

# Production par BioEngineering des biomolécules

L3 Biochimie, Microbiologie

Module assuré par:

Sabri Bousbia, *PhD, MCA*

Centre universitaire Abdelhafid Boussouf, Mila, Algérie

## 2- Semestre 2 :

Unité d'Enseignement	VHS	V.H hebdomadaire				Coeff	Crédits	Mode d'évaluation	
	15 sem	C	TD	TP	TV/Pers			Continu	Examen
<b>Unités d'Enseignement Fondamentales (UEF)</b>									
<b>UEF1 : Biothérapies</b>		<b>9h00</b>							
Matière 1 : Biothérapie génique	67h30	3h	1h30		82h30	3	6	40%	60%
Matière 2 : Biothérapie cellulaire	67h30	3h	1h30	-	82h30	3	6	40%	60%
<b>UEF2 : Dynamique structurale des macromolécules</b>									
Matière 1 : Dynamique structurale des macromolécules	67h30	3h	1h30	-	82h30	3	6	40%	60%
<b>Unités d'Enseignement Méthodologiques</b>									
<b>UEM1(O/P) Bio-Engineering</b>									
Matière 1 : Production par Bio-Engineering des biomolécules	60h	3h	1h	-	65h	3	5	40%	60%
Matière 2 : Applications en Biotechnologies	45h	1h30	1h30		55h	2	4	40%	60%
<b>Unité d'Enseignement Découverte (UED)</b>									
<b>UED1 (O/P) Valorisation des biomolécules</b>									
Matière1 : Valorisation des biomolécules à usage thérapeutique	45h	1h30		1h30	5h00	2	2	40%	60%
<b>Unité d'Enseignement Transversale (UET)</b>									
<b>UET1 (O/P) Législation</b>									
Matière1 : Législation	22h30	1h30	-	-	2h30	1	1		100%
<b>Total Semestre 2</b>	<b>375h</b>	<b>16h30</b>	<b>7h00</b>	<b>1h30</b>	<b>375h</b>	<b>17</b>	<b>30</b>		

## **1. Introduction**

## **2. Définition des protéines recombinantes**

## **3. Production par bioingénierie**

### **3.1. Du gène à la protéine**

3.1.1. Vecteurs d'expression : pET ; pBAD ; pQE ; pGEX

3.1.2. Promoteurs:  $P_{lac}/P_{lacUV5}$  ; tac/trc ; T7 ; arap<sub>BAD</sub> ; Promoteur pL

### **3.2. Les systèmes d'expression**

3.2.1. Système bactérien

3.2.2. Système fongique

3.2.3. Système cellule d'insecte et baculovirus

3.2.4. Système 'cell-free'

3.2.5. Système mammifère

3.2.6. Comparaison des systèmes d'expression

### **3.3. Les bioprocédés**

3.3.1. Procédé upstream

3.3.2. Procédé downstream

### **3.4. Tags d'affinité et purification**

### **3.5. Production des biomolécules dans les organismes transgéniques**

### **3.6. Exemples de protéines biopharmaceutiques produites par bioingénierie**

- Activateur tissulaire de plasminogène

- Facteur VIII

- Insuline

- Hormone de croissance humaine

- Interférons

- Erythropoïétine

- Facteurs de croissance

## **4. Optimisation des protéines par bioingénierie**

4.1. Augmentation de la thermostabilité

4.2. Augmentation de l'activité enzymatique

4.3. Modification des exigences des cofacteurs métallique

4.4. Optimisation de la spécificité des enzymes

# Introduction

- Les biomolécules sont des molécules organiques naturellement présentes dans les organismes vivants.
- Elles constituent les composants structurels et fonctionnels de base des cellules vivantes.
- Les biomolécules peuvent être classées en petites molécules (lipides, phospholipides, glucides, sucres, vitamines et hormones), en monomères (acides aminés, nucléotides, monosaccharides) et en polymères (peptides, oligopeptides, polypeptides, **protéines**, acides nucléiques oligosaccharides, polysaccharides et lignine).
- Certaines des biomolécules importantes jouent un rôle dans le stockage de l'énergie, le métabolisme, les molécules précurseurs, les éléments constitutifs, les messagers et les catalyseurs.
- Les sources importantes de biomolécules précieuses sont obtenues à partir de plantes, de microbes et d'algues.

# **La production de protéines recombinantes**

# La production de protéines recombinantes

- À l'origine, les protéines d'intérêt thérapeutique étaient extraites de sources naturelles telles que le sang, le placenta, ou autres extraits de tissus humains ou animaux. Toutefois, les quantités restaient limitées.
- L'essor de la génomique et des technologies de l'ADN recombinant dans les années 1970 a permis le développement de nouveaux procédés de production de protéines, mettant en oeuvre 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.

# La production de protéines thérapeutiques

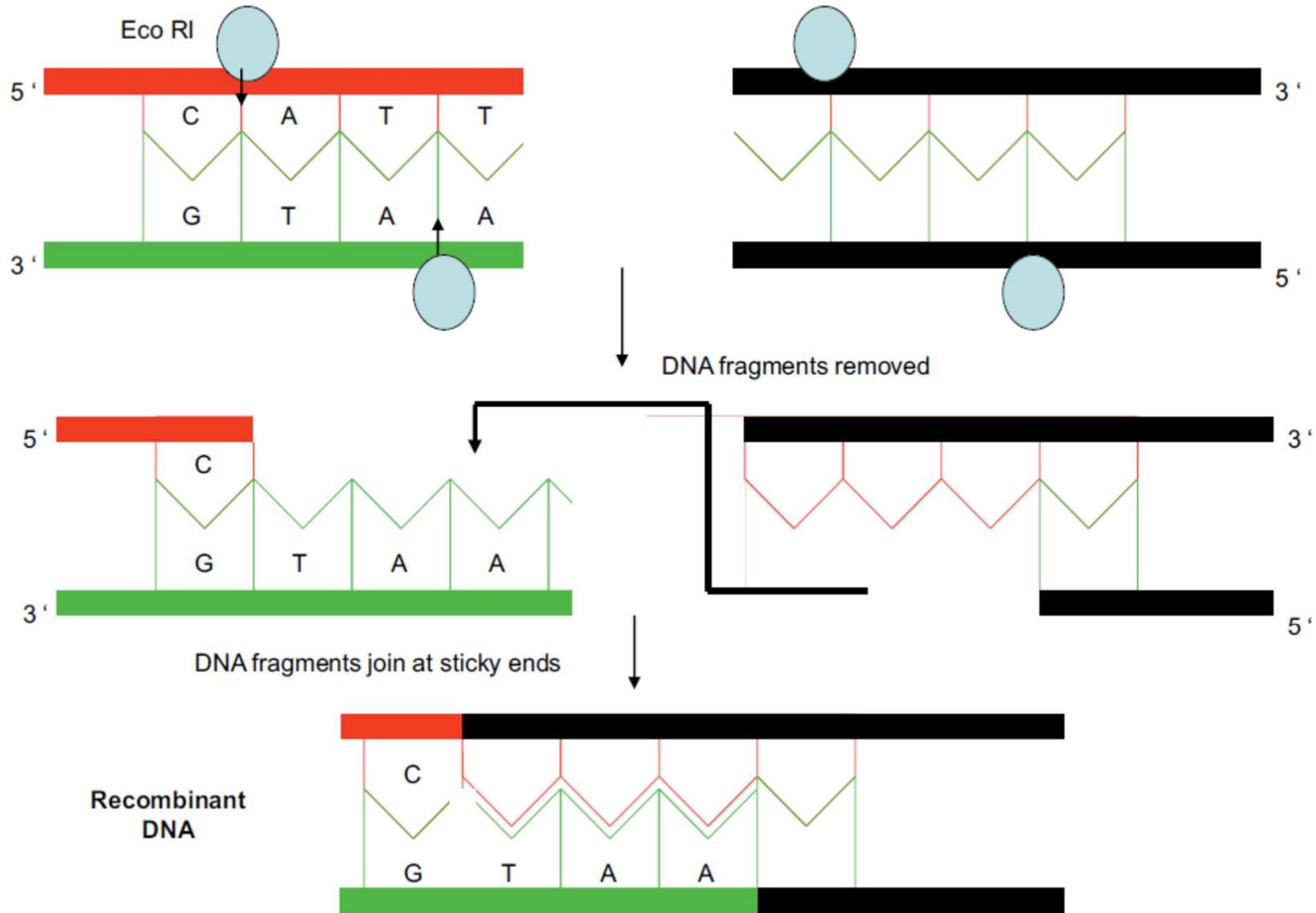
## Exp:

Protéines thérapeutiques	Fonctions	Exemples d'applications
Cytokines	Régulation du système immunitaire	
	Ex : Interférons Interleukines	Hépatite B Cancers, VIH
Enzymes	Catalyse de réactions biochimiques	
	Ex : Enzymes lysosomales Lipase gastrique	Maladies lysosomales Mucoviscidose
Hormones	Régulation des fonctions de l'organisme	
	Ex : Insuline Œstrogènes	Diabète Traitements de substitution hormonale
Facteurs de coagulation	Régulation de la coagulation du sang	
	Ex : Facteur VIII	Hémophilie
Antigènes	Stimulation du système immunitaire pour la production d'anticorps spécifiques	
	Ex : Hémagglutinine Antigène HB	Vaccin contre la grippe Vaccin contre l'hépatite B
	Allergènes	Désensibilisation et traitement des allergies
Anticorps	Régulation du système immunitaire	
	Ex : Anticorps polyclonaux Anticorps monoclonaux	Sérothérapie, prévention des maladies hémolytiques du second nouveau-né Cancérologie, transplantation d'organes, maladies auto-immunes

# Protéines recombinantes

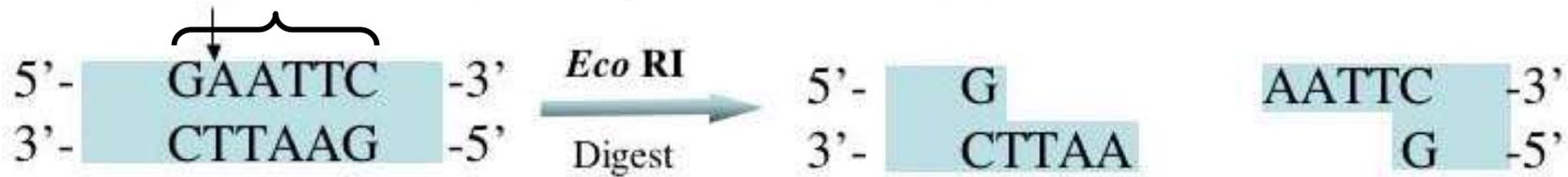
- **Définition**
- Une protéine recombinante (ou hétérologue) est une protéine produite par une cellule dont le matériel génétique a été modifié par recombinaison génétique.
- Un gène codant une protéine d'intérêt est introduit dans le génome de l'espèce productrice (bactéries, cellules mammifères en culture, animaux transgéniques, etc.).
- Les protéines recombinantes peuvent être purifiées et utilisées à des fins thérapeutiques, industrielles ou bien encore dans les activités de recherche.

