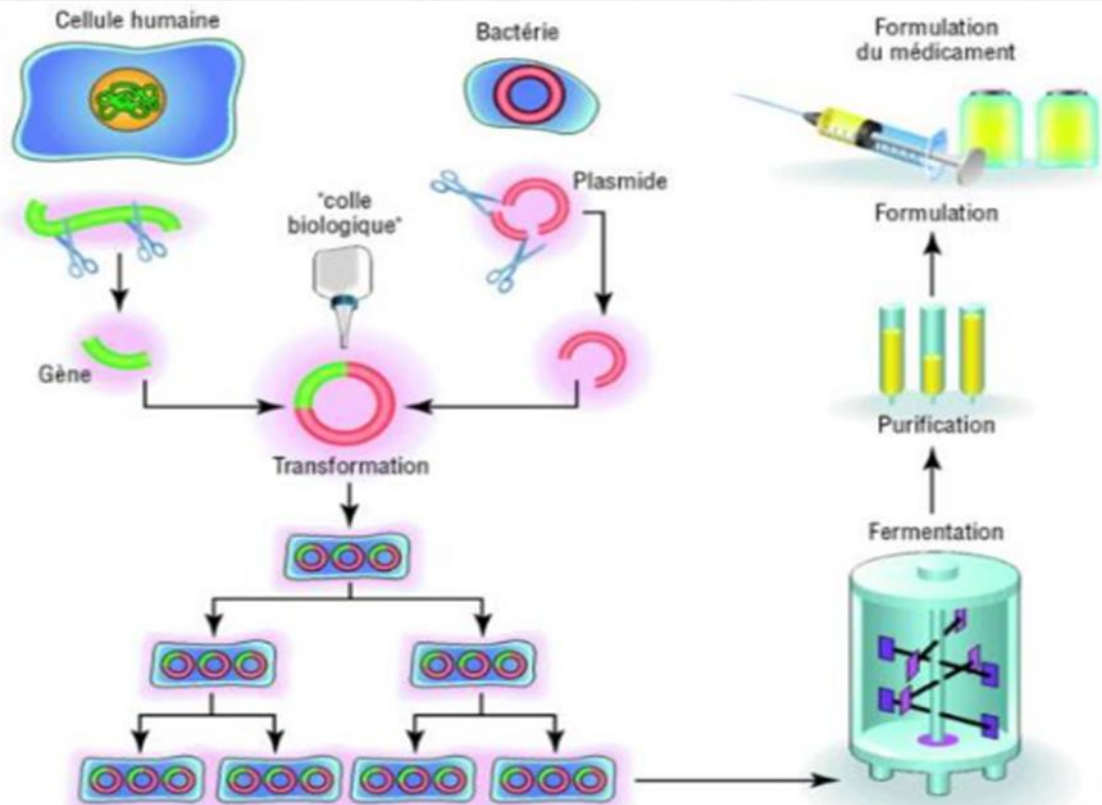


À l'origine, les protéines d'intérêt thérapeutique étaient extraites de sources naturelles telles que le sang, le placenta, ou d'autres extraits de tissus humains ou animaux. Les limitations que cela impliquait en termes d'approvisionnement ainsi que le risque de contamination par des virus - Sida, hépatite B notamment, ont motivé le développement des produits dérivés d'organismes génétiquement modifiés.

En effet, la technologie de l'ADN recombinant (**ADN modifié par recombinaison génétique**) a permis le développement de nouveaux procédés de production de protéines, mettant en œuvre divers systèmes d'expression. L'identification, puis le clonage des gènes gouvernant la synthèse des principales protéines humaines d'intérêt thérapeutique ont rendu possible la production de celles-ci par des organismes hétérologues, et ces protéines, dites recombinantes, ont constitué une alternative aux protéines d'extraction. Aujourd'hui, la quasi-totalité des protéines thérapeutiques produites sont des protéines recombinantes avec un profil thérapeutique mieux défini ce qui a conduit à une plus grande efficacité et à des traitements plus sûrs pour le patient.



## La production de protéines thérapeutiques via la technique de l'ADN recombinant

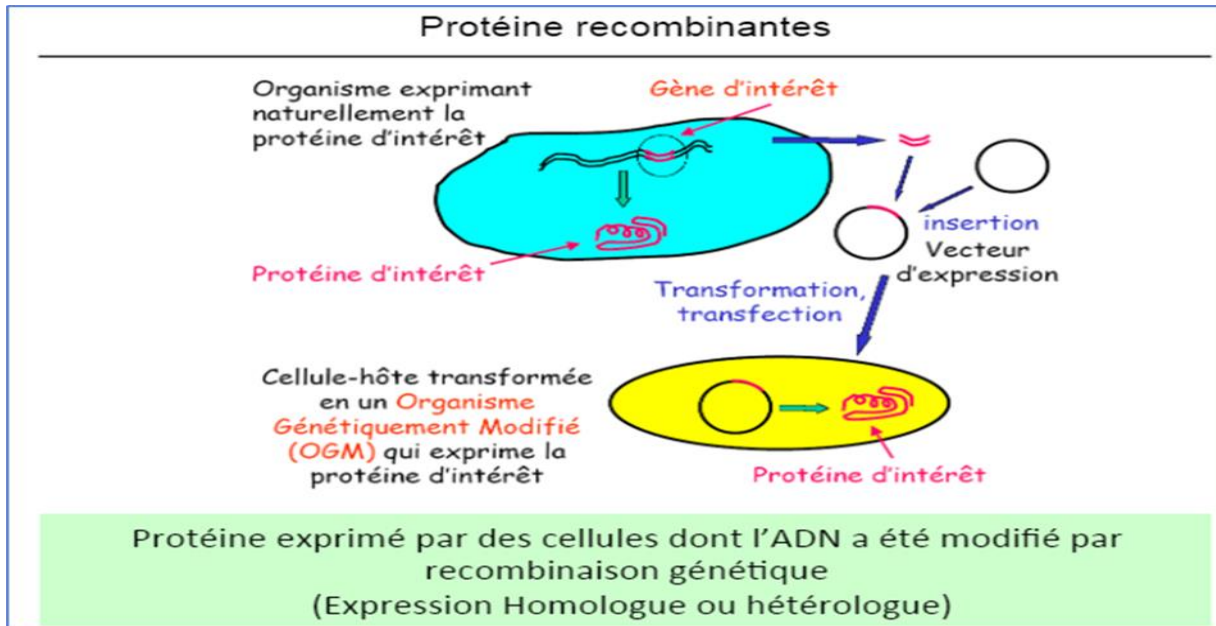


Idée / Source : Fédération européenne des associations de l'industrie pharmaceutique (EFPIA)

*La production de protéines thérapeutiques via la technique de l'ADN recombinant  
(Fédération européenne des associations de l'industrie pharmaceutique - EFPIA)*

## Production des protéines recombinantes

Un système efficace de production de protéines recombinantes s'appuie sur **un gène ou un ADN complémentaire (ADNc)** codant la protéine d'intérêt, **un vecteur d'expression transportant le gène d'intérêt** et **une cellule hôte** qui « exécute » les instructions fournies par la séquence codante pour synthétiser la protéine.



**1/Vecteur d'expression:** En biologie, un vecteur est un **plasmide ou virus** utilisé en **génie génétique** pour **insérer des gènes** dans une cellule. **Molécule d'acide nucléique** dans laquelle il est possible d'**insérer des fragments d'acide nucléique étranger**, pour ensuite les introduire et les maintenir **dans une cellule hôte**.

Le vecteur doit être stable et en grand nombre de copies dans la cellule hôte. Pour exprimer une protéine recombinante, un vecteur doit avoir **un promoteur**, une **origine de réplication**, une **séquence codante**, un **marqueur de sélection**, un **site de liaison aux ribosomes** et un **signal de recombinaison**. Certains vecteurs peuvent contenir des gènes pour des protéines de fusion (TAG) qui facilitent la purification de la protéine recombinante par chromatographie d'affinité. Par exemple, le **bactériophage lambda** et les **plasmides** sont des vecteurs utilisés pour cloner des fragments d'ADN **dans *Escherichia coli***.

### **Aperçu d'un vecteur d'expression typique destiné aux cellules animales**

Un vecteur d'expression contient des éléments essentiels à l'expression des gènes dans la cellule animale. Les **vecteurs les plus communs sont des plasmides**. La figure ci-contre illustre les séquences d'ADN que l'on retrouve le plus souvent dans un vecteur d'expression destiné aux cellules animales. Ce plasmide contient des séquences d'ADN importantes permettant sa multiplication (pUC ori) et sa sélection (Ampr) dans un hôte bactérien. Le promoteur SV40 rend possible l'expression de gènes au sein d'une cellule animale. La séquence de liaison au ribosome (RBS) assure le succès de la traduction tout comme la séquence de polyadénylation (polyA SV40), qui augmente la durée de vie de l'ARN messager transcrit. Le gène d'intérêt est inséré à l'intérieur du site de clonage multiple (SCM). La séquence codant une résistance à la néomycine (Neo) permet de sélectionner les cellules qui contiennent le vecteur et d'obtenir des cellules transfectées de façon stable.

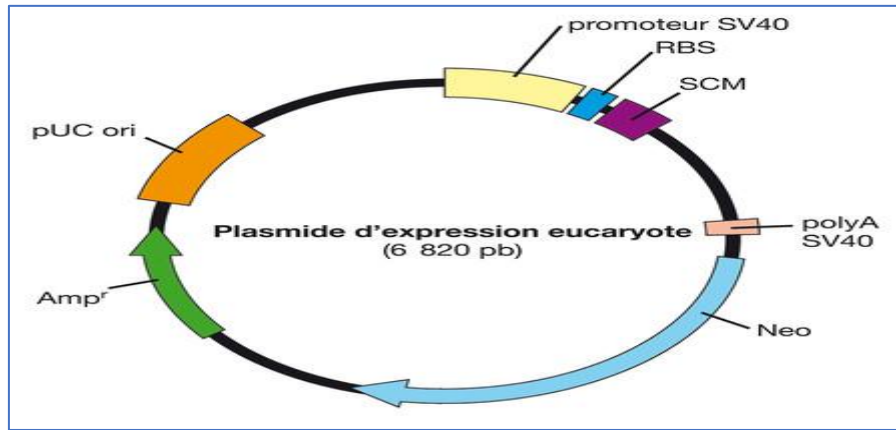


Figure . du manuel *Culture cellulaire animale et végétale*, d'Antoine Campeau-Péloquin et Sophie Roy.

**2/ Cellule hôte:** Pour obtenir de grande quantité d'ADN cloné l'hôte idéal doit:

- Se développer rapidement dans un milieu de culture peu onéreux.
- Être non pathogène.
- Être capable d'incorporer l'ADN.
- Être stable en culture
- Possède des enzymes appropriées pour la réplication du vecteur.

Les hôtes répondant à ces critères sont des microorganismes eucaryotes ou procaryotes dont les génomes sont bien connus car entièrement séquencés, génétiquement manipulables.

