

## Expression des protéines recombinantes

**Définition :** Une protéine recombinante est une protéine produite par des cellules dont l'ADN a été modifié par recombinaison génétique.

### Expression en système procaryote

Une fois qu'un gène a été cloné, on désire souvent obtenir une protéine à partir d'un fragment d'ADN. La technique utilisée dépendra tout d'abord de l'objectif final.

Par exemple, si on a besoin de la protéine pour faire un anticorps polyclonal, on n'a généralement pas besoin d'avoir une protéine active mais par contre la protéine doit être facilement purifiable. Dans ce cas on s'orientera vers la production en bactérie, en fabriquant des protéines de fusion ou la protéine d'intérêt est fusionnée avec une autre protéine facilement purifiable. Si le but est d'obtenir une protéine pour des études de biochimie ou de biologie cellulaire, la purification n'est pas toujours le problème prédominant par contre l'activité de la protéine (enzyme, récepteur) devient plus importante. Enfin, pour des études de biochimie ou de biophysique, on peut avoir besoin de grande quantité de protéines purifiées comme par exemple pour les études de biochimie structurale.

On va pouvoir exprimer des protéines soit dans des systèmes procaryotes, soit dans des systèmes eucaryotes. Il faudra donc tenir compte des différences entre ces deux systèmes.

Par exemple chez les procaryotes, l'initiation s'effectue par reconnaissance d'une séquence particulière (RBS). Donc, si on veut exprimer un clone provenant d'une cellule eucaryote dans une bactérie, il faudra incorporer cette séquence en amont de l'ATG d'initiation.

### Ce qui est indispensable pour produire une protéine dans E. coli.

Pour exprimer un gène dans E. coli, il doit être inséré dans un vecteur qui contient plusieurs éléments :

- 1) une origine de réplication, un marqueur sélectionnable pour trier et maintenir les bactéries ayant incorporé le vecteur et un polylinker, comme tous les plasmides servant de vecteur,
- 2) un promoteur, si possible contrôlable, dont l'induction produira une grande quantité d'ARNm à partir du gène cloné,
- 3) les séquences responsables de la traduction telle qu'une séquence de liaison au ribosome, ces séquences doivent être bien positionnées.

### Les séquences responsables de la traduction

Chez les procaryotes, l'initiation s'effectue par reconnaissance d'une séquence particulière (RBS, ribosome binding site). Cette séquence est composée d'une séquence riche en purine (Shine-Dalgarno, SD, AGGAGG) et du codon d'initiation qui doit être proche, idéalement, l'extrémité 3' de la boîte de Shine-Dalgarno doit être à 6 bases de l'ATG. Si on veut exprimer un clone provenant d'une cellule eucaryote dans une bactérie, il faudra incorporer cette séquence en amont de l'ATG d'initiation. On peut dans certains cas ajouter en 3' du gène des séquences responsables de l'arrêt de la transcription. Au niveau d'un signal de terminaison, la polymérase libère la matrice et l'ARN monocaténaire néosynthétisé.

La présence d'un codon stop est indispensable pour assurer la terminaison de la traduction. Il en existe trois UAA, UAG et UGA. Le plus souvent on met plusieurs codons stop à la suite à la fin de la séquence codante pour éviter les dérapages et donc les protéines plus longues.

La production d'une protéine eucaryote dans un système bactérien n'est souvent envisageable que

- Pour des protéines dont l'activité n'est pas requise (pour faire un anticorps par exemple),
- Pour les protéines dont les modifications post-traductionnelles ne sont pas importantes. En effet, les procaryotes sont incapables de réaliser la plupart des modifications catalysées par des enzymes,
- Pour les protéines dont le repliement n'est pas dû à des protéines comme c'est le cas pour la plupart des protéines synthétisées dans le réticulum endoplasmique.
- Dans les autres cas on utilisera un système eucaryote

### Les avantages de la fusion traductionnelle

1) la protection de la protéine hétérologue contre la protéolyse : les peptides courts sont rapidement protéolysés dans E. coli ; sous forme de fusion, ils seront stabilisés (cas de peptides hormonaux comme la somatostatine, l'insuline, etc...)

2) la purification de la protéine hétérologue : on peut purifier la protéine de fusion sur une colonne d'affinité portant des anticorps dirigés contre la protéine porteuse. Ces mêmes anticorps serviront à purifier n'importe quelle fusion avec la protéine porteuse.

3) la localisation cellulaire du produit final. Les fusions à la  $\beta$ -galactosidase s'accumuleront dans le cytoplasme. Les taux d'expression sont souvent si élevés que les protéines de fusion peuvent dépasser 50 % des protéines totales et conduire à la précipitation du produit dans le cytoplasme : on parle de corps d'inclusion, visibles en microscopie optique ou électronique sous forme d'une grosse « boule » dans la cellule (les corps d'inclusion posent souvent un problème de renaturation).