# ADN Recombinant

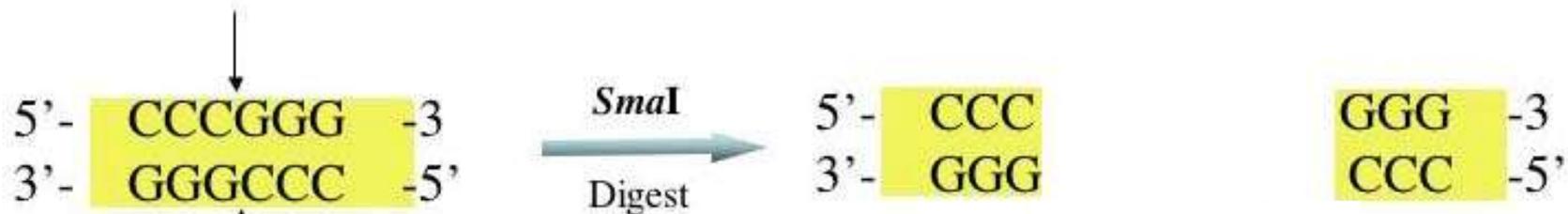


# Enzymes de restriction type II

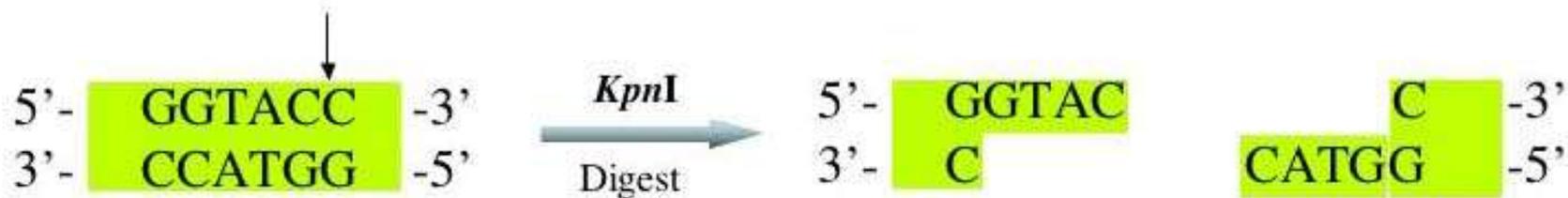
## Site de Restriction



a) Extrémités cohésives (5' Sortante)



b) Extrémités franches



c) Extrémités cohésives (3' Sortante)

# Enzymes de restriction type II

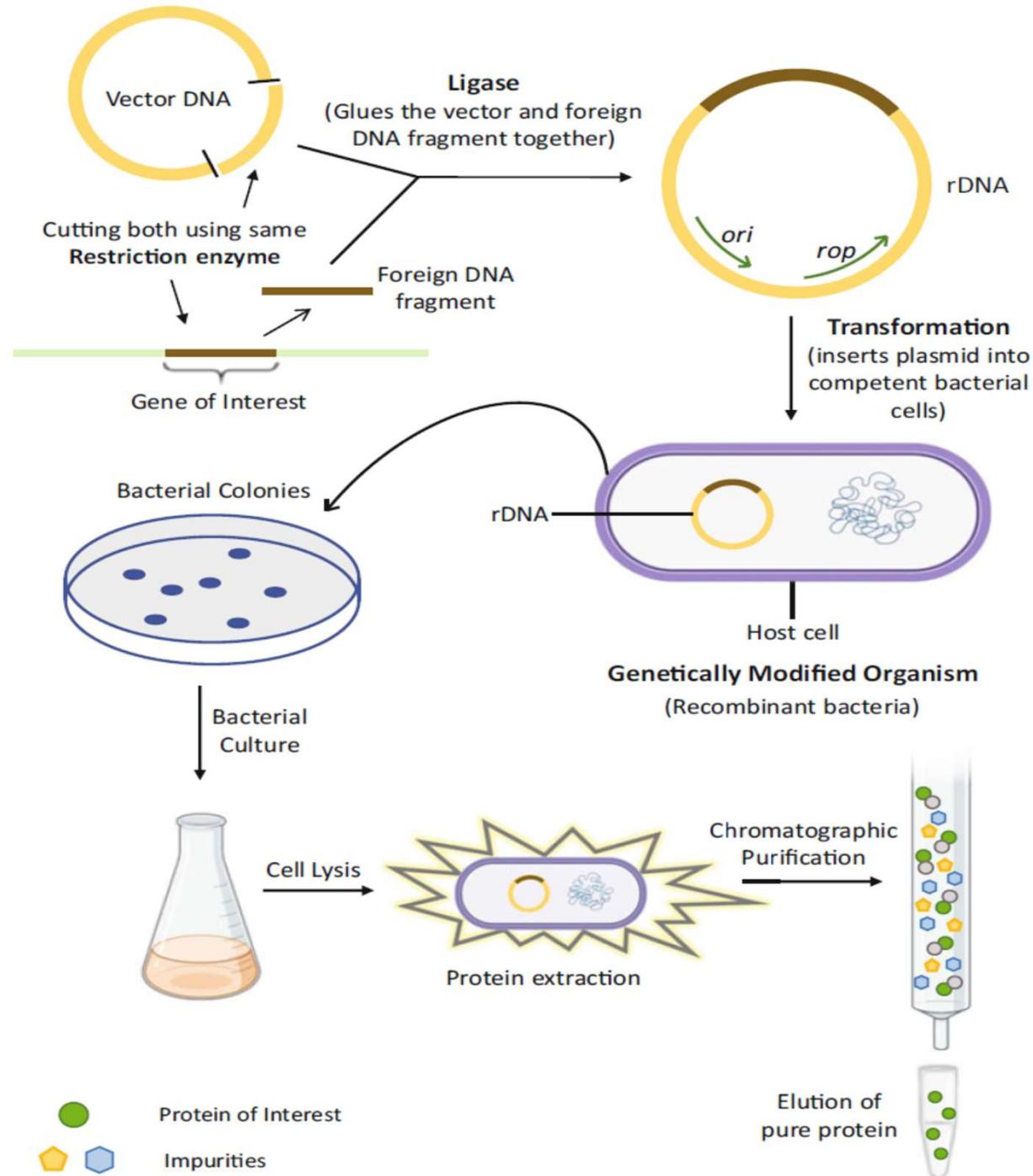
Enzyme	Source	Site de restriction	Coupure		Extrémités (après coupure)
<i>EcoRI</i>	<i>Escherichia coli</i>	5'GAATTC 3'CTTAAG	5'---G	AATTC---3' G---5'	Cohésives
<i>BamHI</i>	<i>Bacillus amyloliquefaciens</i>	5'GGATCC 3'CCTAGG	5'---G	GATCC---3' G---5'	Cohésives
<i>HindIII</i>	<i>Haemophilus influenzae</i>	5'AAGCTT 3'TTCGAA	5'---A	AGCTT---3' A---5'	Cohésives
<i>TaqI</i>	<i>Thermus aquaticus</i>	5'TCGA 3'AGCT	5'---T	CGA---3' T---5'	Cohésives
<i>NotI</i>	<i>Nocardia otitidis</i>	5'GCGGCCGC 3'CGCCGGCG	5'--GC	GGCCGC--3' CG--5'	Cohésives
<i>AluI</i>	<i>Arthrobacter luteus</i>	5'AGCT 3'TCGA	5'---AG	CT---3' GA---5'	Franches
<i>BglIII</i>	<i>Bacillus globigii</i>	5'AGATCT 3'TCTAGA	5'---A G	ATCT---3' A---5'	Cohésives
<i>HaeIII</i>	<i>Haemophilus aegyptius</i>	5'GGCC 3'CCGG	5'---GG	CC---3' GG---5'	Franches
<i>HhaI</i>	<i>Haemophilus haemolyticus</i>	5'GCGC 3'CGCG	5'---GCG	C---3' GCG---5'	Cohésives
<i>PstI</i>	<i>Providencia stuartii</i>	5'CTGCAG 3'GACGTC	5'---CTGCA	G---3' ACGTC---5'	Cohésives
<i>SmaI</i>	<i>Serratia marcescens</i>	5'CCCGGG 3'GGGCCC	5'---CCC	GGG---3' CCC---5'	Franches

# **Production par bioingénierie: Du gène à la protéine**

# La production de protéines thérapeutiques:

- Les protéines recombinantes sont produites à partir d'un transgène introduit dans un organisme hétérologue par génie génétique. Ce procédé biotechnologique se déroule suivant cinq grandes étapes :
- 1. la construction d'un vecteur d'expression (en général, plasmide) jouant le rôle de transporteur du gène codant pour la protéine d'intérêt ;
- 2. la modification d'une cellule hôte ou d'un organisme complexe par le vecteur d'expression ;
- 3. la production de ce nouveau système d'expression, qui exécute les instructions fournies par le transgène, en l'occurrence synthétise la protéine d'intérêt ;
- 4. la séparation de la protéine du milieu de culture, quand elle est sécrétée, ou son extraction du milieu intracellulaire le cas échéant, puis sa purification ;
- 5. la caractérisation de la molécule et la vérification de son degré de pureté.

# Produire une protéine recombinante



# La méthodologie de clonage

- **Couper** le vecteur de clonage avec une enzyme de restriction de choix (par exemple EcoRI)
- **Couper** l'ADN d'intérêt avec la même enzyme de restriction donnant mêmes extrémités cohésives,
- **Mélanger** le vecteur de clonage restreinte et de l'ADN d'intérêt ensemble.
- **Ligaturer** les fragments ensemble en utilisant l'ADN ligase
- **La Transformation: Insérer** de l'ADN ligaturé dans un hôte de choix -
- **Cultiver** des cellules hôtes dans des conditions restrictives,
- **Croître** sur des plaques contenant un antibiotique (**la sélection**)

# Types de vecteurs de clonage

- Différents types de vecteurs de clonage sont utilisées pour différents types d'expériences de clonage.
- La production d'une protéine recombinante est réalisée en utilisant un vecteur d'expression

# Types de vecteurs de clonage

- Types de Vecteurs de clonage
  - Plasmides Bacteriens
  - **Vecteurs d'expression**
  - Bacteriophage
  - Cosmides
  - Vecteurs viraux
  - Bacterial Artificial Chromosomes (BAC)
  - Yeast Artificial Chromosomes (YAC)
  - Human artificial chromosome

# Vecteurs d'expression

- Les vecteurs d'expression : permettent aux expérimentateur de contrôler l'expression des gènes clonés basé sur le contrôle transcriptionnel
- Prévoient des niveaux élevés d'expression de la protéine
- Des promoteurs forts
- Des terminateurs de transcription efficaces sont utilisés pour empêcher l'expression d'autres gènes sur le plasmide

# Produire des protéines recombinantes:

Les différents systèmes d'expression

La production de protéines recombinantes dans:

- **Les systèmes d'expression**

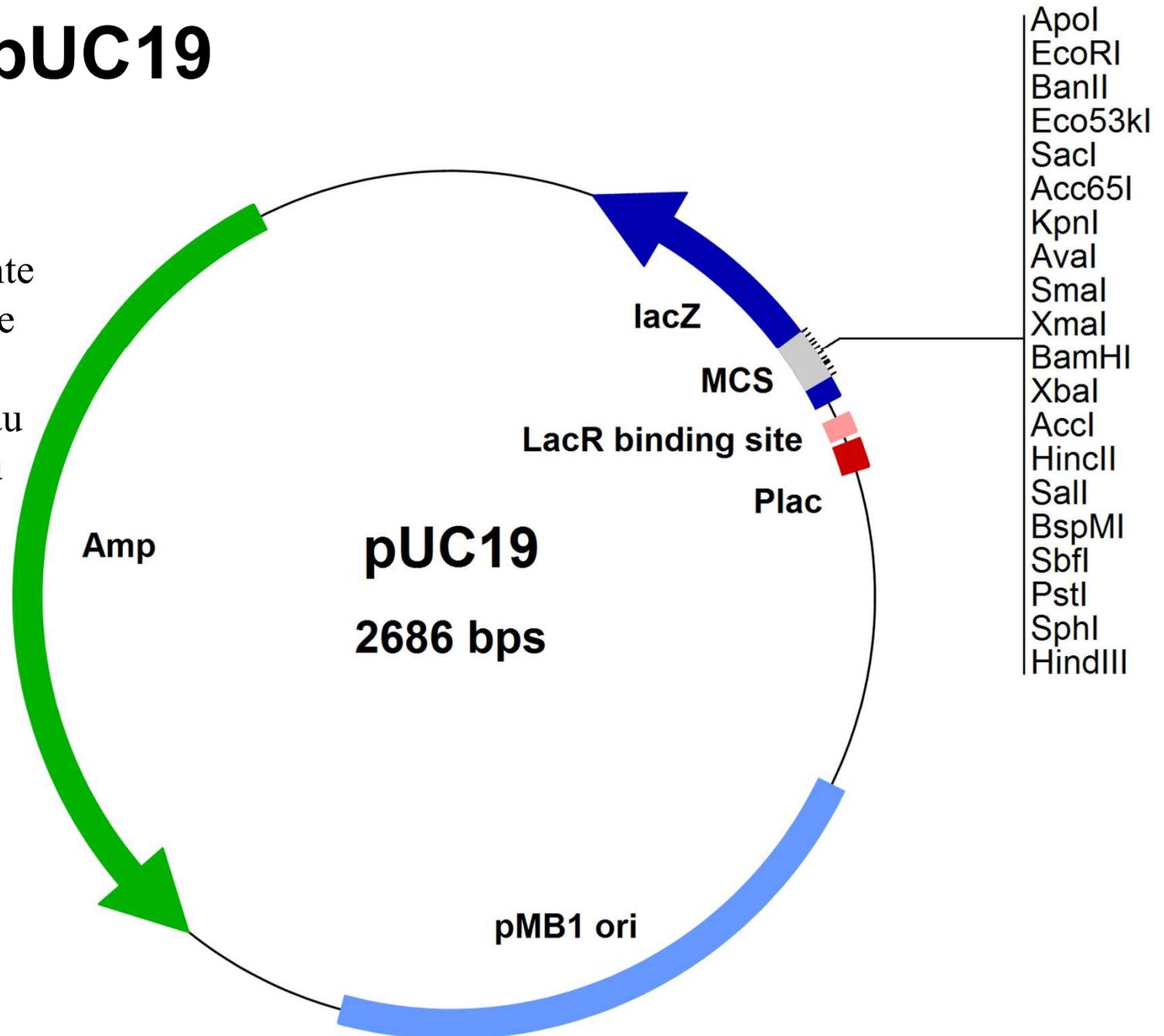
- Système bactérien
- Système fongique
- Système cellule d'insecte et baculovirus
- Système 'cell-free'
- Système mammifère