**3/ Expression de la protéine:** La protéine doit être soluble (sécritée dans le milieu de culture).

**4/ Extraction et purification de la protéine recombinante:** La protéine recombinante doit être soluble et facile à extraire et purifier.

### Modifications post- traductionnelles

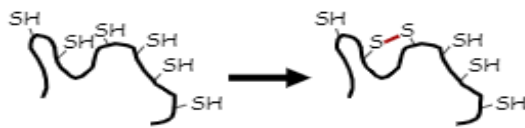
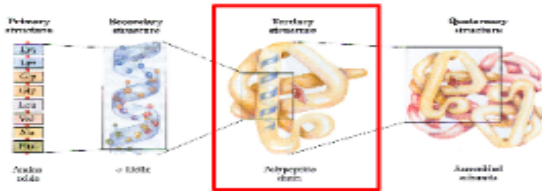
Ces protéines peuvent subir de multiples modifications posttraductionnelles, assurant ainsi leur stabilité et leur activité *in vivo*. Ces modifications incluent notamment la protéolyse, la désamination, la sulfatation et la glycosylation. Ces modifications se produisent après la traduction de l'ARN messenger en une chaîne d'acides aminés et sont spécifiques du type cellulaire considéré. Parmi les modifications requises pour assurer les fonctions biologiques de la protéine, la glycosylation représente la modification la plus importante.



# Modifications post-traductionnelles

## Ponts disulfure

Structure tertiaire de la protéine



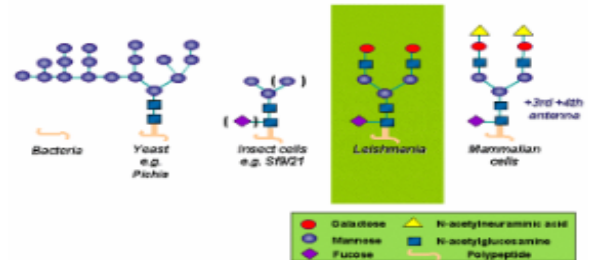
Pas dans le cytoplasme de *E. coli* !

## N-Glycosylation

Stabilité de la protéine

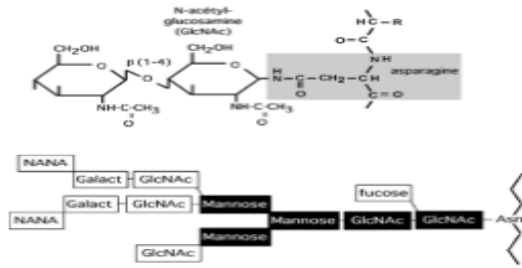
Fonction

Immunogénicité



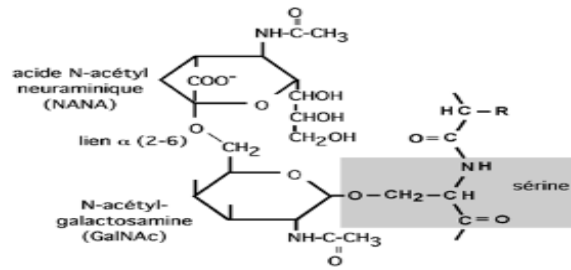
Consensus : Asn-X-Ser/Thr

## Les glycoprotéines



### N-glycosylation

N-X-S(T) (X = tout aa sauf P)



### O-glycosylation

Les chaînes en N peuvent former de véritables arborisations

Les chaînes liées en O sont courtes et variables (1, 2 ou 3 résidus glucidiques)

## Autres modifications PT des protéines

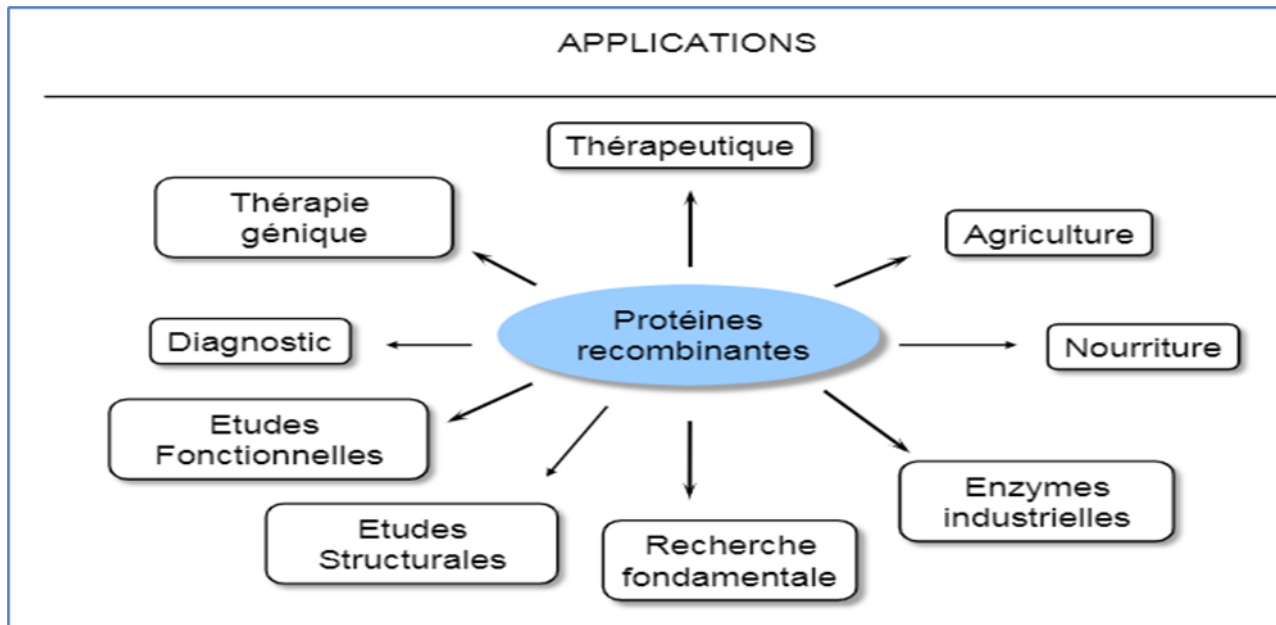
modification	target site	cellular process
Phosphorylation	Ser, Thr, Tyr	signalling, activation
Methylation	Arg, Lys, His, Glu	prot. repair, chemotaxis
Hydroxylation	Pro, Lys, Asn, Asp	collagen structure
Sulfation	Tyr	protein-protein intera'n
Prenylation	Cys	signalling, oncogenesis
Palmitoylation	Cys	membrane association
Myristoylation	N-terminus	membrane association
Acetylation	Lys, N-terminus	gene expression
Sulfation	Tyr	protein-protein intera'n
Amidation	C-terminus	bioactive peptides
Ubiquitination	Lys	degradation/other
Truncation	various	activation

Les modifications post traductionnelles (MPT) non souhaitées pour les protéines à visée thérapeutique incluent l'agrégation des protéines, l'oxydation des méthionines, la désamination de l'asparagine, l'ajout de lipides...etc. De telles modifications ne représentent pas uniquement un challenge au niveau des procédés de production mais peuvent avant tout avoir des conséquences indésirables pour le patient car elles peuvent réduire l'efficacité de la protéine et également entraîner une réaction du système immunitaire.

## APPLICATIONS

Parmi la grande diversité de pathologies visées, les traitements contre les cancers par des anticorps monoclonaux, les maladies inflammatoires, le nanisme par l'hormone de croissance et les diabètes par l'insuline.

## AUTRES APPLICATIONS



## ✓ Etapes et stratégies de la production des protéines recombinantes

### Stratégie générale

#### 1. Clonage moléculaire (transgénèse)

Le clonage moléculaire consiste à isoler un fragment d'ADN complémentaire (codant une protéine d'intérêt) et à l'insérer dans un vecteur de clonage. Celui-ci, de nature plasmidique ou virale, est ensuite introduit dans des cellules hôtes adaptées. Grâce à leur croissance dans un milieu sélectif, ces cellules permettent l'amplification du transgène.

#### 1.2. Isolement d'un fragment d'ADN d'intérêt

- Pour obtenir des fragments d'ADN complémentaire (ADNc, 3'→5') codant la protéine d'intérêt (Ptl), il faut extraire les ARN totaux des cellules exprimant la Ptl et les rétrotranscrire grâce à une ADN synthétase ARN dépendante (transcriptase inverse).
- Les ARN messagers (ARNm) contenus dans le pool d'ARN totaux sont reconnus par la transcriptase inverse qui, en présence de nucléotides, génère de l'ADNc monocaténaire (hybridé avec l'ARNm) qui sera ensuite amplifié par PCR (Polymerisation Chain Reaction).
- Au début de la PCR, la température élevée permet de dénaturer l'hybride ARNm-ADNc et d'activer la Taq polymérase (ADN synthétase ADN dépendante, enzyme thermostable).
- Lors du premier cycle d'amplification, l'enzyme synthétise le brin d'ADN sens (5'→3') en utilisant uniquement l'ADNc comme matrice. Dès le deuxième cycle de PCR, l'ADN bicaténaire précédemment obtenu est dénaturé et chacun des brins sert de matrice pour la synthèse de l'autre.
- Au total, l'amplification aboutit à  $2^n$  copies d'ADN bicaténaire.