Les fusions à la protéine A de *Staphylococcus aureus*, une protéine naturellement sécrétée, sont dirigées sur l'espace périplasmique par le peptide signal de la protéine A : séparées des autres protéines de la cellule (= pré-purifiées), isolées des protéases endocellulaires (= plus stables), ces protéines peuvent aussi être mieux repliées.

### Quelques problèmes rencontrés lors de l'expression hétérologue

Problème rencontré	Solutions possibles
<b>promoteurs non reconnus</b> différences de séquence de promoteurs chez procaryotes / eucaryotes, différents phylums de bactéries	Vecteur comportant un promoteur reconnu par l'hôte d'expression. Clonage de la séquence codante en aval de ce promoteur - une fusion transcriptionnelle
<b>Présence d'introns</b> gène eucaryote, expression chez un procaryote	expression du cDNA
<b>Toxicité</b>	promoteurs inductibles plasmides à nombre de copie contrôlé
<b>Biais de codons</b> Utilisation de codons différents selon les différents organismes	gènes de synthèse (totale ou partielle)
<b>Protéolyse</b>	hôtes protéase-moins protéines de fusion sécrétion du produit
<b>Localisation</b>	signaux de sécrétion protéines de fusion
<b>Absence de modification</b>	Changer d'hôte - levure - cellules animales

### Expression en système eucaryote

#### Les levures

Plusieurs levures sont utilisées pour exprimer les protéines recombinantes. Les plus souvent utilisées sont *Saccharomyces cerevisiae* et *Pichia pastoris*.

#### *Saccharomyces cerevisiae* (levure de boulangerie)

Cette levure présente deux avantages : d'une part, on dispose de vecteurs qui peuvent se maintenir dans les cellules comme des plasmides bactériens, et, d'autre part, on peut facilement intégrer un gène par recombinaison dans le chromosome.

#### Origine de réplication ou ARS

L'origine de réplication de l'épisome 2- $\mu$ m est le plus souvent utilisé,

#### Les marqueurs de sélection

De nombreuses souches de laboratoire sont auxotrophes pour des acides aminés ou des nucléotides, c'est à dire qu'elles ne peuvent pas pousser en l'absence du composé dans le milieu. Les plus utilisés sont les gènes URA3 ou LEU2 responsables de la synthèse de l'uracile ou de la leucine.

#### Le promoteur

On peut utiliser des promoteurs des gènes codant pour l'alcool déshydrogénase (ADH1) ou pour 3-phosphoglycerate kinase (PGK) qui sont exprimés à très haut niveau.

### Sécrétion des protéines

Dans le cas où on désirerait faire sécréter une protéine cytoplasmique, il suffirait de faire une fusion avec un peptide signal d'une protéine de la levure tel que celui de l'invertase.

- *Pichia pastoris* permet l'expression de protéines membranaires qui sont difficiles à obtenir avec *S. cerevisiae*.
- La plupart des modifications post-traductionnelles sont effectuées par *S. cerevisiae*.

### Transformation

On dispose de plusieurs méthodes pour faire rentrer un ADN dans la levure. On peut utiliser soit une transformation chimique avec le l'acétate de lithium, soit une transformation avec du  $\text{CaCl}_2$  sur des sphéroblastes (les sphéroblastes sont obtenus par une lyse de la paroi des levures par une lyticase), soit par électroporation.

### Avantage de la levure

On a quelques millénaires d'expérience et on sait qu'il n'y a pas de toxicité pour l'homme si bien qu'on peut produire dans la levure des protéines destinées à être utilisées comme vaccin. Par exemple, le vaccin contre l'hépatite B a été fait dans la levure.

Avantages de la production de protéines dans *P. pastoris*

- La production de protéine est de 10 à 100 fois plus importante qu'avec *S. cerevisiae*
- Elle permet l'expression de protéines membranaires qui sont difficiles à obtenir avec *S. cerevisiae*

### Clonage et production de l'insuline

L'insuline est constituée de deux chaînes polypeptidiques (51 acides aminés en tout) ; les chaînes A et B 21 et 30 acides aminés, respectivement et reliées entre elles par 2 ponts disulfures et 1 pont disulfure intr chaîne dans la chaîne A. Elle exerce un effet hypoglycémiant.

L'insuline fut cristallisée en 1926.

La première protéine à être complètement séquencée en 1955,

La première à être synthétisée chimiquement en 1958 et

La première protéine humaine produite par biotechnologies en 1979 et commercialisée en 1982.