# Les Vecteurs d'expression plasmidiques

## • Le vecteur pUC19

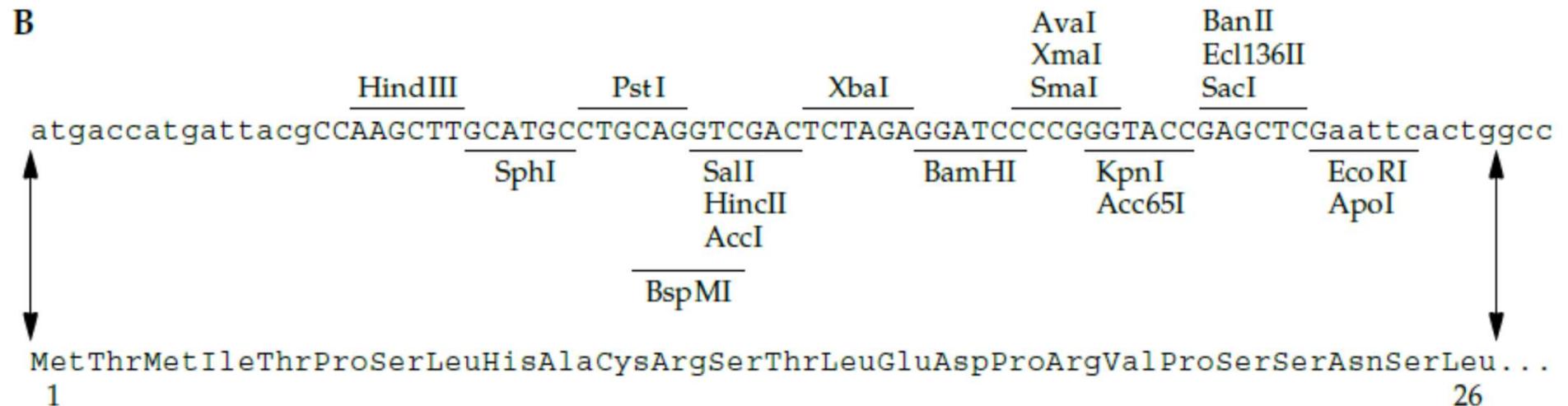
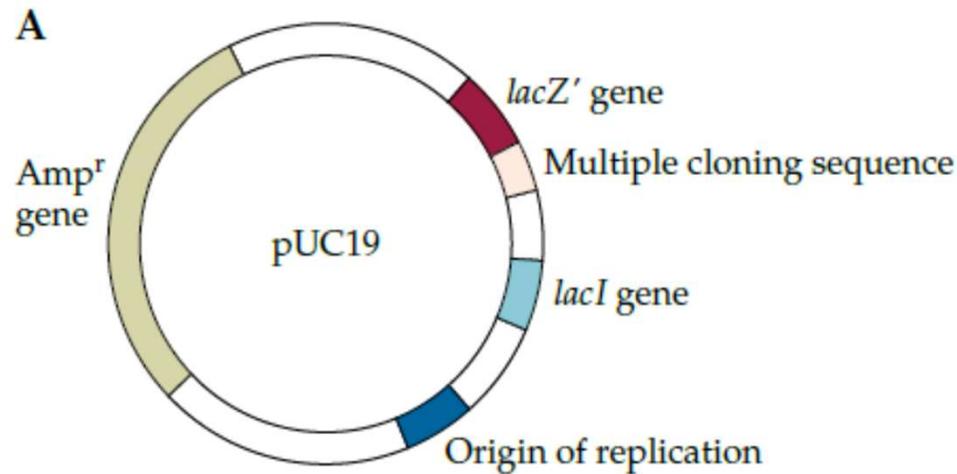
Clonage dans un plasmide (pUC):

- Réplication indépendante de celle du chromosome bactérien (ori)
- Polylinker de clonage au niveau d'une portion du gène de la  $\beta$ -galactosidase *E. coli* (LacZ').
- Sélection des bactéries transformées (plasmide avec ou sans insert): ampR
- Sélection des bactéries transformées (plasmide avec insert): crible blanc/bleu en présence d'IPTG et de X-gal

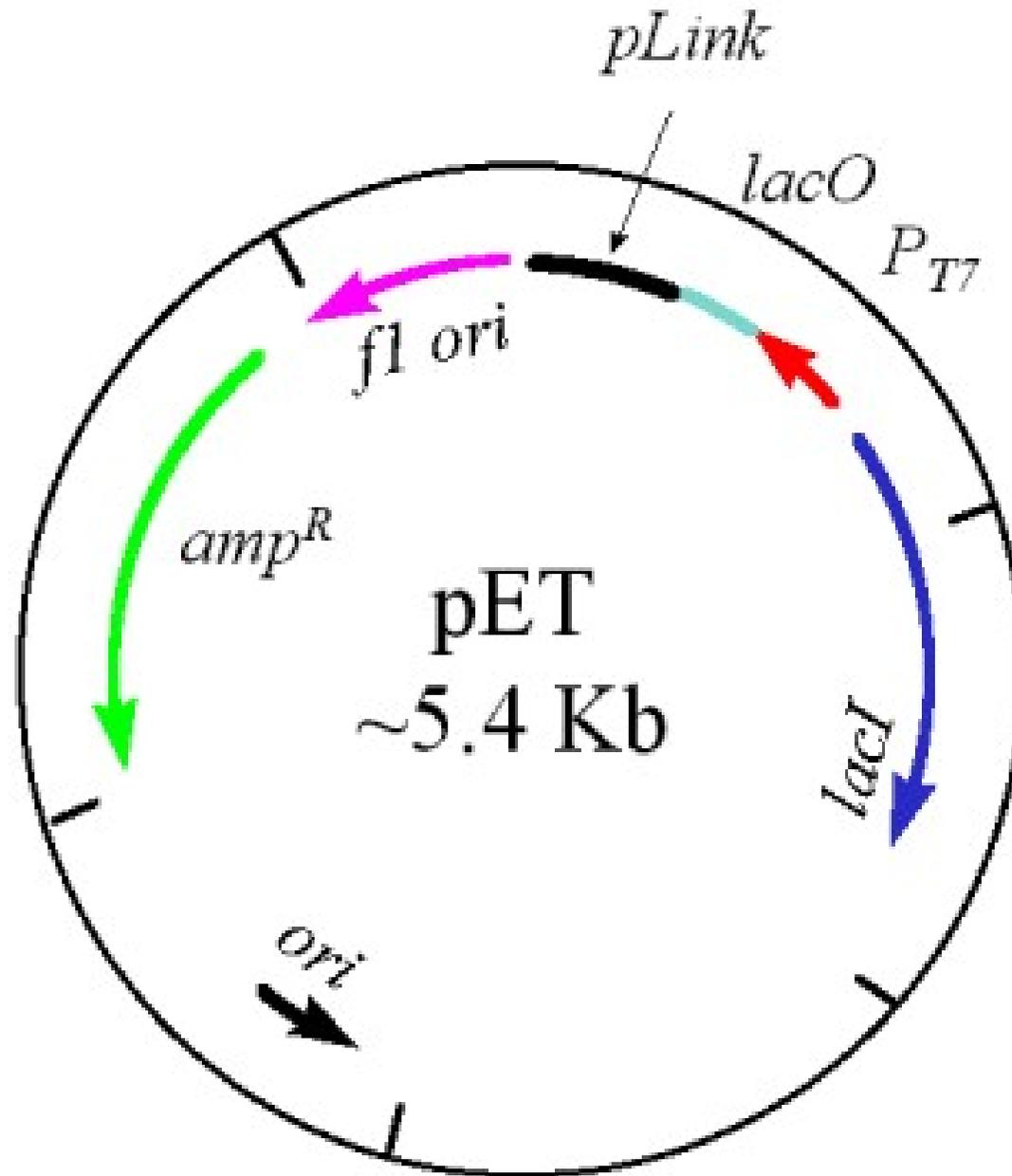


# Les vecteurs de seconde génération

- **Le vecteur pUC19**



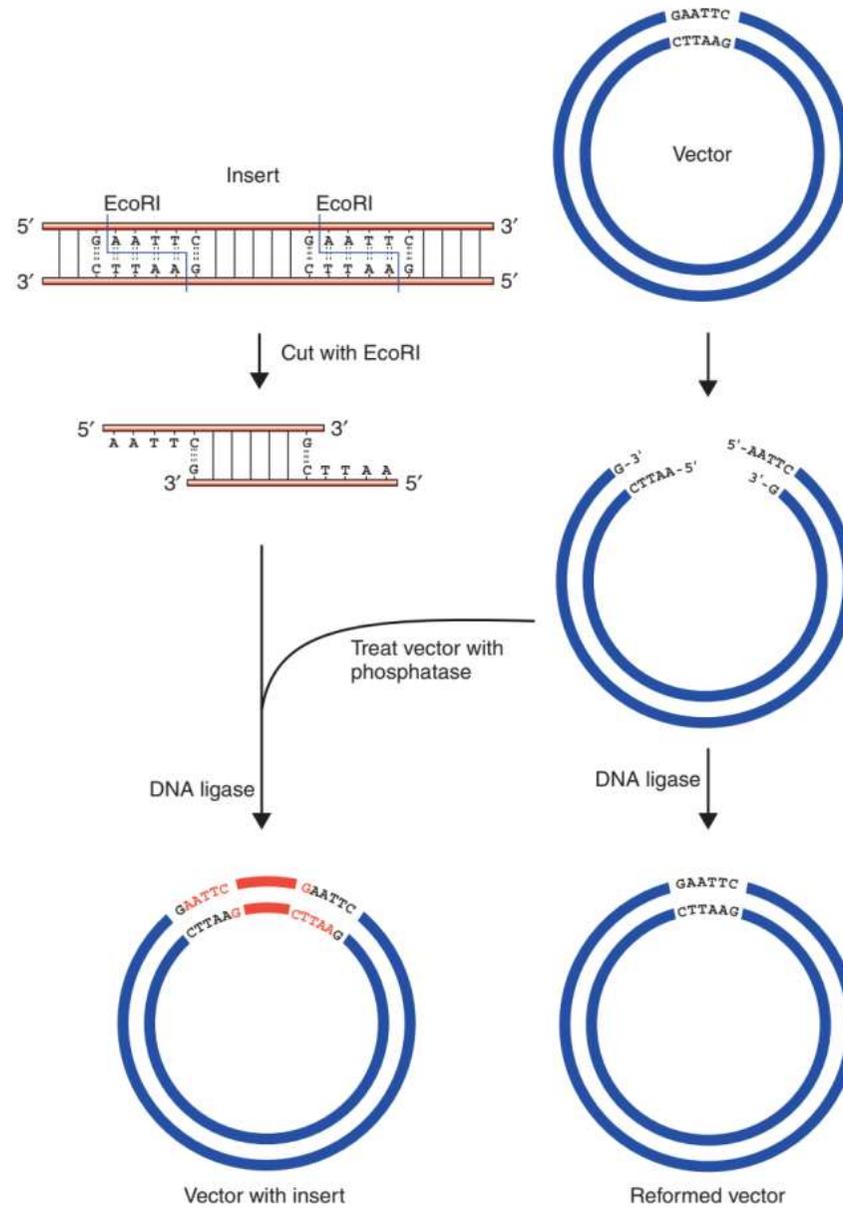
# Vecteurs d'expression



# Clonage du gène d'intérêt

- Si nous souhaitons insérer des séquences d'ADN étrangères dans ce vecteur, nous devons le couper pour produire un ADN linéaire sur lequel nous pouvons attacher d'autres séquences d'ADN à l'aide d'ADN ligase.