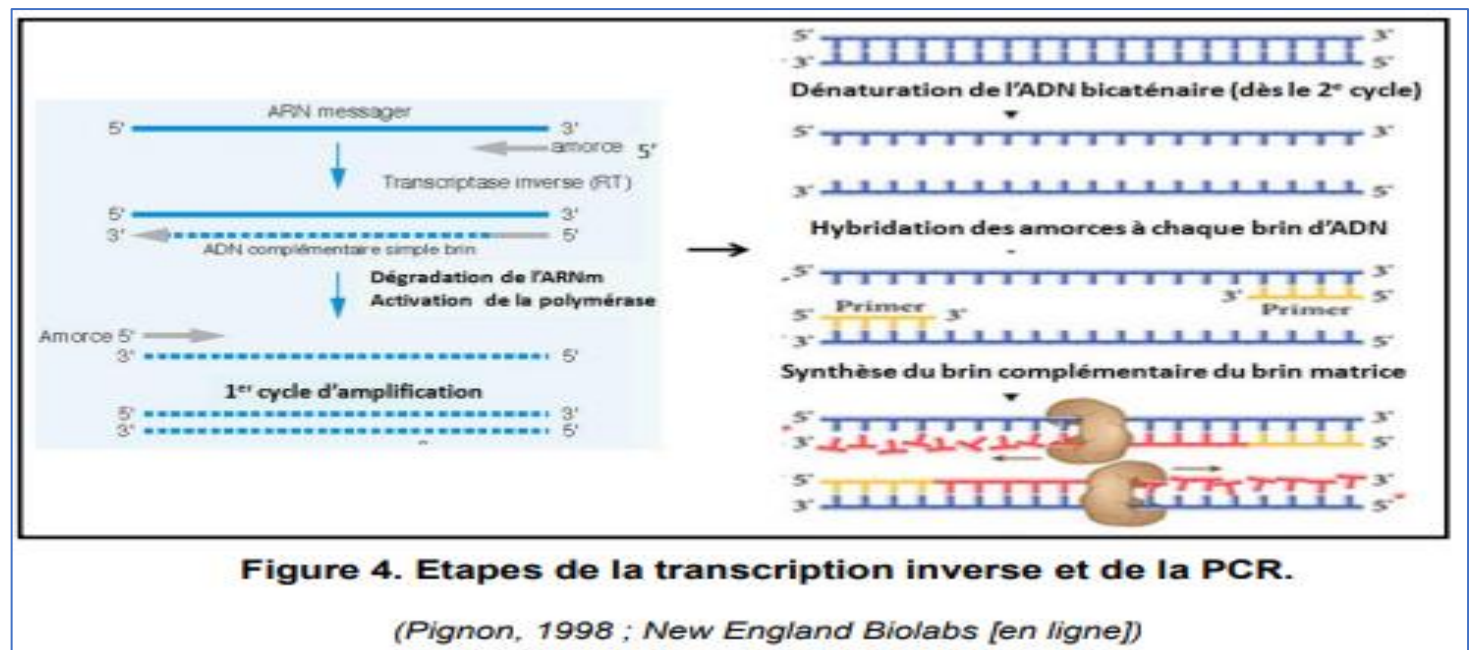
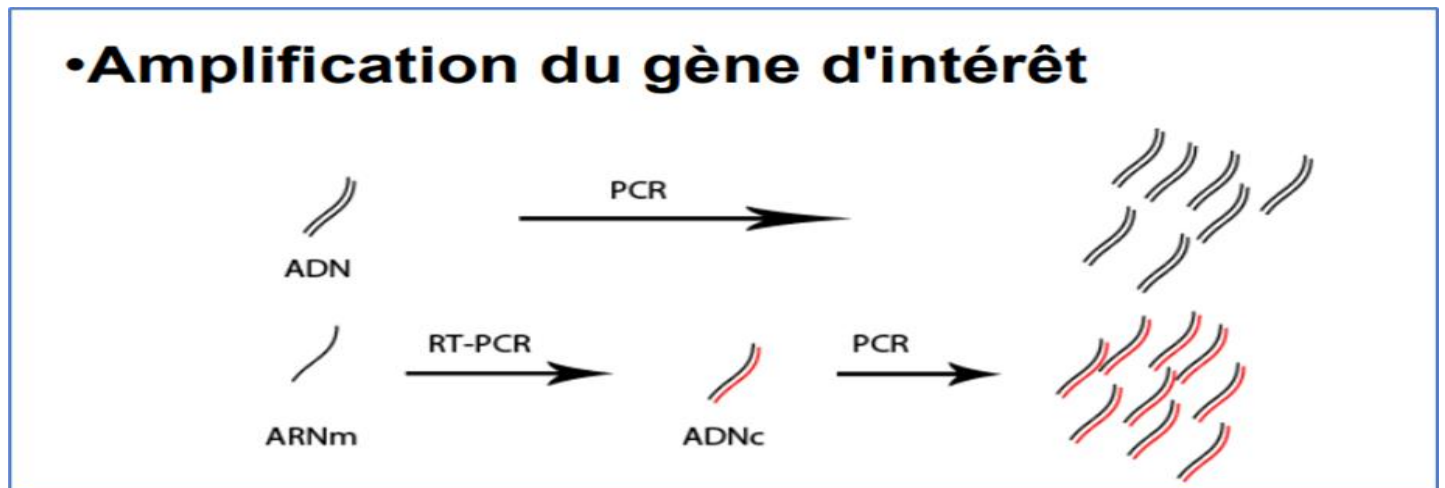


### Notions :

#### La PCR (Réaction de polymérisation en chaîne)

C'est une réaction enzymatique qui permet de sélectionner puis d'amplifier en une très grande quantité un fragment d'**ADN** particulier, présent en très faible quantité au départ, parmi des millions d'autres fragments.

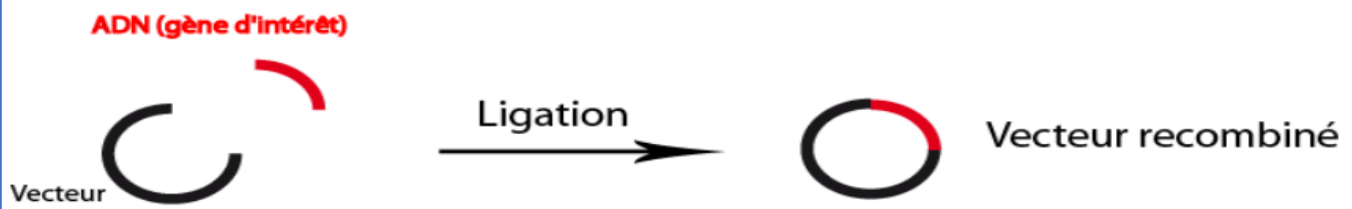
La **RT-PCR** est une technique qui permet de faire une PCR (réaction en chaîne par polymérase) à partir d'un échantillon d'ARN. L'ARN est tout d'abord rétrotranscrit grâce à une enzyme appelée transcriptase inverse, qui permet la synthèse de l'ADN complémentaire (ADNc). Ce dernier est ensuite utilisé pour réaliser une PCR.



### 1.3. Insertion du transgène dans un vecteur de clonage

Le gène d'intérêt est alors inséré dans un vecteur qui doit être capable de se répliquer et de répliquer le gène étranger qui lui est rattaché.

## • Insertion dans un vecteur



Vecteur	Cellule hôte	Taille du transgène (kb)
Plasmides	Bactérie	0,5 - 10
Bactériophages (virus infectant les bactéries)	Bactérie	10 - 20
Bacterial Artificial Chromosome	Bactérie	50 - 300
Yeast Artificial Chromosome	Levure	100 - 2000
Virus (adénovirus, rétrovirus, lentivirus, etc)	Cellule animale	variable

**Tableau II. Différents vecteurs de type ADN.**

(Perrin-Schmitt, 2011)

Les vecteurs couramment utilisés étant les plasmides (voir cours biologie cellulaire)

Les plasmides possèdent plusieurs caractéristiques justifiant de leur usage pour la transgénèse.

- En effet, compte tenu de leur petite taille,
- faciles à isoler, à purifier, et à réintroduire dans une autre cellule hôte.
- Leur origine de répllication permet le maintien du transgène dans la cellule hôte recombinante (CHR) et leurs marqueurs de sélection (gènes de résistance à des antibiotiques, gène lacZ, etc.) facilitent le criblage des CHR.
- De plus, les plasmides possèdent un site dit de « multi clonage » reconnu par de nombreuses enzymes de restriction. Celles-ci ont pour rôle d'exciser le vecteur et le transgène de façon spécifique et complémentaire afin que la ligase, une autre enzyme, puisse catalyser leur recombinaison (ligation) (figures 6 à 8).

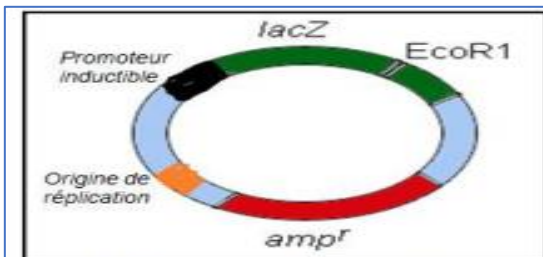


Figure 6. Plasmide contenant les gènes amp<sup>r</sup>, lacZ et le site de restriction EcoRI.

(Pronovost M, 2013 [en ligne])

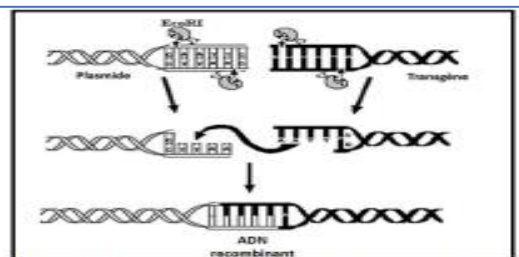


Figure 7. Mode d'action d'une enzyme de restriction. Exemple d'EcoRI.

(Pronovost M, 2013 [en ligne])

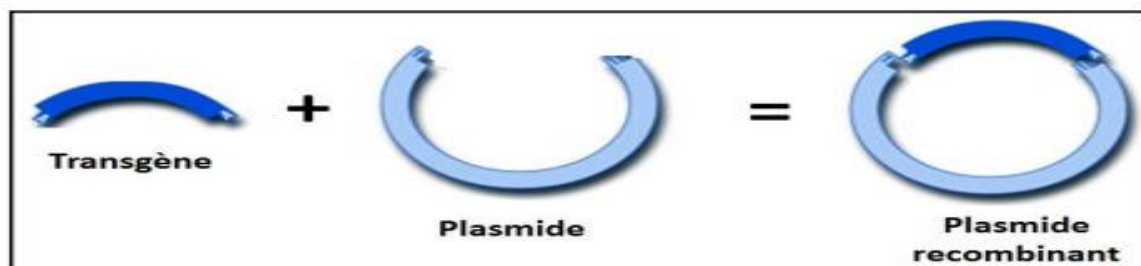
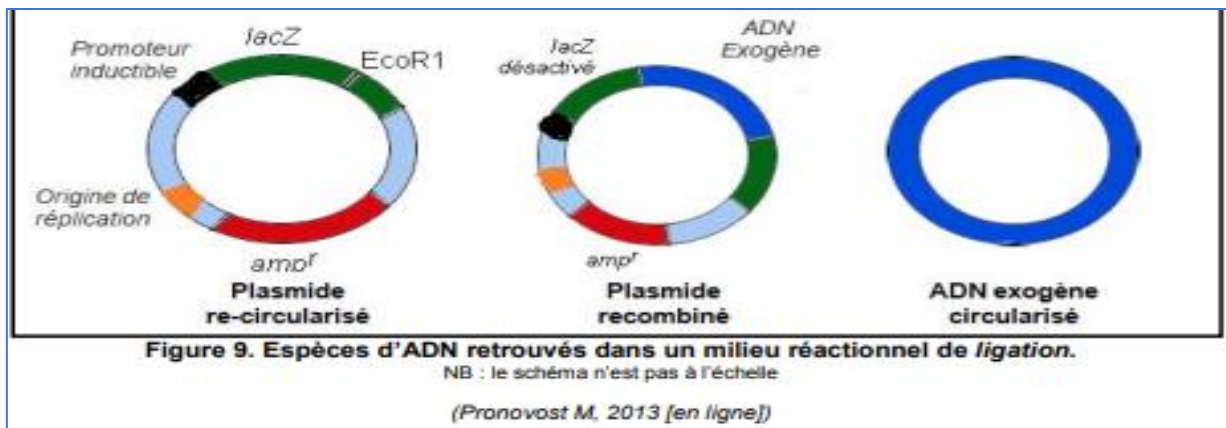


Figure 8. Recombinaison entre un transgène et un plasmide sous l'action de la ligase.