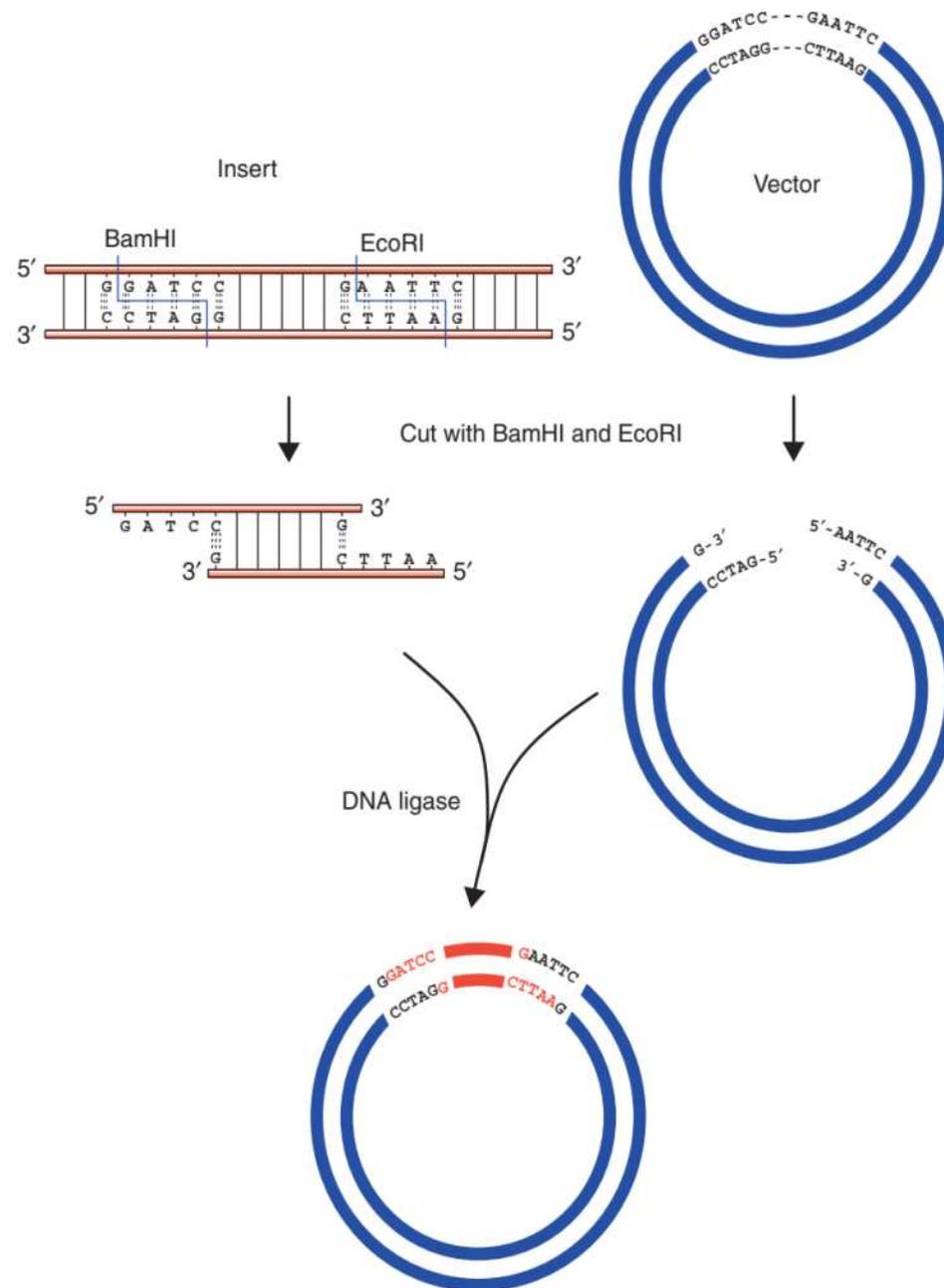
# Clonage du gène d'intérêt



# Clonage du gène d'intérêt

- Il est souvent possible aussi de considérer l'insertion d'ADN dans le vecteur à l'aide de deux enzymes de restriction différentes.
- Couper un tel vecteur par exemple avec *BamHI* et *EcoRI* donnera deux fragments
  - un grand, comprenant la majorité du vecteur,
  - et un petit, représentant l'ADN entre les sites de restriction des enzymes de restriction.
- En présence d'ADN ligase, aucun de ces fragments n'est capable de se ligaturer car les extrémités d'ADN ne sont pas compatibles entre elles.
- La digestion de la séquence d'ADN de l'insert linéaire avec *BamHI* et *EcoRI* donne un fragment qui contient une extrémité compatible *BamHI* et une extrémité compatible *EcoRI*.

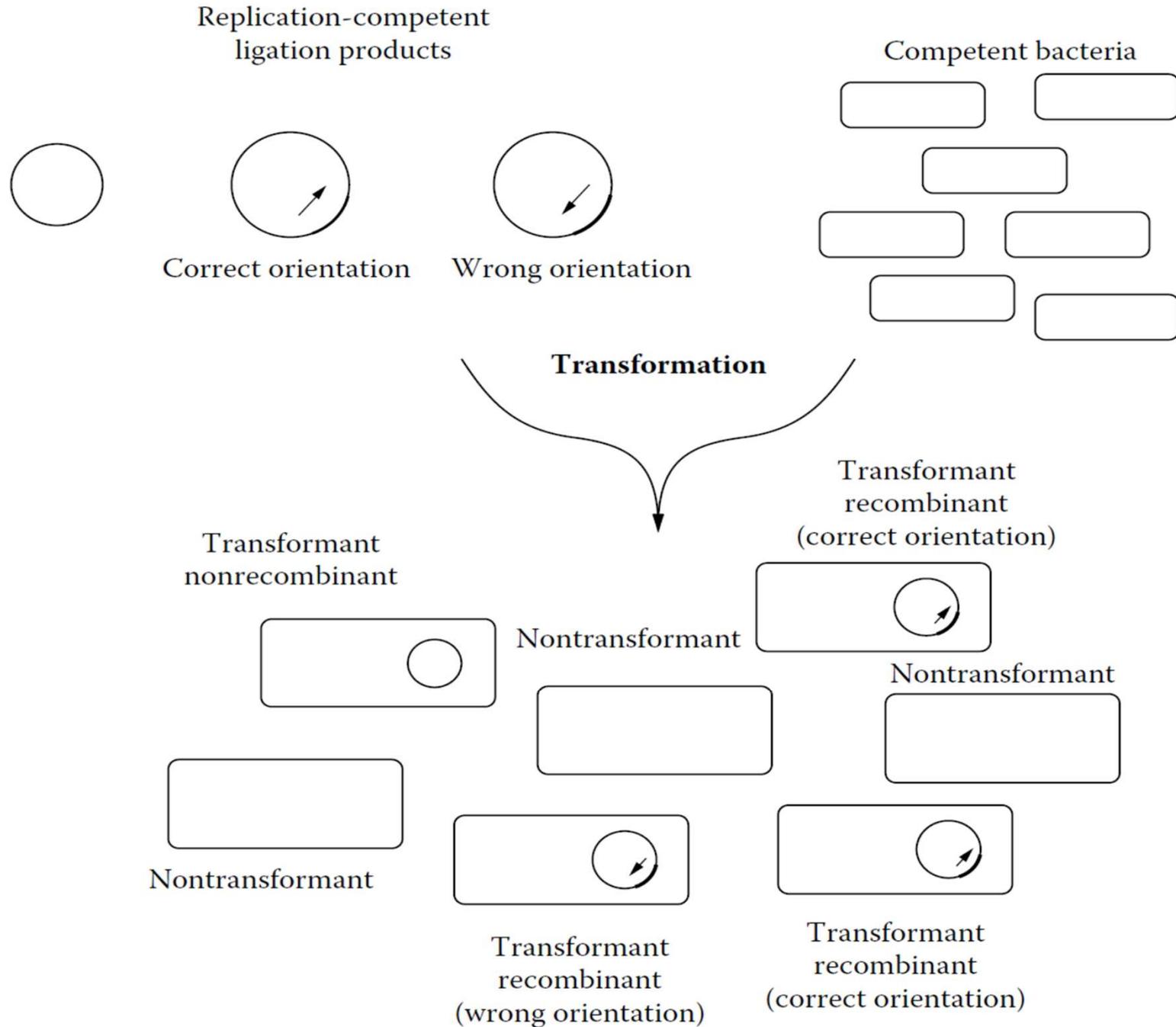
# Clonage du gène d'intérêt



# Clonage du gène d'intérêt

- La transformation est le fait de faire entrer le vecteur dans l'hôte approprié
- Le processus de transformation aboutit à l'insertion d'une molécule d'ADN dans la cellule hôte.
- Tous les vecteurs plasmidiques et bactériophages couramment utilisés pour cloner des fragments d'ADN étrangers permettent l'insertion d'une seule molécule de vecteur dans la cellule hôte.
- Cette molécule unique peut être amplifiée plusieurs fois dans l'hôte, mais toutes les molécules résultantes sont identiques.

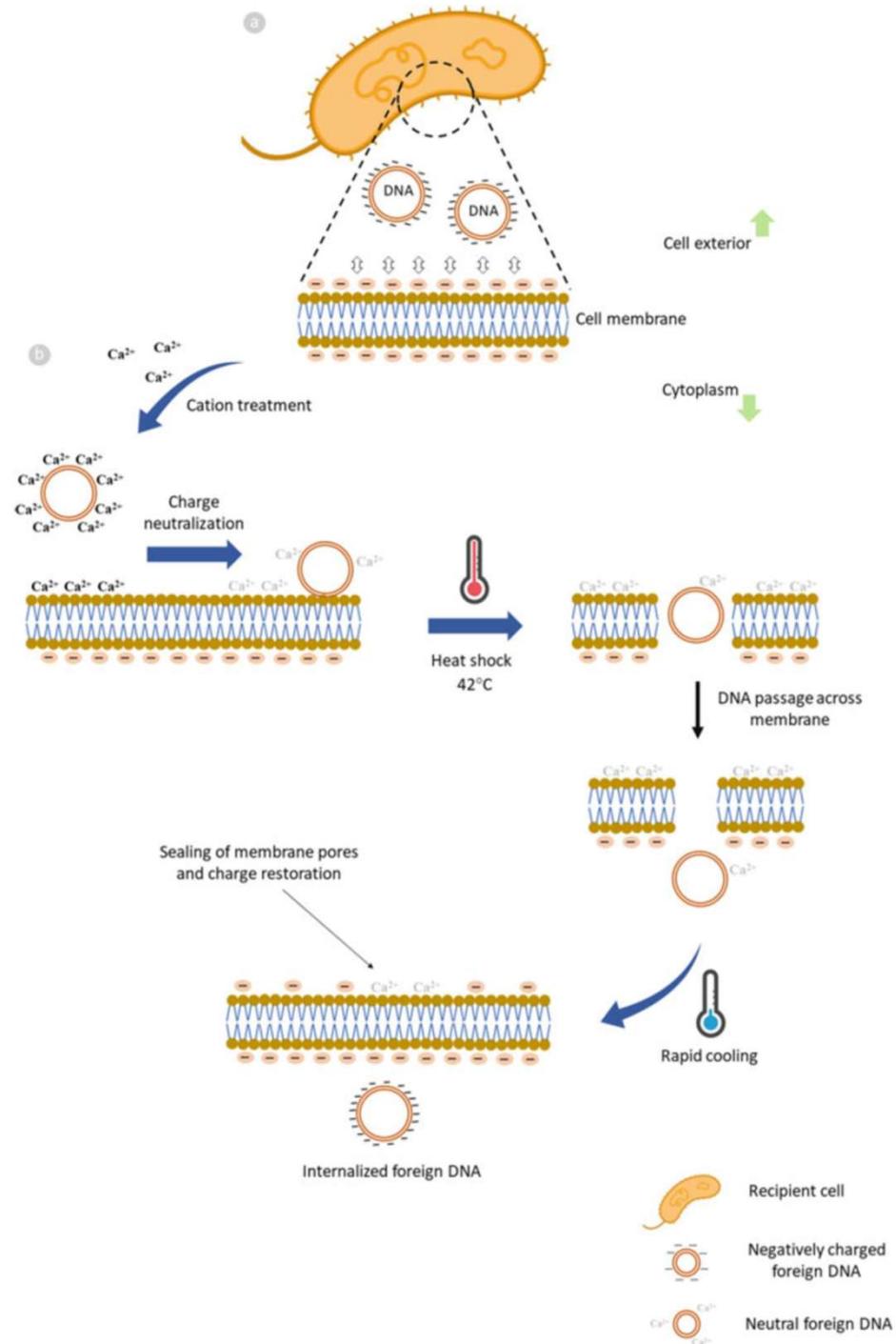
# Transformation bactérienne



# Clonage du gène d'intérêt

- La transformation des cellules compétentes par choc thermique est réalisée en mélangeant l'ADN plasmidique avec les cellules, en incubant pendant 2 min à 42°C est souvent utilisé afin d'élargir les nanopores créés dans les membranes puis le mélange est directement transféré sur glace pendant 90 secondes et ainsi stimuler l'entrée du plasmide dans les cellules.
- Les cellules transformées sont généralement incubées dans un bouillon nutritif à 37°C pendant 60 à 90

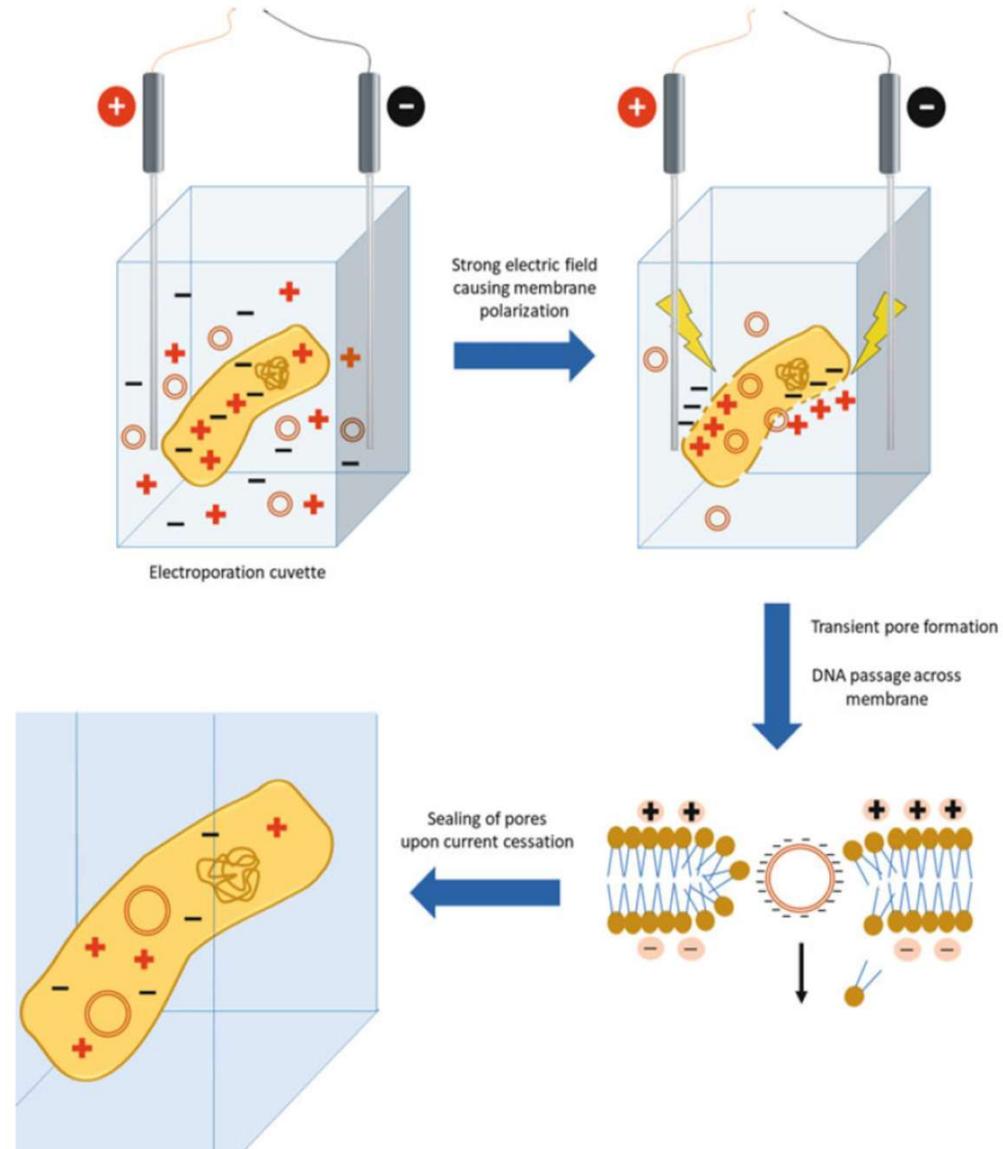
# Transformation bactérienne



# Clonage du gène d'intérêt

- Des efficacités de transformation accrues ont été observées en utilisant des impulsions électriques à haute tension dans un processus appelé électroporation.
- La transformation par choc électrique (ou l'électroporation) est l'utilisation d'une impulsion de champ électrique pour induire des pores microscopiques dans une membrane biologique

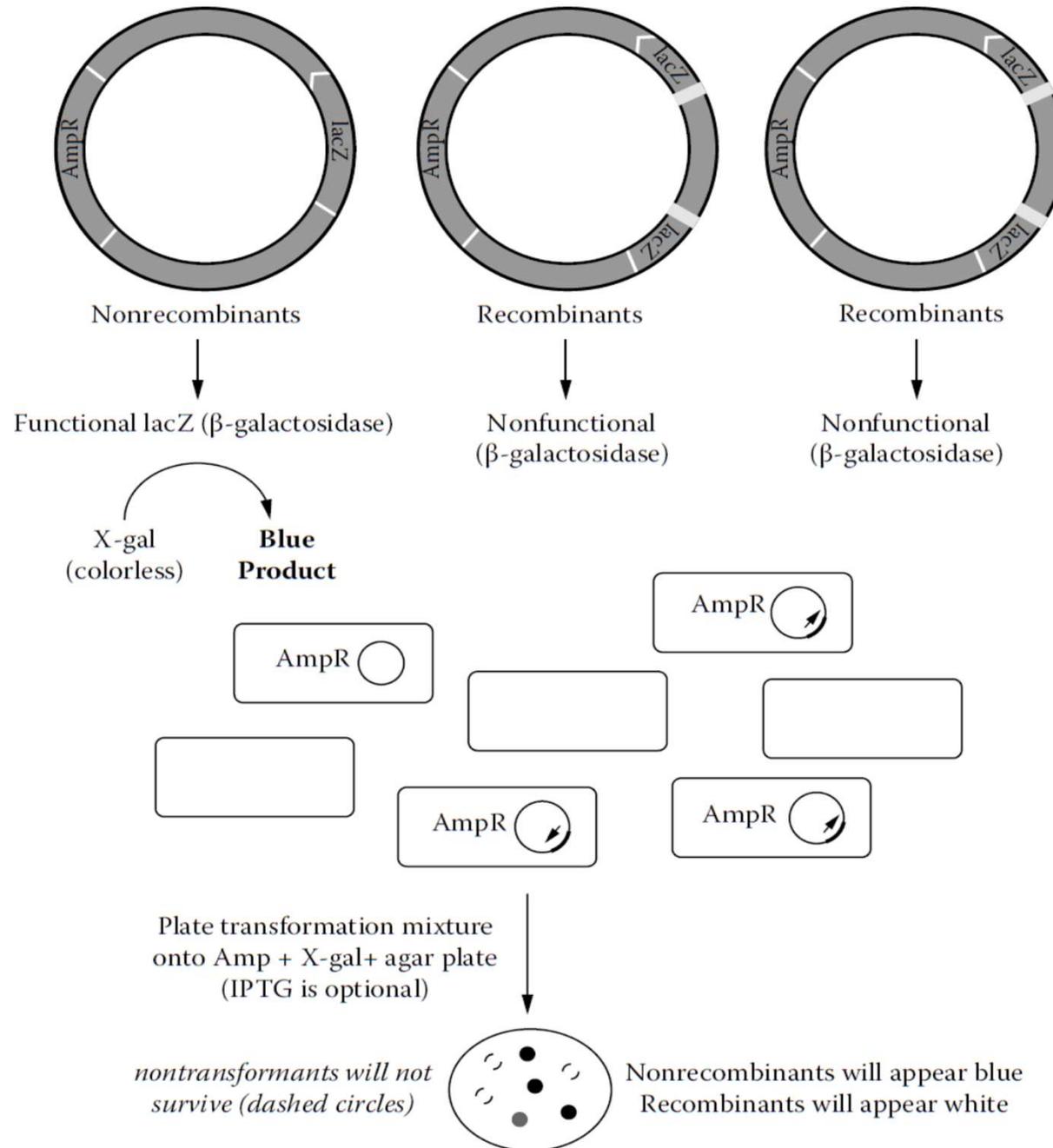
# Transformation bactérienne

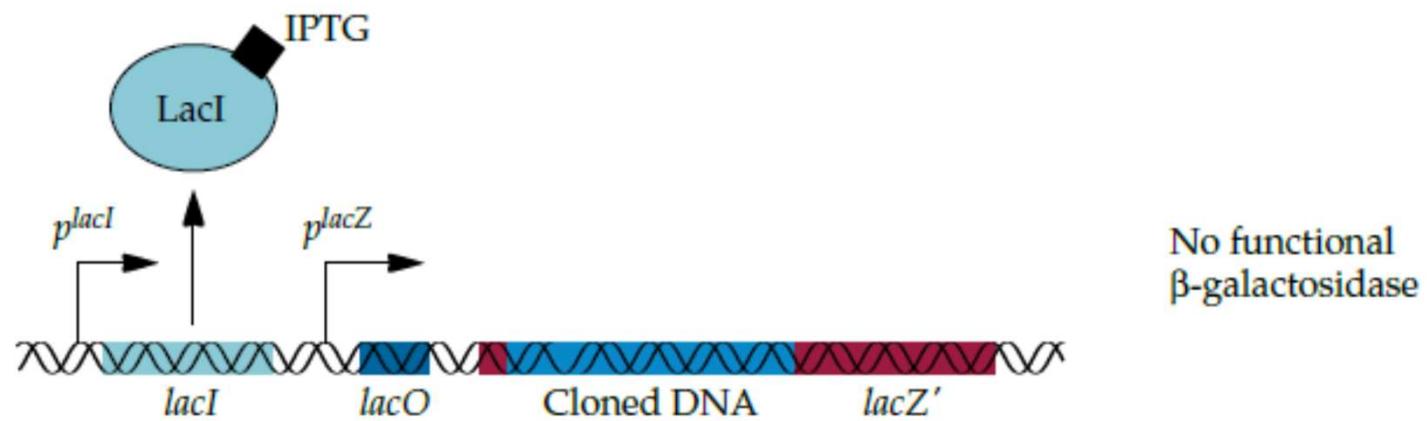
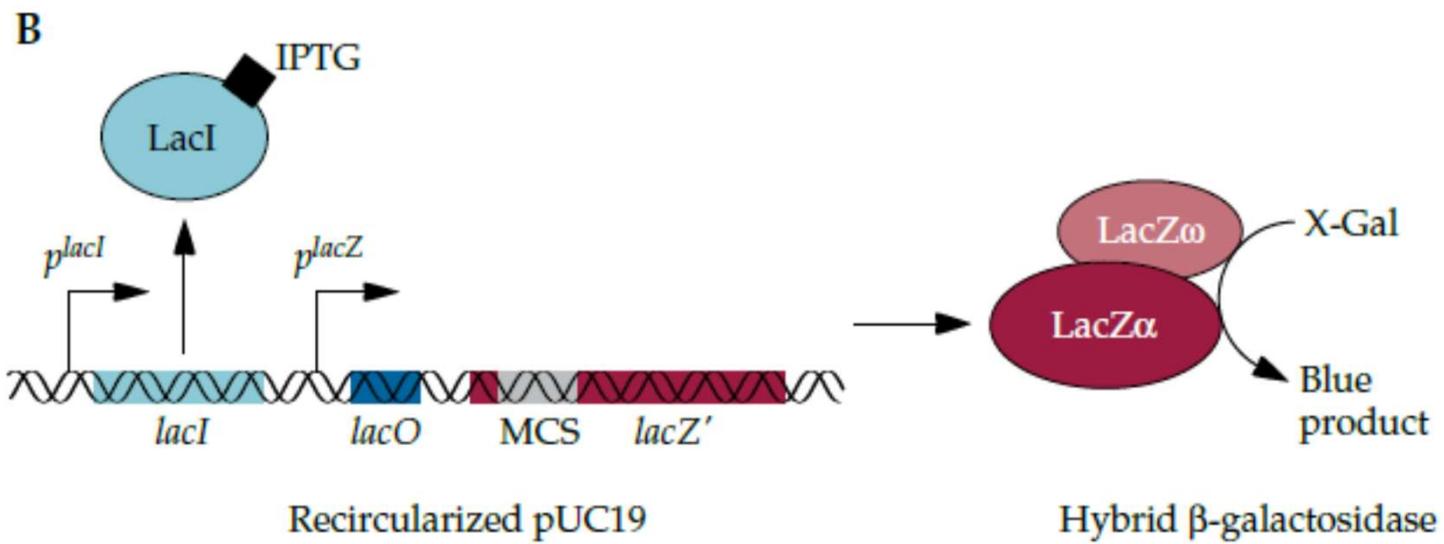