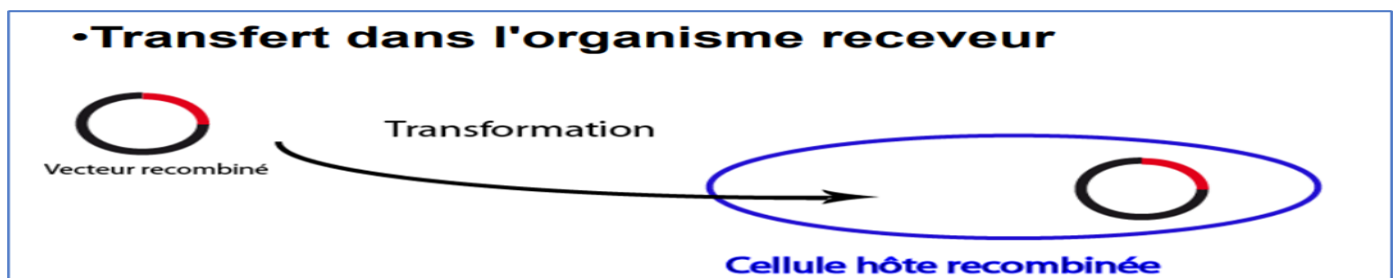
(Pronovost M, 2013 [en ligne])

Il faut noter que la ligation n'est pas effective à 100%. Il existe 3 « espèces » d'ADN dans le milieu post-ligation : le transgène (circularisé), le plasmide vide re-circularisé (si les extrémités sont compatibles) et le plasmide recombinant (figure 9).



#### 1.4. Insertion de l'ADN recombinant dans une cellule hôte

La molécule d'ADN recombinant obtenue (vecteur + gène d'intérêt) est ensuite introduite dans une cellule hôte, qui exécute les instructions qui lui sont fournies par le gène d'intérêt, et réalise les modifications post-traductionnelles dans l'objectif de fabriquer la protéine recherchée. L'hôte peut être une bactérie (E. coli), une levure, une cellule de mammifère ou encore une plante ou un animal transgénique ;



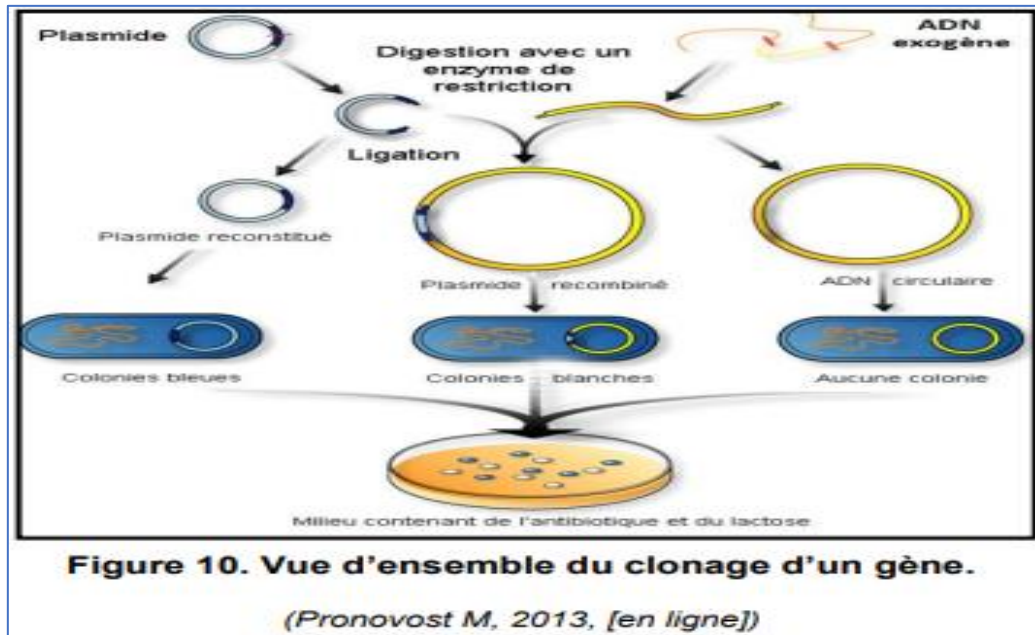
#### 1.5. Sélection des cellules hôtes recombinantes

L'insertion de l'ADNr dans les cellules hôtes n'est pas effective à 100% et à ce stade, il existe trois population de cellules :

- Les non modifiées,
- Celles qui ont internalisé le vecteur non recombinant (sans le transgène)
- Et celles qui ont internalisé le vecteur recombinant (avec le transgène).

Pour les séparer, elles sont cultivées dans un milieu sélectif contenant un antibiotique.

Dans le cas d'un système d'expression bactérien, le milieu sélectif peut aussi contenir du lactose si le vecteur utilisé contient le gène *lacZ*. Ce gène code une  $\beta$ -galactosidase, enzyme capable d'hydrolyser des  $\beta$ -galactosides et d'engendrer des colonies bactériennes bleues. Comme l'illustre la figure 10, seules les bactéries ayant internalisé un plasmide sont capables de croître sur milieu sélectif, car résistantes à l'antibiotique. Celles qui parmi elles, ont reçu le plasmide reconstitué, donc qui possèdent un gène *lacZ* intact, se colorent en bleu. Les autres, celles qui ont reçu le plasmide recombinant n'ont plus de gène *lacZ* (car le transgène a été inséré à cet endroit) et demeurent donc blanches.



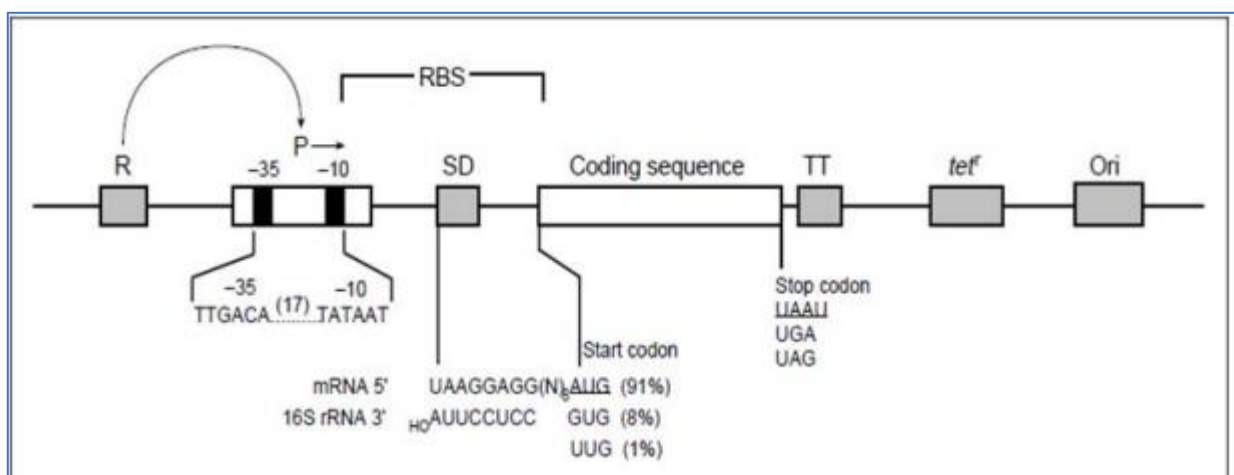
Une fois isolées, les CHR sont cultivées en masse afin de pouvoir extraire et purifier une quantité importante de plasmides contenant l'ADNr qui sera ensuite **sous-cloné dans un vecteur plasmidique d'expression**, lui-même inséré dans des cellules hôtes.

Plusieurs éléments doivent être pris en compte afin de garantir un bon niveau d'expression, une bonne solubilité et une bonne stabilité de la PtR à produire. Parmi ces éléments figurent **le choix du plasmide**, de **la souche bactérienne** et des **conditions de culture**, tous conditionnés par les propriétés de la PtR à exprimer (structure, taille, propriétés physico-chimiques, présence d'acides aminés codés par des codons rarement utilisés par l'hôte, présence de ponts disulfures, toxicité pour l'hôte, etc).

### 1.5.1. Garantir un bon niveau d'expression des protéines recombinantes

#### 1.5.1.1. Choix du plasmide d'expression

Un plasmide d'expression doit obligatoirement comporter les éléments suivants: une **origine de réplication**, des **marqueurs de sélection par antibiotique** ainsi que des éléments **d'initiation et d'arrêt de la transcription** et de la **traduction**. Il peut aussi contenir une ou plusieurs séquences codant des **partenaires de fusion**, éléments facilitant l'expression, la solubilité et la purification de la PtR .



**R** : répresseur du promoteur. **P** : promoteur (séquence -35 à -10 par rapport au codon d'initiation). **RBS** : site de fixation du ribosome lors de l'initiation de la traduction (la sous-unité 16S du ribosome fixe la séquence de Shine Delgarno (**SD**)). **Start / Stop codon** : codon d'initiation/de terminaison de la traduction, respectivement. **TT** : élément de terminaison de la transcription et de stabilisation des ARNm. **TET** : gène de résistance à un antibiotique (ici, la tétracycline). **Ori** : origine de réplication.

(Hannig and Makrides, 1998; Jana and Deb, 2005; Makrides, 1996)



## Dans le cas d'un système d'expression bactérien

### 1.5.1.2. Choix de la souche de production de protéines recombinantes

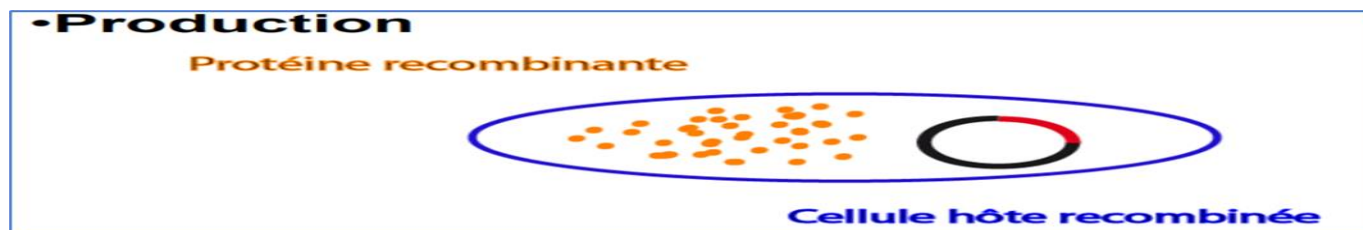
Il existe une large gamme de souches d'*E. coli* génétiquement modifiées pour faciliter la stabilité du plasmide et des ARNm, l'expression, la solubilité et la stabilité des PtR.