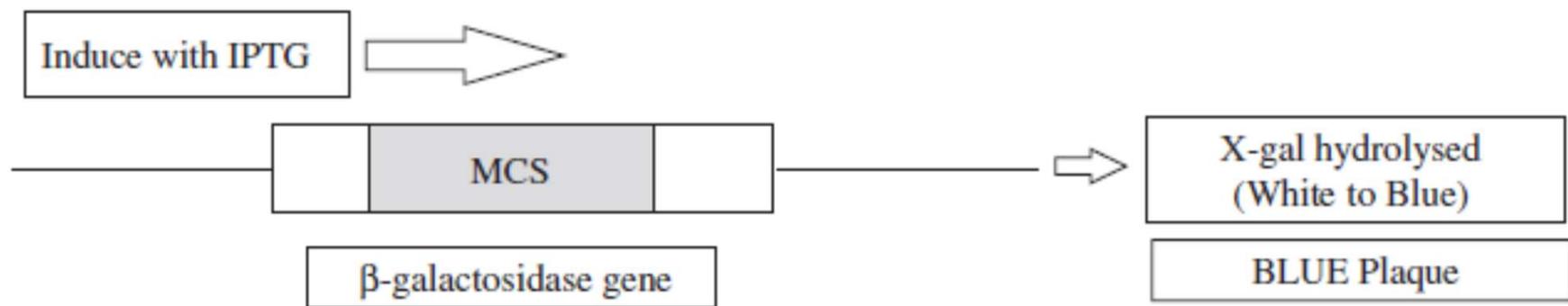
# Clonage du gène d'intérêt

- **Sélection**
- Il existe de nombreuses façons différentes de cribler (sélectionner) les clones recombinants, selon le vecteur utilisé pour le clonage. La méthode la plus couramment utilisée pour le criblage des transformants, par exemple, est la sélection par des antibiotiques, bien que d'autres méthodes telles que le criblage bleu-blanc

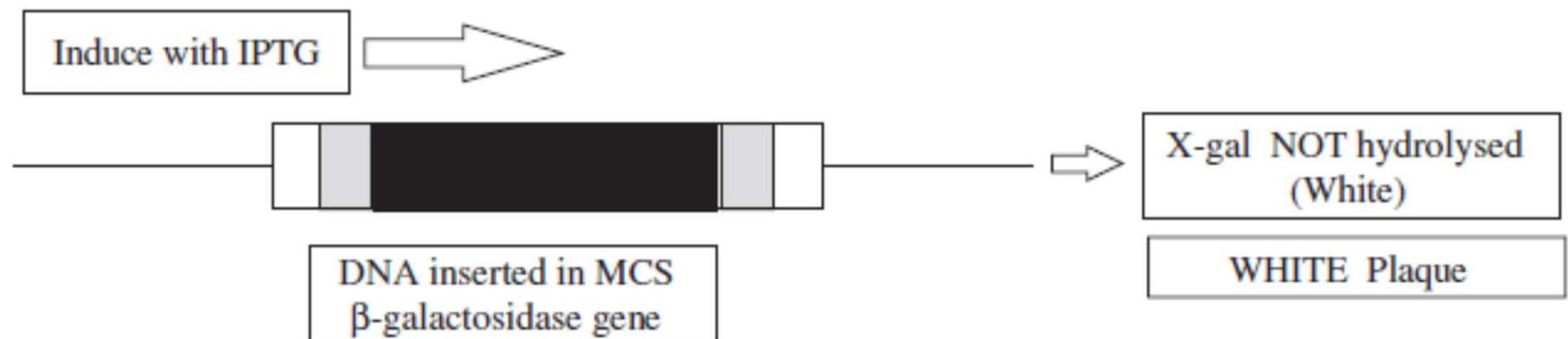
# Sélection des clones recombinants





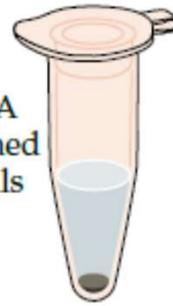


Recombinant vector (insert within MCS)

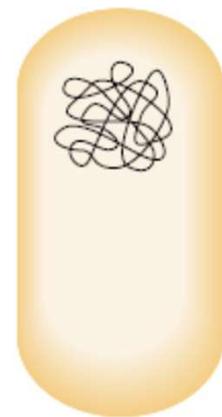
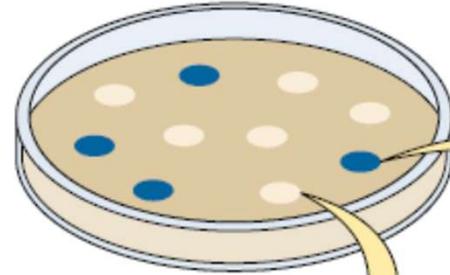


A

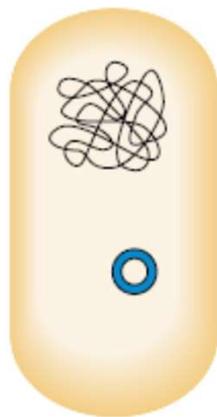
Mixture of pUC19-target DNA ligation products is transformed into competent *E. coli* host cells



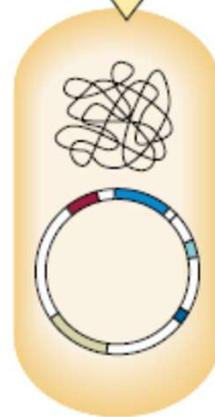
Transformation mixture is plated on medium containing ampicillin, X-Gal, and IPTG



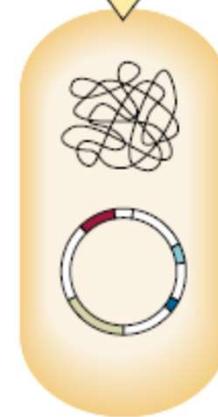
No plasmid  
Amp<sup>S</sup>, no growth  
on ampicillin



Circularized  
target DNA  
Amp<sup>S</sup>, no growth  
on ampicillin



pUC19 with cloned  
target DNA  
Amp<sup>r</sup>, no  $\beta$ -galacto-  
sidase activity



Recircularized  
pUC19  
Amp<sup>r</sup>,  $\beta$ -galacto-  
sidase activity



La croissance sur la plaque: bleu = ADN pas inséré et blanc = ADN inséré.  
Certains% de vecteurs digérés sera recircularisé sans insert. Retirez 5' phosphates avec la phosphatase alcaline pour empêcher la recircularisation (ce qui élimine aussi des plaques bleues).

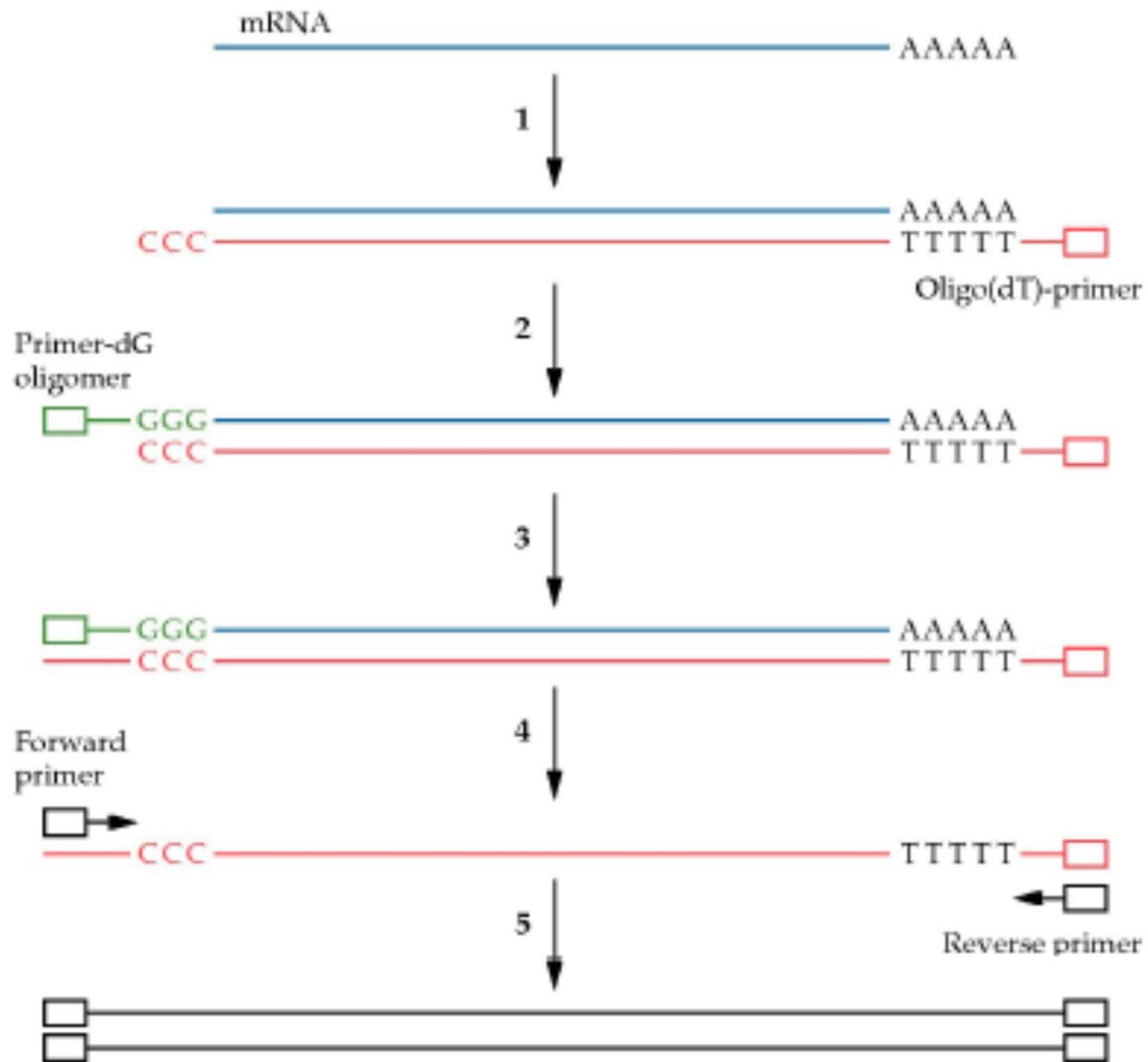
# Cloner à partir de l'ARNm

- En théorie, l'ADN de n'importe quel organisme peut être cloné. L'ADN cible peut être obtenu directement à partir de l'ADN génomique, dérivé de l'acide ribonucléique messenger (ARNm), sous-cloné à partir d'ADN précédemment cloné ou synthétisé in vitro.
- L'ARNm est transformé en ADN double brin avant clonage

# Expression des gènes de mammifères chez les bactéries

- Ceci peut être accompli en utilisant une transcriptase inverse pour synthétiser des ADNc à partir de l'ARNm mature codant pour la protéine d'intérêt

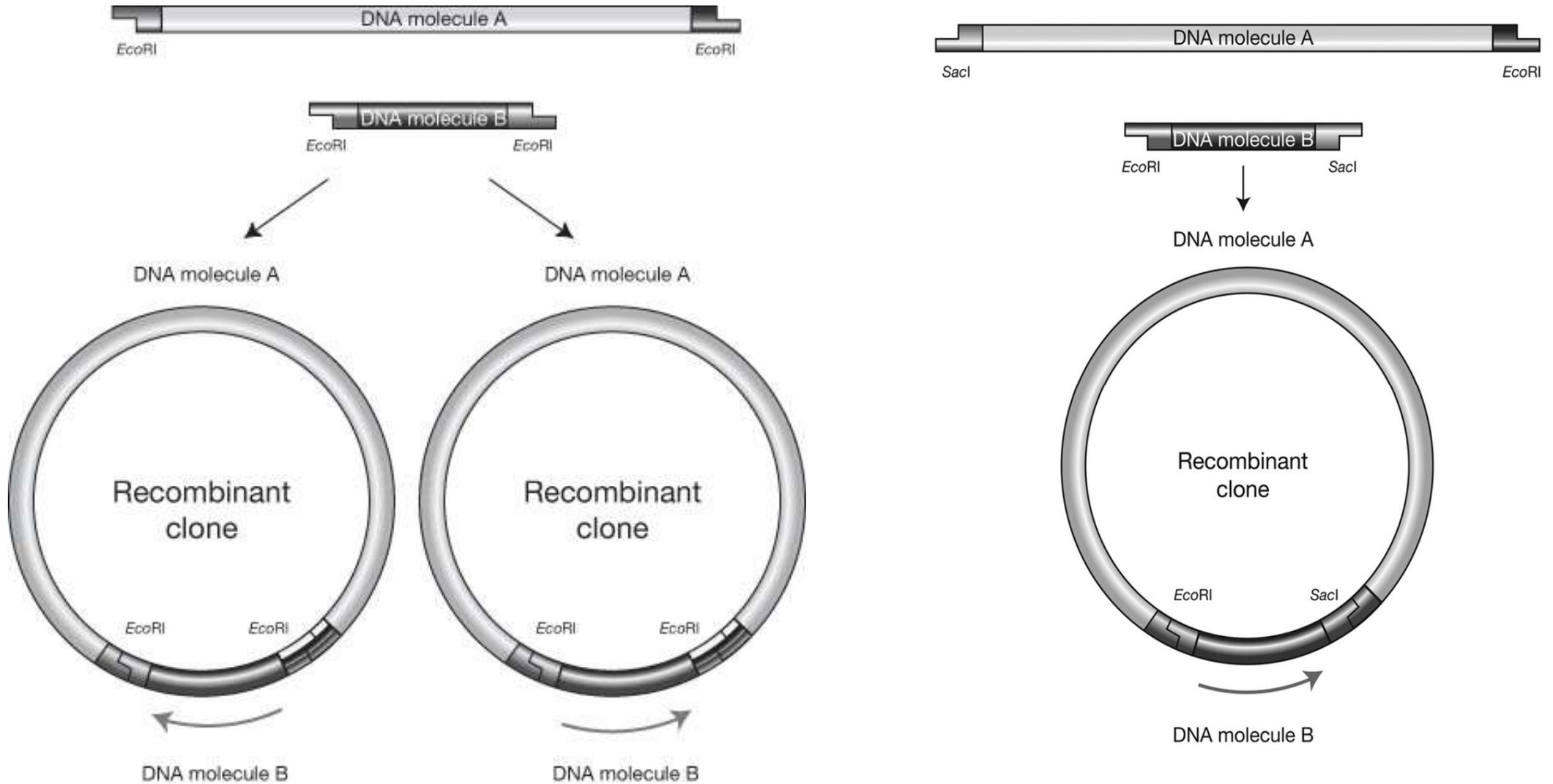


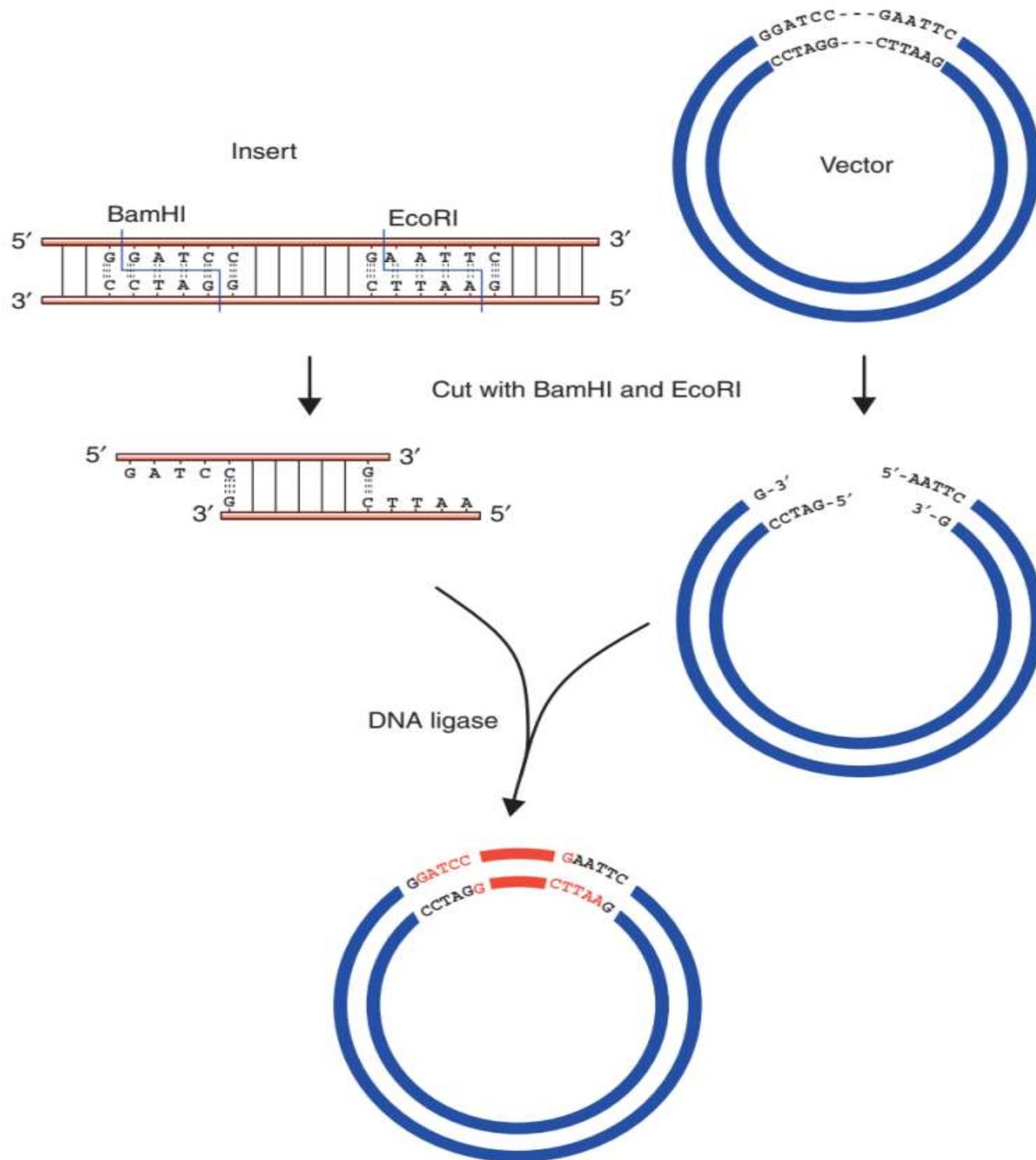


# Direction du clonage

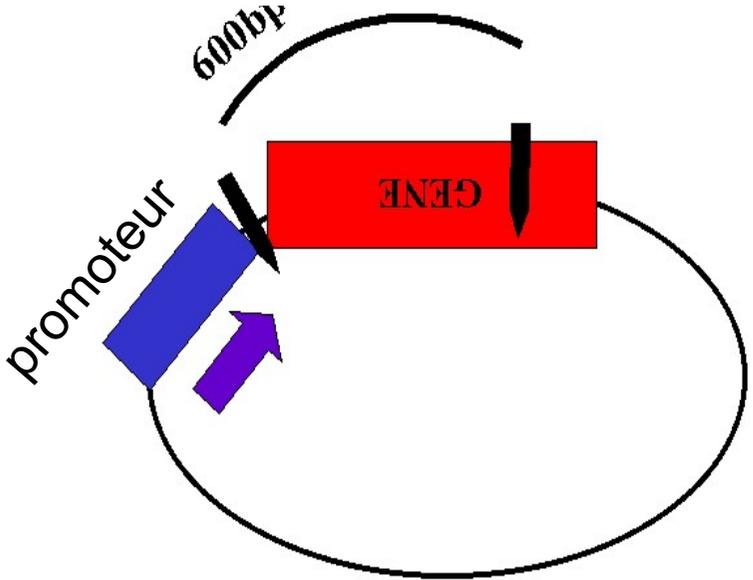
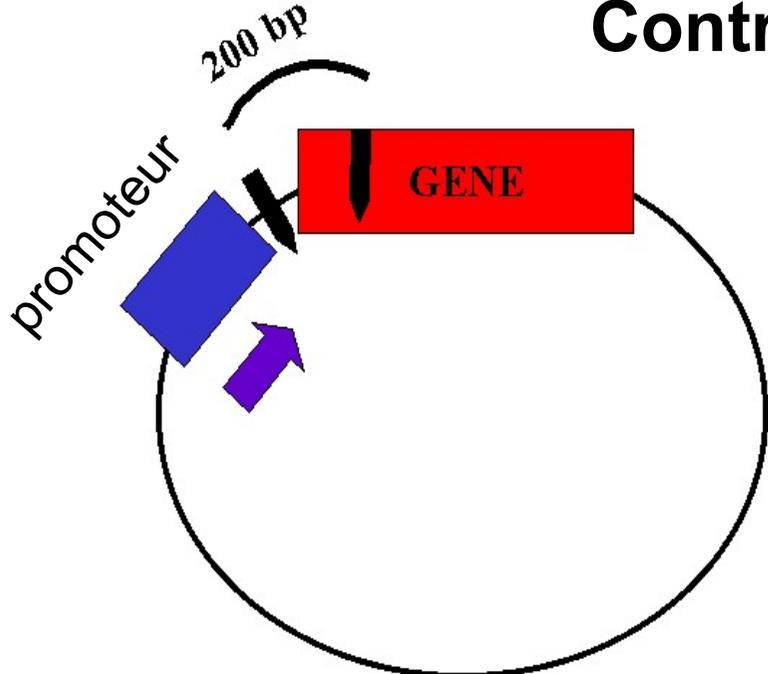
- Souvent, on souhaite insérer un ADN étranger dans une orientation particulière
- Cela peut être fait en faisant deux clivages avec deux enzymes de restriction différentes
- Coupure de l'ADN étranger par les mêmes deux enzymes de restriction
- ADN étranger ne peut être insérée que dans une direction