- La plupart sont déficientes en protéases (EX/ souches BL21(DE3), sont déficientes en **ompT** et **lon**, deux puissantes protéases, les protéines recombinantes produites dans ces souches ne subiront pas de protéolyse),
- Certaines possèdent des gènes codant des ARNt supplémentaires rares (EX : souches BL21 Codon Plus(DE3)-RIL, BL21 Codon Plus(DE3)-RP/RPIL et BL21(DE3)-Rosetta qui contiennent un plasmide codant des ARNt reconnaissant jusqu'à 6 codons rares (AGA/AGG = Arg ; Ile = AUA ; Leu = CUA ; Pro = CCC ; Gly = GGA),
- D'autres facilitent la formation de ponts disulfures (EX : souches AD494 et BL21-Origami, déficientes en deux réductases (thiorédoxine réductase (trxB-) ± glutathion réductase (trxB-/gor-) respectivement), permettent la formation de ponts disulfures dans le cytoplasme de la bactérie. Les PtR fusionnées à la thiorédoxine ou à la GST produites par ces souches sont bien repliées, plus solubles et faciles à purifier
- D'autres encore sont résistantes à la toxicité de certaines protéines (EX : souche BL21-AI possède dans son génome le gène de la T7 ARN polymérase qui est sous le contrôle du promoteur araBAD dont la régulation est très fine. Le mécanisme expliquant leur résistance à la toxicité des Pt exprimées reste inconnu.

**A la fin du clonage**, une partie des CHR est conservée à -80°C ou dans de l'azote liquide, sous forme de banque cellulaire, pour garantir la reproductibilité des lots d'ADNr cloné et des PtR exprimées.

### 1.6. Production de la protéine recombinante

Il s'agit alors de passer par une phase de production, généralement en fermenteur (en ce qui concerne les micro-organismes et cellules de mammifères), afin de produire les volumes de protéine nécessaires ;



- La protéine doit enfin être récupérée puis purifiée, cette étape faisant appel à des techniques plus classiques (séparation, extraction, purification) mais pouvant se révéler longue et couteuse ;
- Il s'ensuit des essais biochimiques et immunologiques (contrôle qualité) permettant de contrôler la pureté de la protéine obtenue, son activité, etc.





## 2. Systèmes d'expression des protéines recombinantes

Les principaux types de cellules ou d'organismes utilisés pour la transgénèse sont les bactéries, les levures, les cellules d'insectes associées au baculovirus, les cellules de mammifères, les animaux et plantes transgéniques



### Récapitulatif des systèmes d'expression de protéines recombinantes

Le tableau ci-dessous récapitule les principaux avantages et inconvénients des systèmes d'expression de PtRIT habituellement utilisés.

**Le système d'expression est choisi** en tenant compte de la nature et de l'utilisation finale de la protéine. Si la PtR exprimée nécessite toute forme de modifications post-traductionnelles, le choix d'un système bactérien ne sera pas optimal.

**Un système eucaryote** sera plus adapté : levure, cellules d'insectes ou de mammifères en fonction du degré de complexité des modifications requises.

**Par exemple, pour la sécrétion efficace** de protéines comportant des ponts disulfures, la levure est suffisante, tandis que pour exprimer des protéines glycosylées, les cellules d'insectes sont plus adaptées que la levure.

**Les cellules de mammifères** ou les animaux transgéniques, bien que coûteux, demeurent très étudiés. De gros efforts sont portés à l'amélioration de leur productivité car ils permettent de produire des molécules thérapeutiques complexes.

**Pour la production** de petites PtR ne requérant aucune modification post-traductionnelle, les bactéries constituent le système idéal : le procédé est rapide, le coût total de la production est peu élevé, la production est facilement industrialisable et les contraintes réglementaires ne sont pas drastiques.

Hôtes	Avantages	Inconvénients
<b>Bactéries</b> ( <i>E. coli</i> )	<ul style="list-style-type: none"> <li>- Bonne documentation sur le génome</li> <li>- Culture facile, peu chère</li> <li>- Expression rapide du gène d'intérêt</li> <li>- Rendements importants</li> </ul>	<ul style="list-style-type: none"> <li>- Pas de modifications post-traductionnelles</li> <li>- Faible potentiel de sécrétion</li> <li>- Formation de corps d'inclusions (agrégats de protéines)</li> <li>- Présence de protéases et d'endotoxines</li> </ul>
<b>Levures</b> ( <i>S. cerevisiae</i> )	<ul style="list-style-type: none"> <li>- Culture facile, peu chère</li> <li>- Potentiel de sécrétion</li> <li>- Rendements importants</li> <li>- Modifications post-traductionnelles élaborées (repliement, glycosylation, ponts disulfure)</li> </ul>	<ul style="list-style-type: none"> <li>- Difficulté à sécréter des grosses molécules</li> <li>- Parfois problème d'isolement de la PtR</li> <li>- Glycosylation simple</li> <li>- Glycosylation immunogène (hypermannosylation)</li> <li>- Utilisation de méthanol dans la culture (<i>P. pastoris</i>)</li> </ul>
<b>Cellules d'insectes</b>	<ul style="list-style-type: none"> <li>- Potentiel de sécrétion, rendements importants</li> <li>- Modifications post-traductionnelles élaborées (repliement, glycosylation, ponts disulfure, acylations, phosphorylations, etc.)</li> <li>- Flexibilité de la taille de la PtR à produire</li> <li>- Expression simultanée de plusieurs gènes</li> <li>- Pas de risque biologique</li> </ul>	<ul style="list-style-type: none"> <li>- Parfois problème d'isolement de la PtR</li> <li>- Parfois production de PtR mal repliées</li> <li>- Glycosylation parfois incorrecte (hypo ou hyper)</li> <li>- Présence de protéases</li> </ul>
<b>Cellules de mammifères</b>	<ul style="list-style-type: none"> <li>- Glycosylation + modifications complexes</li> <li>- Potentiel de sécrétion</li> </ul>	<ul style="list-style-type: none"> <li>- Culture spécifique, chère, difficile à grande échelle</li> <li>- Parfois problème d'isolement de la PtR</li> <li>- Risque biologique, immunogénécité</li> </ul>
<b>Animaux transgéniques</b>	<ul style="list-style-type: none"> <li>- Glycosylation + modifications complexes</li> <li>- Sécrétion dans liquides biologiques, rendements importants</li> </ul>	<ul style="list-style-type: none"> <li>- Coût élevé, problème éthique + gestion des troupeaux</li> <li>- Parfois problème d'isolement de la PtR</li> </ul>
<b>Plantes transgéniques</b>	<ul style="list-style-type: none"> <li>- Culture à grande échelle, simple et peu chère</li> <li>- Glycosylation complexe</li> <li>- Pas de risque biologique</li> </ul>	<ul style="list-style-type: none"> <li>- Difficulté d'extraction et de purification</li> <li>- Glycosylation parfois immunogène (xylose)</li> <li>- Problème éthique et environnemental</li> </ul>

**Tableau IX. Principaux avantages et inconvénients des systèmes d'expression de PtRIT.**

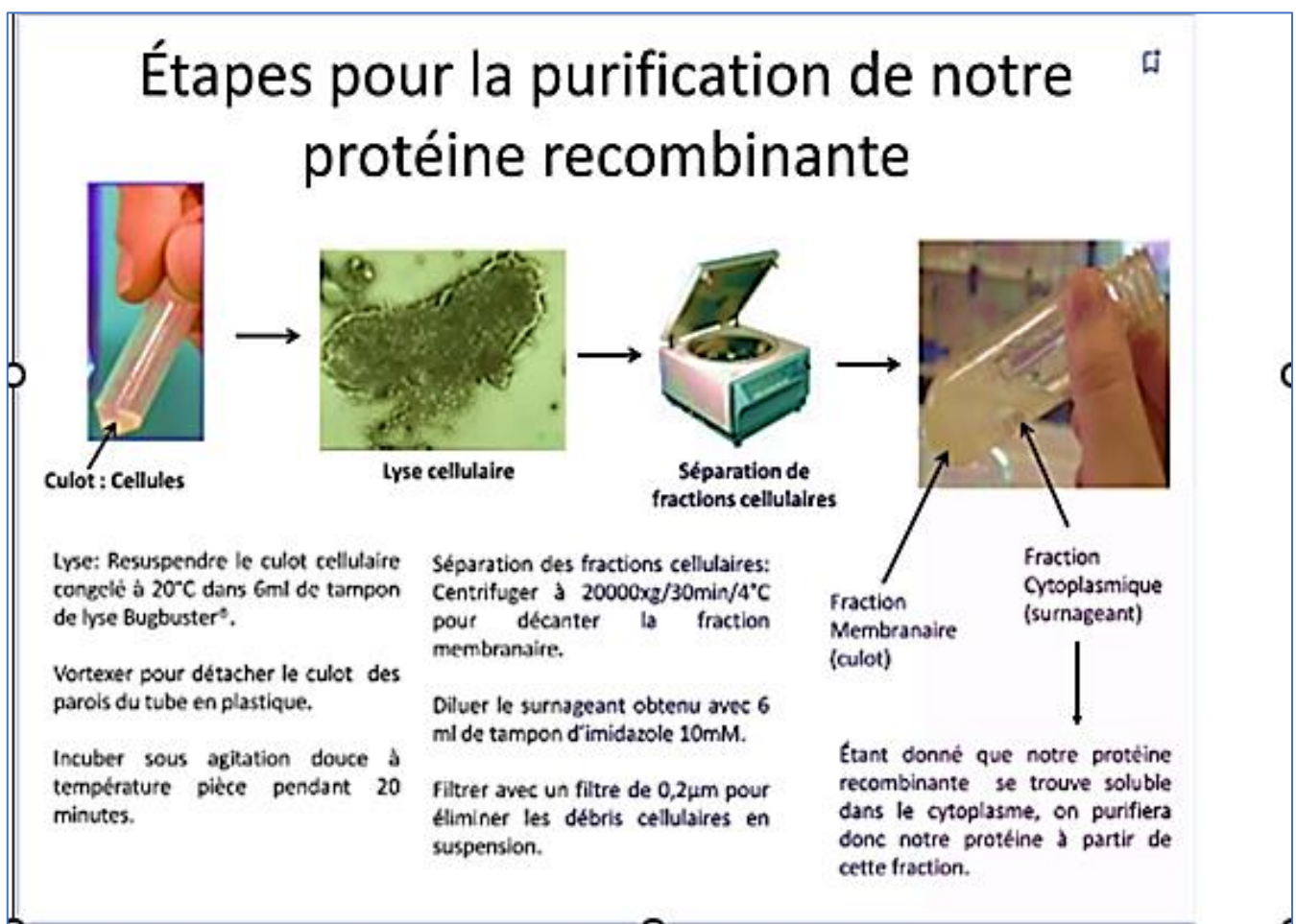
(Demain and Vaishnav, 2009; Philippe et al., 2009)

## Bioprocédés, procédés de purification, formulation

Les procédés d'extraction ou de purification de biomolécules, regroupent les opérations de séparation des constituants des milieux, d'extraction des biomolécules, de leur purification et de leur stabilisation/formulation. Ces procédés sont nécessaires afin de compléter les bioprocédés et récupérer les molécules d'intérêt afin de les valoriser.

### 1. Etape préalable à la purification des protéines recombinantes : Clarification

L'étape préalable à la purification consiste à isoler les Pt solubles des constituants solides. Si la PtR est sécrétée dans le milieu de culture, une simple centrifugation ou une filtration est suffisante pour obtenir un échantillon clarifié prêt à être purifié. Cependant, dans la plupart des cas, la PtR doit être extraite de la cellule ou de l'organisme hôte après une lyse physique, mécanique, enzymatique ou thermique.



Les conditions de lyse doivent permettre d'éviter l'oxydation et la protéolyse, mais aussi la contamination par l'ADN génomique. Le lysat obtenu est ensuite centrifugé et le surnageant est recueilli pour la purification.

## 2. Purification des protéines recombinantes

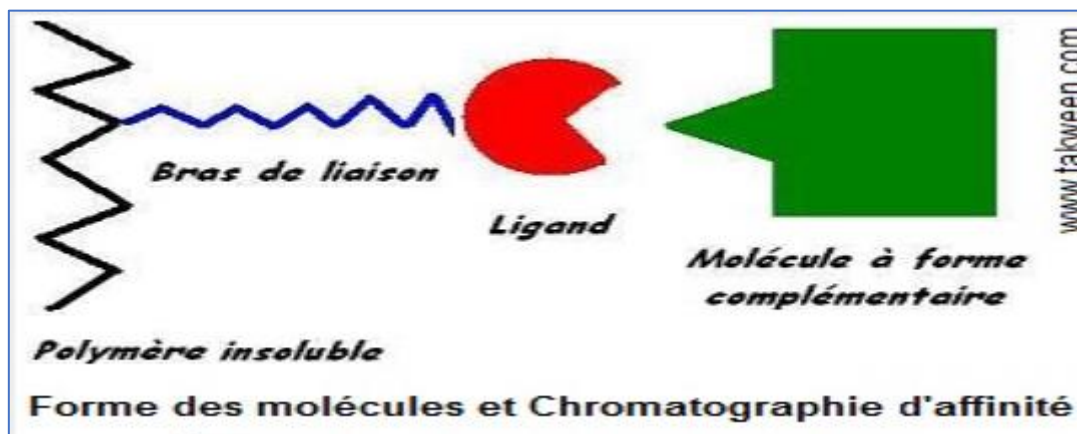
La purification d'une PtR permet non seulement d'isoler celle-ci du reste des Pt de la cellule ou de l'organisme hôte, mais aussi de la concentrer. Cette séparation est possible grâce à la différence de propriétés physicochimiques entre la PtR et les Pt de la cellule ou de l'organisme hôte (affinité pour un ligand, charge ionique, taille, hydrophobicité, etc.).

La méthode la plus utilisée est la chromatographie d'affinité qui, en général, suffit pour obtenir un bon degré de pureté de la PtR.

Néanmoins, si un plus grand degré de pureté est nécessaire, d'autres techniques peuvent être envisagées : chromatographie liquide haute performance, chromatographie d'échange d'ions (CEI), chromatographie d'exclusion stérique (CES).

### 2.1. Chromatographie d'affinité

La chromatographie d'affinité permet d'isoler la PtR grâce à son affinité pour un ligand (une immunoglobuline, un composé protéique ou osidique, un ion métallique, etc) immobilisé sur une matrice solide. Cette technique repose sur la fixation spécifique et réversible de la PtR au ligand lié à la matrice. Le ligand peut soit se lier directement à la protéine d'intérêt ou à un tag qui est attaché de manière covalente à la protéine.



La chromatographie d'affinité est une méthode de purification plus robuste. Généralement, elle est utilisée dans les premiers stades de purification.

Pour la chromatographie d'affinité, la phase stationnaire est réalisée sur une matrice inerte (agarose ou polyacrylamide) fixée de manière covalente à un ligand qui se lie spécifiquement à une protéine ou un groupe de protéines.

Il existe deux types de chromatographie d'affinité dont la chromatographie d'affinité sélective et la chromatographie d'affinité non sélective.

- **Chromatographie d'affinité sélective**: un ligand spécifique d'une protéine ou d'un tag lié de manière covalente est utilisé.

- **Chromatographie d'affinité non sélective**: comme la protéine A, G, L pour les immunoglobulines, ou de l'héparine pour des protéines liant l'ADN, ou la lectine pour les glycoprotéines, **le ligand se lie à un groupe de protéines avec des capacités de liaison similaires.**

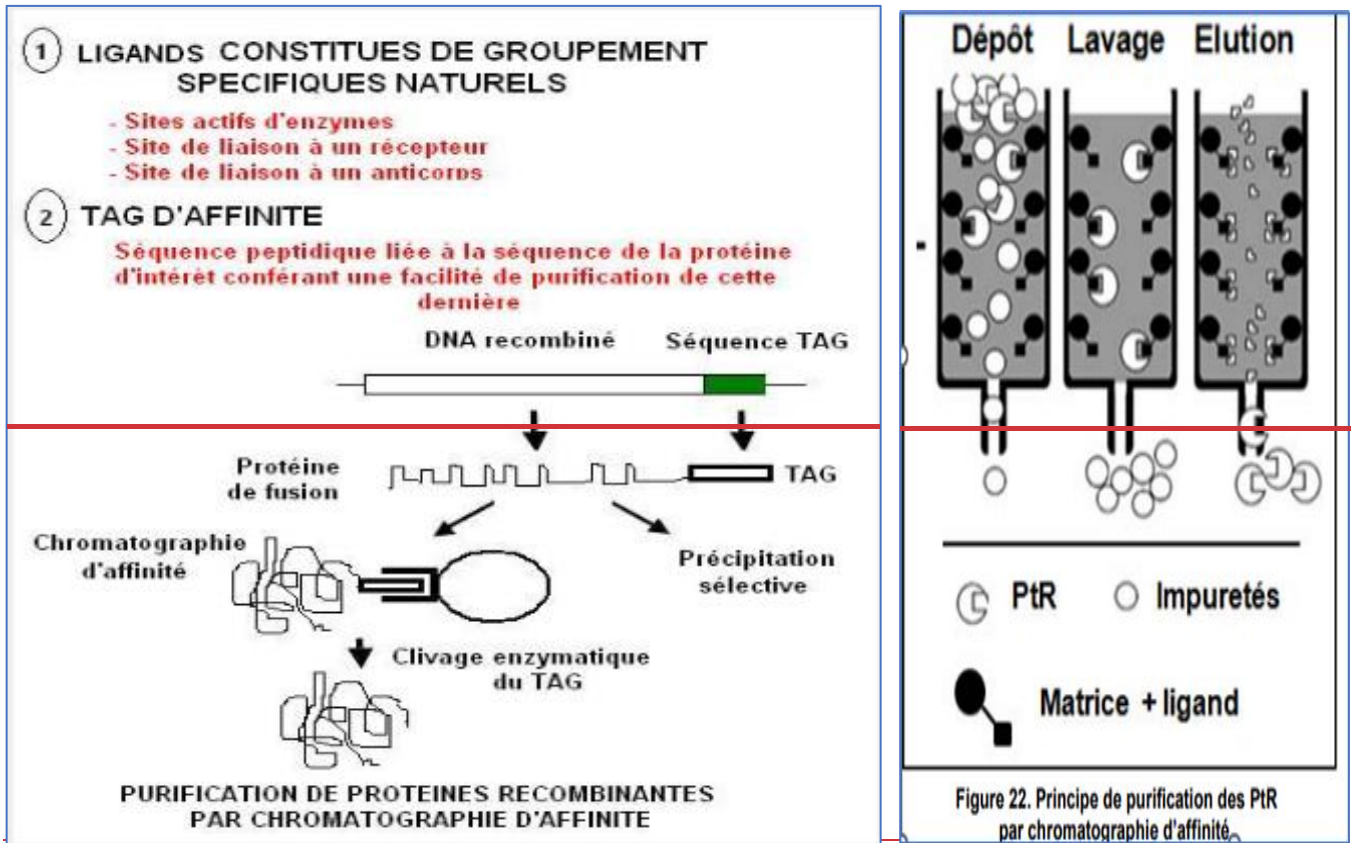
En chromatographie d'affinité, les protéines sont chargées sur la colonne dans des conditions qui influencent la liaison entre la protéine (ou tag) et son ligand.

Le lavage de la protéine liée se fait dans des conditions ne perturbent pas l'interaction spécifique, mais pouvant perturber les interactions non spécifiques entre des protéines contaminantes et la phase stationnaire.

L'élution de la protéine liée est fait avec un tampon contenant une molécule concurrente ou des conditions perturbant les interactions protéine / protéine. La molécule concurrente se lie au ligand et déplace la protéine d'intérêt.



Exemple d'application de la chromatographie d'affinité. Purification de [protéines recombinantes](#)



### 3. La formulation des protéines recombinantes

Les protéines recombinantes sont complexes et extrêmement sensibles à toute modification de leur environnement. La stabilité d'une protéine est donc un facteur à surveiller durant tout le cycle de vie du médicament et c'est pour cela que beaucoup de recherches sont en œuvre afin de l'améliorer. La conformation la plus stable (de moindre énergie) étant celle de l'état natif (N).

De nombreux facteurs sont susceptibles d'influencer la stabilité d'une protéine. De plus les protéines et peptides sont susceptibles d'avoir des sites multiples de réaction conduisant à leur instabilité.

Les principales réactions menant à l'instabilité d'une protéine ou d'un peptide peuvent être d'ordre chimique ou physique.

**Les instabilités chimiques** sont définies comme toute réaction impliquant une modification (clivage ou formation) d'une ou plusieurs liaisons de la protéine et conduisant à une nouvelle entité chimique.

**Les instabilités physiques** quant à elles engendrent un changement dans les structures secondaires et d'ordres plus élevés de la protéine.

Instabilités physiques	Instabilités chimiques
<ul style="list-style-type: none"> <li>• Dénaturation</li> <li>• Adsorption</li> <li>• Aggrégation</li> <li>• Précipitation</li> </ul>	<ul style="list-style-type: none"> <li>• Hydrolyse</li> <li>• Oxidation</li> <li>• Racémisation</li> <li>• Isomérisation</li> <li>• <math>\beta</math>-Élimination</li> </ul>

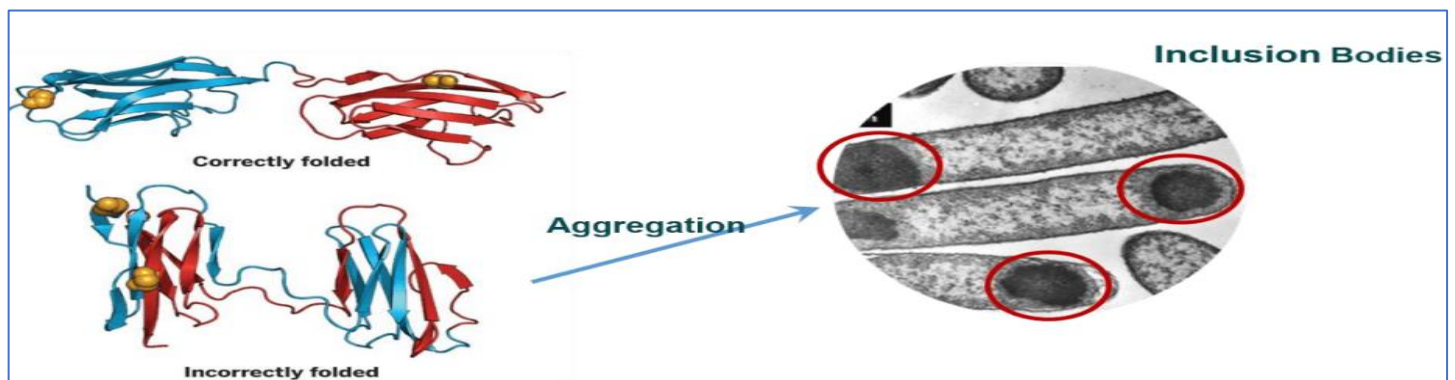
Figure 7 : Instabilités des protéines et peptides

La formulation de produits biopharmaceutiques vise à stabiliser la conformation des protéines, à promouvoir leur efficacité et à prévenir les problèmes de sécurité, tels que l'immunogénicité. Leur stabilité est critique quelle que soit l'étape de fabrication du principe actif et doit être strictement contrôlée.

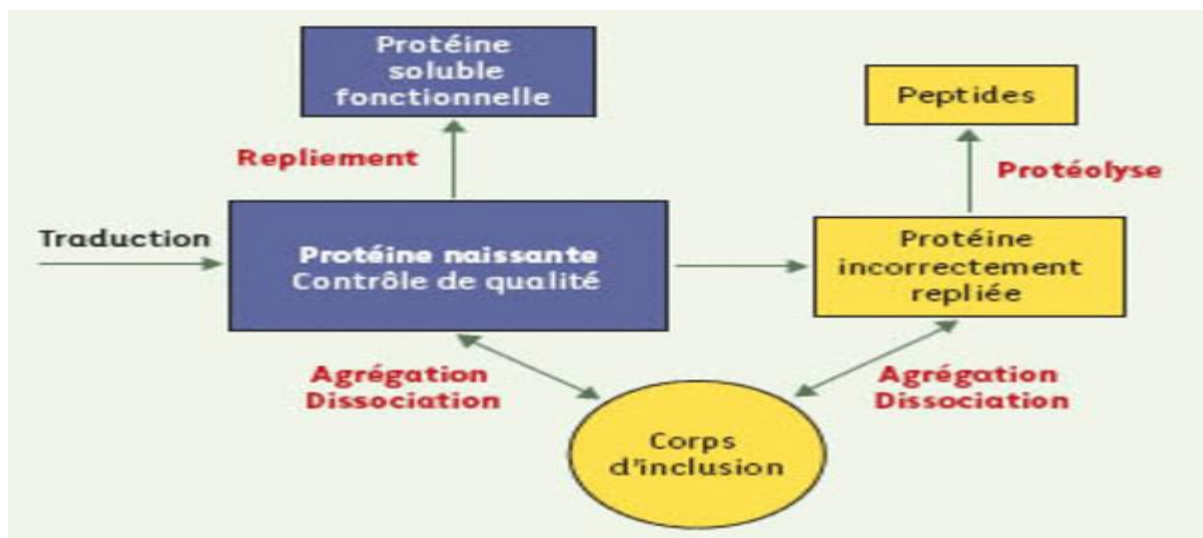
### A. Formation des corps d'inclusion

Pour remplir leur fonction, les protéines doivent adopter une conformation tridimensionnelle précise, l'état natif, acquise au cours du repliement cellulaire. De nombreuses protéines ne sont pas produites dans leur état natif, mais s'agrègent dans un état le plus souvent biologiquement inactif, que l'on appelle corps d'inclusion. Ce phénomène d'agrégation est le principal obstacle à la production de protéines.

Le mécanisme moléculaire de l'agrégation des protéines recombinantes suggère que les chaînes polypeptidiques naissantes, adoptant des conformations partiellement repliées et instables (les intermédiaires de repliement), peuvent emprunter deux voies alternatives et compétitives : le repliement correct, qui conduit à la forme native, ou le repliement incorrect, qui conduit à l'agrégation.



Les corps d'inclusion sont des agrégats protéiques non ordonnés, relativement homogènes et denses ; ils sont de plus en plus considérés comme des  **systèmes en équilibre dynamique entre association et dissociation**  d'espèces partiellement ou incorrectement repliées (Figure). La concentration très élevée des chaînes polypeptidiques naissantes dans une cellule recombinante favorise les réactions d'agrégation aux dépens de la réaction productive vers l'état natif.



Dans le contexte cellulaire, le repliement est facilité par un ensemble d'acteurs :

**Les chaperons**, qui assistent les protéines en cours de repliement



**Des catalyseurs**, qui interviennent sur les étapes lentes de ce processus, **les oxydoréductases**, qui catalysent la **formation des ponts disulfures**, et **les peptidyl-prolyl isomérases**, qui catalysent **l'isomérisation *cis-trans*** des liaisons peptidiques.

En agissant sur les conformations intermédiaires, les protéines chaperons contrôlent directement le processus de repliement, en collaboration avec les protéases cellulaires. En effet, les protéines incorrectement repliées sont également les substrats cellulaires de systèmes de dégradation en assurant l'élimination rapide.

Dans certains cas, l'agrégation résulte de la **disponibilité limitée des chaperons** et induit une **réponse au stress** dans la cellule qui **produit des quantités non physiologiques** d'une **protéine inutile**. Cependant, la formation des **corps d'inclusion** peut avoir un effet bénéfique en **protégeant la protéine recombinante** de la dégradation par les **protéases** cellulaires.

## **B. Renaturation des protéines *in vitro***

Bien qu'il n'existe pas d'approche universelle favorisant le repliement correct des protéines qui s'agrègent, plusieurs techniques permettant l'obtention d'une forme native sont couramment utilisées, avec des taux de succès très variés. Parmi elles, on distingue

Celles qui visent à renaturer la protéine *in vitro* à partir des corps d'inclusion purifiés

Et celles qui interviennent directement sur l'expression cellulaire afin de minimiser les réactions d'agrégation.

La purification des corps d'inclusion est rarement une étape critique, dans la mesure où ces agrégats sont facilement isolés par centrifugation après lyse cellulaire et contiennent la protéine recombinante à un degré de pureté très élevé.

L'étape suivante de solubilisation est en revanche souvent plus délicate, car elle nécessite la dénaturation chimique complète de la protéine agrégée par des dénaturants forts comme l'urée 8 M ou la guanidine 6 M, éventuellement en présence d'un agent réducteur.

Enfin, intervient l'étape critique d'élimination du dénaturant qui, réalisée par dialyse ou forte dilution, favorise la renaturation de la protéine. À cette étape, certaines petites molécules peuvent être additionnées à la solution de renaturation, comme des détergents, des alcools, des amides, ou de l'arginine. En général, ces additifs, dont le mécanisme d'action est souvent peu documenté, diminuent l'agrégation des protéines en inhibant leurs interactions intermoléculaires. De même, les paramètres de température et de concentration de protéine sont contrôlés pour favoriser le repliement correct aux dépens de l'agrégation. Pour les protéines possédant une séquence polyhistidine, la chromatographie de chélation métallique constitue un troisième moyen permettant l'élimination du dénaturant tout en maintenant la protéine fixée. Cette méthode permet une séparation physique des protéines lors de la renaturation sur la colonne, et limite donc les réactions d'agrégation.

Les inconvénients de ces diverses techniques de renaturation *in vitro* sont souvent les faibles taux de renaturation obtenus, les difficultés d'optimisation des conditions expérimentales spécifiques de chaque protéine et la précipitation de la protéine renaturée. Pour ces raisons, il est souvent souhaitable d'intervenir en amont, sur le système d'expression ou sur le gène codant pour la protéine, afin de promouvoir le repliement cellulaire.

## **C. Minimiser l'agrégation *in vivo***

Les méthodes d'optimisation du repliement productif des protéines recombinantes peuvent être schématiquement présentées en deux catégories.

**Dans la première, les** facteurs cellulaires et les paramètres de culture qui influencent l'expression et le repliement des protéines sont systématiquement testés. Parmi ceux-ci, citons le contexte génétique de la cellule recombinante, la composition du milieu culture, la température de croissance des cellules, le nombre de copies du gène recombinant par cellule et son niveau de transcription ou de traduction, la surexpression des chaperons et l'inactivation des gènes codant pour les protéases.

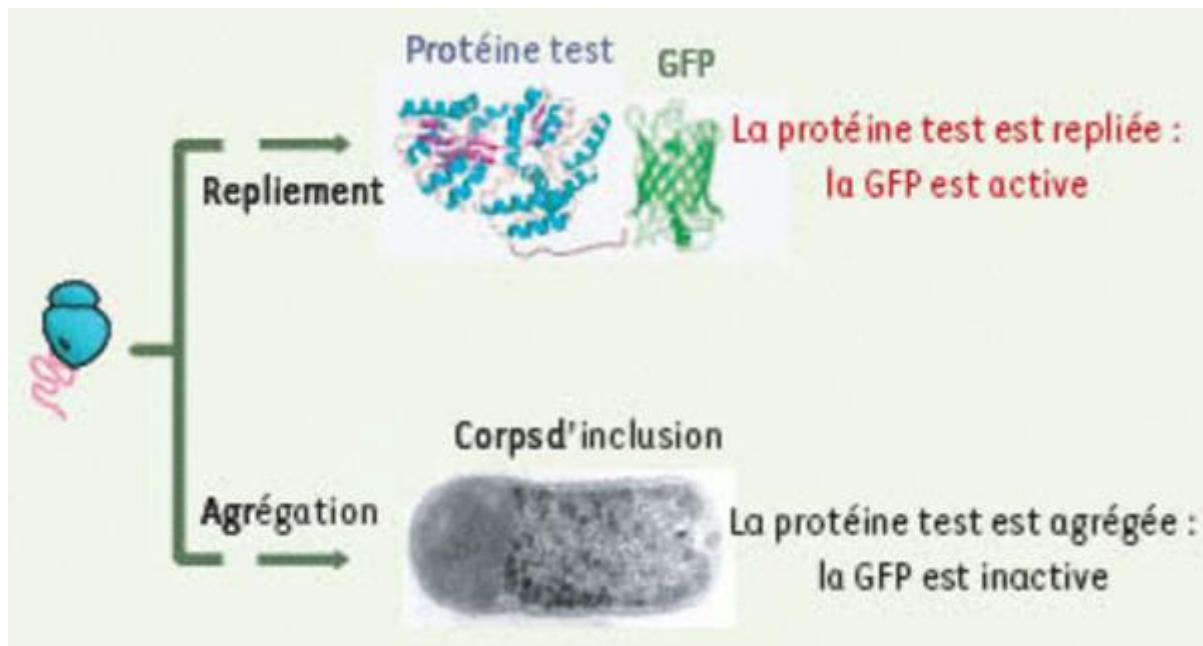
En règle générale, l'élévation de température augmente la vitesse d'agrégation, puisque cette réaction multimoléculaire est dominée par les interactions hydrophobes. En diminuant la température de croissance et le niveau de production, le repliement correct de la protéine sera donc favorisé. La variation de ces deux paramètres donne souvent des résultats encourageants.

Cependant, la manipulation de ces facteurs extrinsèques n'est pas toujours suffisante, et l'on peut recourir à une stratégie qui modifie génétiquement la séquence de la protéine recombinante, par l'ingénierie des protéines.

**La première approche**, utilisée le plus souvent pour faciliter leur purification, est la construction d'une protéine de fusion. Il est fréquemment observé qu'une protéine fusionnée à une protéine d'intérêt par son extrémité aminoterminal a une influence positive sur le repliement de cette dernière. Pour cette application, les systèmes de protéines de fusion les plus utilisées sont la protéine affine du maltose (MBP), la glutathion S-transférase ou encore NusA pour l'expression chez *E. coli*. Cette stratégie augmente la stabilité intracellulaire de la protéine recombinante et simplifie sa purification, qui est réalisée en une simple étape chromatographique sur une colonne d'affinité contenant le ligand immobilisé de la protéine de fusion. Cependant, la conservation de ces propriétés bénéfiques n'est pas toujours assurée après purification et clivage du partenaire de fusion.

**Une seconde approche** est de rechercher, à l'aide des techniques d'évolution moléculaire dirigée, des variants de la protéine d'intérêt qui ne s'agrègent plus. Bien sûr, les systèmes d'expression bactérien se prêtent plus facilement à ce type d'approche expérimentale. Le principe est relativement simple et repose sur **une propriété fonctionnelle spécifique de la protéine de fusion utilisée, indépendante de la protéine cible**. Dans ces approches, la protéine de fusion la plus documentée est **la protéine à fluorescence verte (GFP)**, dont les **propriétés de fluorescence** sont acquises par sa **structure native** et peuvent être directement reliées à la solubilité intracellulaire des protéines (*Figure 3*). En effet, si une protéine d'intérêt fusionnée à la GFP s'agrège, elle entraînera dans les corps d'inclusion la GFP qui sera incapable de se replier : la cellule exprimant la fusion ne présentera pas cette fluorescence verte caractéristique.

À l'inverse, si la **protéine d'intérêt se replie correctement**, la **fluorescence spécifique de la GFP native sera émise par la cellule qui produit cette fusion**. À partir d'une banque de mutants du gène cible, contenant des mutations ponctuelles, les variants fluorescents, donc solubles, seront ainsi isolés.



### Principe d'une fusion par la GFP (green fluorescent protein) pour tester le repliement cellulaire des protéines recombinantes.

La fluorescence spécifique de la GFP est acquise uniquement à partir de sa structure tertiaire correctement repliée. Si le repliement d'une protéine test, fusionnée à la GFP, conduit à sa structure native, les bactéries produisant la fusion seront fluorescentes. Au contraire, si la protéine test s'agrège au cours du repliement, la formation des corps d'inclusion (visualisés par microscopie électronique) bloquera le repliement de la GFP, et les bactéries ne seront pas fluorescentes.

Généralement, les substitutions d'acides aminés qui accélèrent le repliement vers la forme native de la protéine sont localisées en surface, et contribuent à une plus forte solubilité de la protéine sans toutefois modifier sa fonction.

### 3.1. Les enjeux de la formulation

Il est essentiel que la protéine conserve son activité dans l'environnement dans lequel elle sera formulée. La caractérisation de la protéine est donc primordiale avant même d'élaborer sa formule car elle permettra de

- Déterminer son environnement idéal (pH, température etc...) et les facteurs conduisant à son instabilité.
- Ses propriétés physico-chimiques : Masse molaire, AA, point isoélectrique, température de fusion  $T_m$
- Son empreinte, souvent réalisée par Dichroïsme Circulaire, technique qui permet de visualiser le repliement de la protéine. La protéine formulée sera ainsi ré-analysée et comparée à l'empreinte réalisée au temps initial.
- Sa pureté, ses impuretés et contaminants. Les produits de dégradation des protéines thérapeutiques ne sont pas aisément détectés par une seule technique analytique ; les méthodes de détection les plus utilisés sont la spectrométrie de masse et l'HPLC.
- Son activité biologique

Connaître les principales caractéristiques d'une protéine thérapeutique permettra par exemple

- **De connaître sa concentration critique** par le biais d'un test souvent réalisé par dosage colorimétrique.
- **L'agrégation des protéines est concentration-dépendante.** De fortes concentrations protéiques conduisent à l'instabilité de celles-ci mais l'inverse est également possible.
- **La solubilité d'une protéine** est la capacité de ses résidus polaires à interagir avec l'eau. Les résidus hydrophobes sont donc repliés à l'intérieur et la protéine garde sa structure active.

Connaître ces principales caractéristiques permettra également de présélectionner les excipients (Les ajusteurs de pH aussi appelés agents tampons peuvent être des AA tels que la glycine ou l'histidine, L'histidine et la méthionine, ...) qui seront à même de garantir la stabilité de la protéine sous les conditions de stockage qui seront définies. Ces excipients ne doivent pas dégrader la protéine (interactions protéine-excipient néfastes) et d'autre part les excipients doivent être tolérés et autorisés pour être injectés dans le cas de protéines thérapeutiques destinée à la voie injectable.

### 3.2. Les enjeux du procédé de production et du conditionnement primaire

Lors du développement du procédé de production de protéines thérapeutiques, trois paramètres de procédé majeurs impactent la stabilité de la protéine :

**L'agitation, le chauffage et les pressions élevées.**

En effet, les opérations d'agitation et de cisaillement engendrent une exposition des parties hydrophobes dans le milieu environnant de la protéine, initiant l'agrégation.

L'application de hautes pressions favorise également le dépliement des protéines.

Les protéines sont des molécules ayant des propriétés de surface. Elles sont connues pour s'adsorber et s'accumuler aux interfaces, telles que les interfaces solide/liquide ou air/liquide.

Les réactions au niveau de ces interfaces entraînent très souvent une détérioration, des perturbations structurales, des dégradations chimiques et/ou une agrégation de la protéine. Ces interfaces peuvent avoir une surface hydrophile ou hydrophobe. Au contact de surfaces hydrophiles, les protéines s'adsorbent par attractions électrostatiques entre les groupes chargés de la protéine et de la surface en contact avec la protéine, entraînant leur agrégation. Au contact de surfaces hydrophobes, les protéines subissent un réarrangement structural lent dans le but de maximiser les interactions hydrophobes, entraînant encore une fois leur agrégation. Idéalement, toutes les étapes de fabrication du biomédicament devraient faire l'objet d'une analyse de risque afin d'évaluer l'impact d'une quelconque action sur la stabilité de la protéine.

Plus récemment, la réglementation est plus stricte sur les matériaux utilisés lors de la production (en contact direct avec le principe actif), rentrant dans la composition du conditionnement primaire et du dispositif d'administration (comme les seringues par exemple).

En effet, certains matériaux sont susceptibles de relarguer des **composés appelés « leachables »** ; Composés **provenant d'un matériau et migrant vers la formulation** sous des **conditions normales** d'utilisation ou de stockage»

Ou « **extractables** ; Composés qui peuvent être relargués d'un matériau dans des conditions extrêmes (température, temps et solvants).». **Ces composés sont à titre d'exemple les ions métalliques**. Ces composés peuvent **mettre en péril la sécurité et l'efficacité du médicament**.

Ils peuvent en effet **provoquer une toxicité aiguë** à l'injection ou une **toxicité à long terme** lorsque **l'exposition est répétée**. Ces composés **peuvent également interagir avec les protéines** ou **les excipients** engendrant une **altération des propriétés physico-chimiques**, de la **stabilité du principe actif** ou bien de la **demi-vie du biomédicament**

Les industries pharmaceutiques investissent maintenant beaucoup de ressources afin d'identifier, quantifier et minimiser le taux d'impuretés dans leur produit final.

#### **Stratégies d'amélioration des protéines thérapeutiques**

Il est essentiel de comprendre les facteurs qui influencent la solubilité des protéines. Les protéines peuvent être difficiles à solubiliser en raison de leurs structures complexes, de leurs régions hydrophobes et de leur propension à s'agréger. Ces défis peuvent entraîner une biodisponibilité réduite, une efficacité réduite et des problèmes de stabilité dans diverses applications. En conséquence, l'amélioration de la solubilité des protéines n'est pas seulement une entreprise scientifique mais également une nécessité pratique pour de nombreuses industries.

#### **Techniques Avancées Pour L'amélioration De La Solubilité**

Les chercheurs ont développé plusieurs méthodes innovantes pour améliorer la solubilité des protéines. Ces techniques vont du génie génétique à des stratégies de formulation. Voici quelques-unes des approches les plus prometteuses :

#### **Modification génétique et ingénierie des protéines**

- **Evolution dirigée** : Cette technique implique le processus itératif de mutation d'une protéine et de sélection de variantes présentant une solubilité améliorée. L'évolution dirigée imite la sélection naturelle en laboratoire, conduisant à des protéines aux propriétés améliorées.

- **Protéines de fusion** : La fixation d'une protéine ou d'une étiquette peptidique hautement soluble à une protéine cible peut augmenter considérablement sa solubilité. Des étiquettes telles que la protéine liant le maltose (MBP) et la glutathion S-transférase (GST) sont couramment utilisées.
- **Ingénierie des surfaces** : Modifier les acides aminés de surface d'une protéine pour augmenter sa polarité peut améliorer la solubilité. Cette méthode implique souvent de remplacer les résidus hydrophobes par des résidus hydrophiles.

#### Modification chimique

- **Pégylation** : La fixation covalente des chaînes de polyéthylène glycol (PEG) aux protéines peut améliorer la solubilité et la stabilité tout en réduisant l'immunogénicité.
- **Glycosylation** : L'ajout de groupes glucidiques aux protéines peut améliorer leur solubilité et leur stabilité, ainsi que leurs propriétés pharmacocinétiques.

#### Stratégies de formulation

- **Utilisation d'excipients** : Les excipients tels que les sucres, les polyols, les acides aminés et les tensioactifs peuvent stabiliser les protéines et empêcher leur agrégation, améliorant ainsi la solubilité.
- **Ajustements du pH et de la force ionique** : La modification du pH ou de la force ionique de la solution peut déplacer la charge de la protéine, conduisant à une meilleure solubilité.
- **Co-solvants** : L'ajout de co-solvants comme le glycérol ou le diméthylsulfoxyde (DMSO) peut améliorer la solubilité des protéines en modifiant les propriétés du solvant du milieu.

#### Méthodes physiques

- **Homogénéisation haute pression** : Cette technique utilise une pression élevée pour décomposer les agrégats de protéines et améliorer la solubilité.
- **Ultrasons** : Les ondes ultrasoniques peuvent perturber les agrégats de protéines, entraînant une solubilité accrue.