# Clivage avec deux enzymes de restriction différentes





# Si clivage par une seule enzyme de restriction: Contrôle de la bonne orientation



Électrophorèse

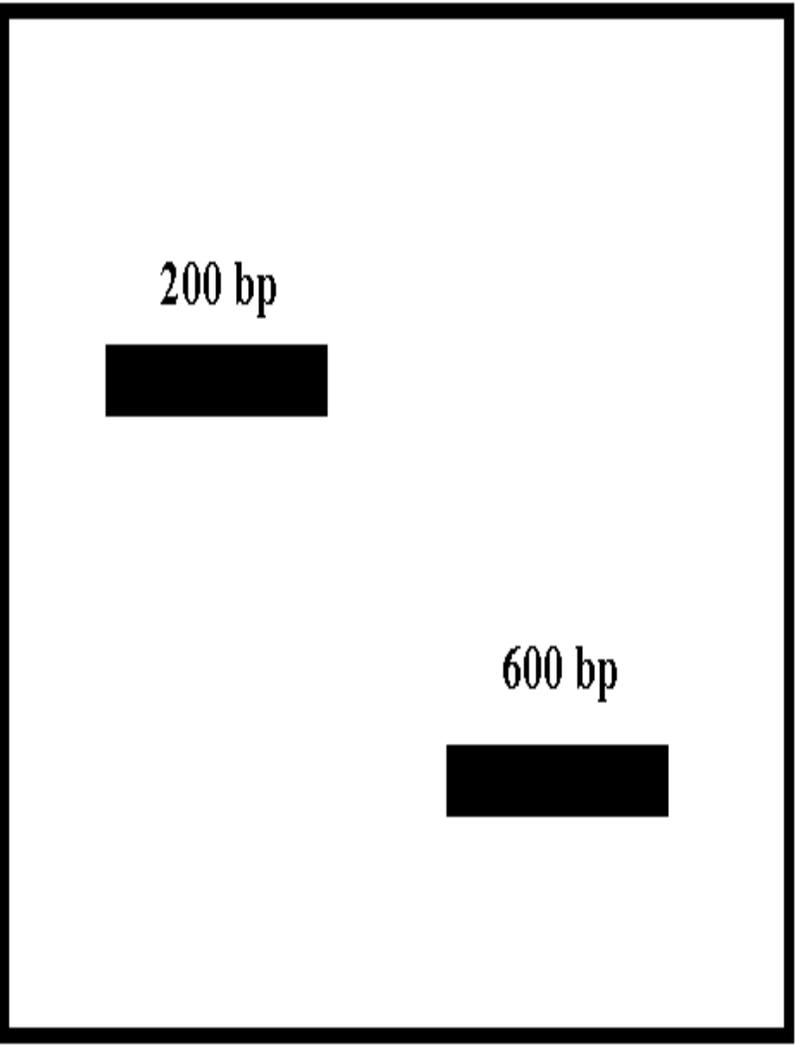
+

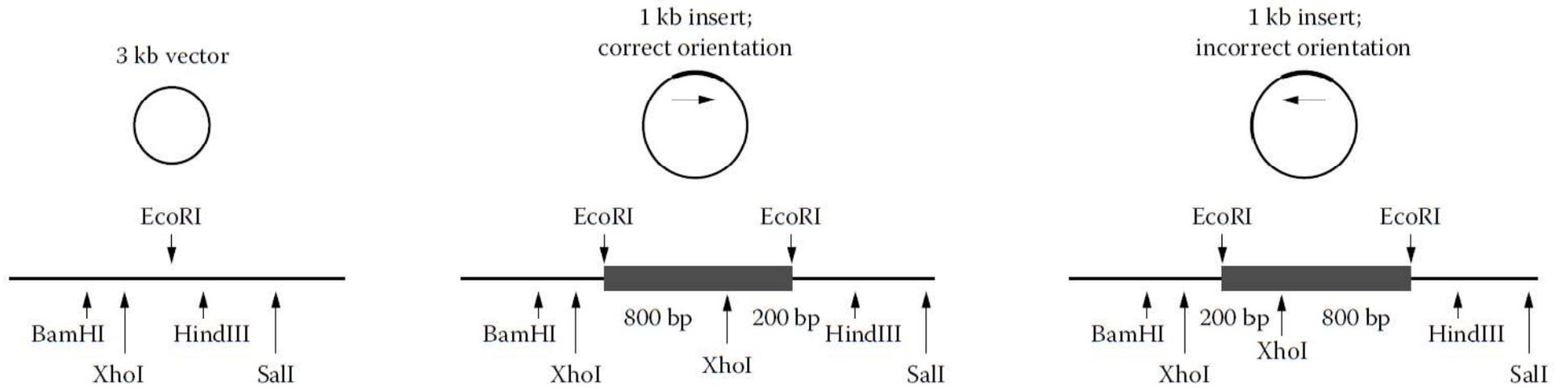


-

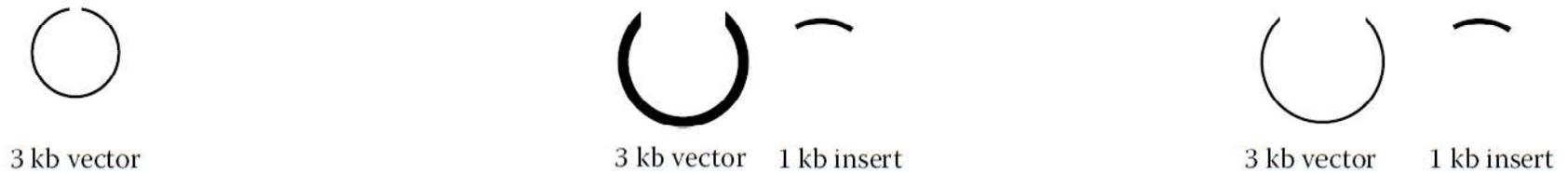
Direction  
correcte

Direction  
incorrecte

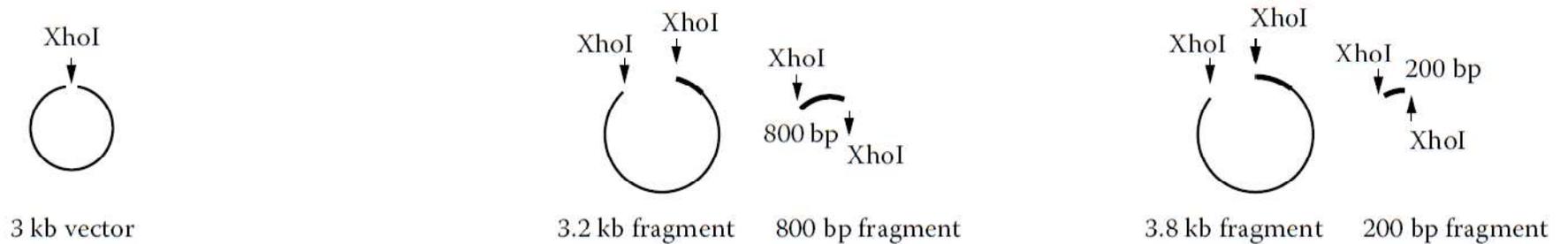




EcoRI digestion of plasmid minipreps from transformant colonies:



XhoI digestion of plasmid minipreps from transformant colonies:



# Production Protéines Recombinantes

Stratégie fondée sur la recombinaison génétique Utilisation d'un couple : Vecteur / Hôte

## → Vecteur d'expression

- Introduire le gène d'intérêt dans la cellule hôte
- Permettre l'expression du gène (*plasmides, virus*)

## → Cellules hôtes

- Hébergement du gène d'intérêt
- Expression du gène → Synthèse de la protéine recombinante (*bactéries, levures, cellules isolées, animaux et plantes transgéniques*)

## → Avantages / Inconvénients

# Utilisations des Systèmes d'expression

## Quelques questions préalables

- Quels types de protéines ? **Solubles / Membranaires**
- Quelles quantités ?  **$\mu\text{g}$ , mg, 100<sup>aine</sup> mg par litre**
- Quels usages ? **Quels niveaux d'exigence/ qualité ?**
- Dispose-t-on de tests fonctionnels ? **Protéines / Enzymes**

## De quelles « garanties » dispose-t-on ?

- Expressions : **Homologues / Hétérologues**
- Conservation de l'usage des codons
- ADN  $\rightarrow$  Epissage des introns  $\rightarrow$  ARNm  $\rightarrow$  Protéine
- Conservation de la fonction
- Importance des modifications post-traductionnelles  
*Glycosylation, Phosphorylation, Ubiquitinylation, S-S, Acétylation, Amidation, Myristylation (C9), Sulfatation,*
- Ancrage membranaire ...
- Assemblage des sous-unités : Homo ou Hétéro-mères

# Les facteurs à considérer lors d'une surexpression

## 1 Vecteurs d'expression : Choix des promoteurs

- Inductibles / Constitutifs

## 2 Cellules hôtes :

- Durée des générations (rendement de biomasse → protéine)
- Culture à haute densité (possibilité d'utilisation de fermenteurs)
- Facilité de manipulation (fragilité)

## 3 Localisation de l'adressage de la protéine

- Cytoplasme / Périplasme / Milieu extérieur (*Sécrétion*)

## 4 Extraction / Purification

- Choix des plasmides (*Protéine Fusion + Etiquette = « Flag »*)
- Corps d'inclusion (*extraction / solubilisation de la protéine*)
- Sécrétion

## 5 Qualité de la protéine

- Stabilité après la synthèse / protéolyse (*via les protéases*)
- Similitude (Identité) structurale avec la protéine « naturelle »
- Similitude de fonction / d'activité

# Vecteurs plasmidiques d'expression

- Après le clonage, le gène étranger (insert) doit être exprimé pour la production de protéines et doit donc être cloné dans un vecteur plasmidique d'expression.
  - Un vecteur d'expression est défini comme un plasmide ou un virus conçu pour délivrer et exprimer un gène d'intérêt à l'intérieur d'un organisme hôte

# Vecteurs plasmidiques d'expression

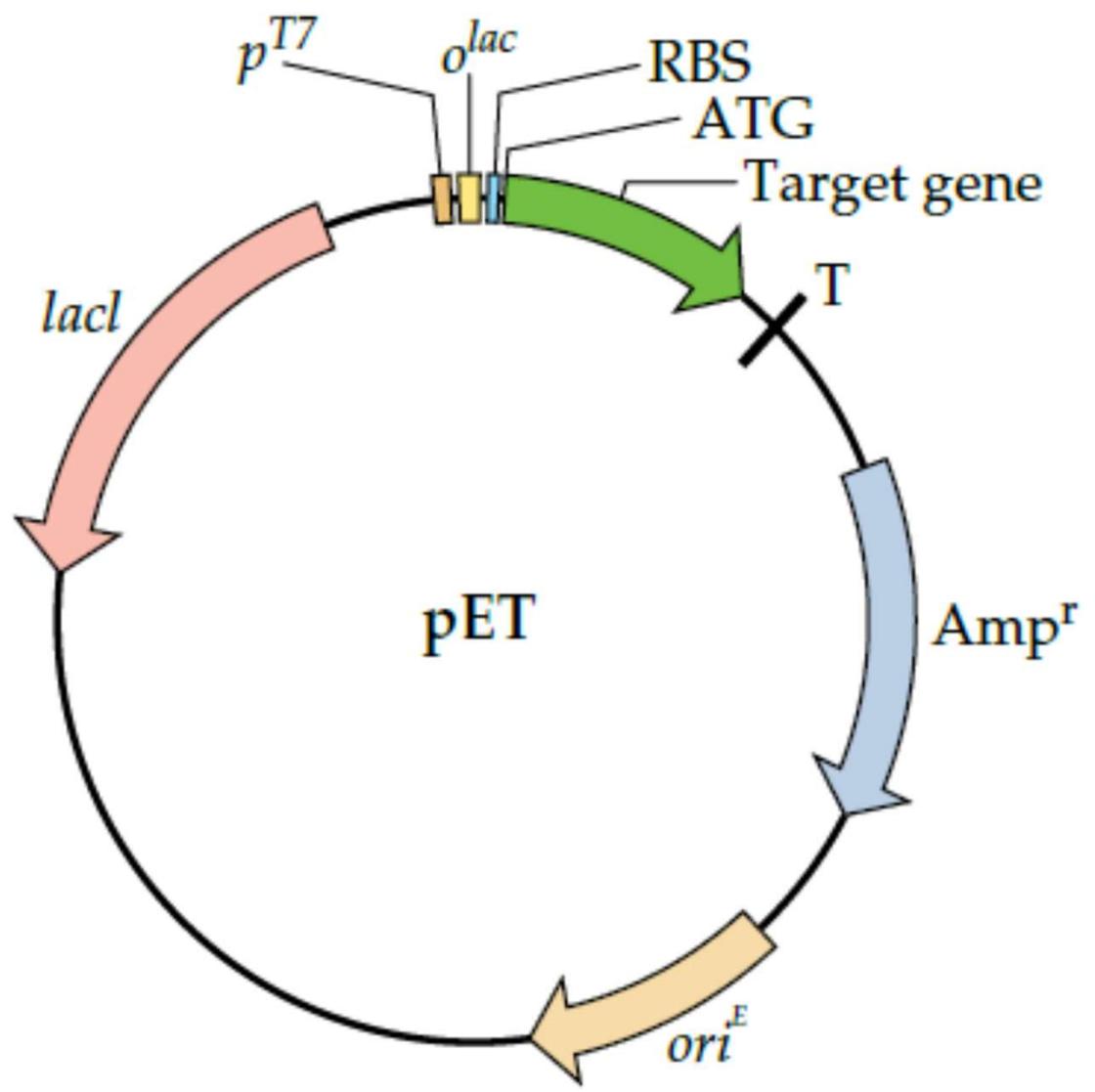
- Contrairement aux plasmides de clonage, ces vecteurs d'expression contiennent des séquences régulatrices appropriées pour la transcription et la traduction du gène d'intérêt.
- Un vecteur d'expression commun doit comprendre les éléments suivants :
  - 1. Séquences régulatrices pour transcrire et traduire le gène d'intérêt.
  - 2. Origine de réplication.
  - 3. Site de clonage multiple (MCS).
  - 4. Marqueur sélectionnable pour la sélection.
  - 5. des Tags pour la séparation et la purification de la protéine exprimée par chromatographie d'affinité.

# Vecteurs d'expression pET

- L'un des systèmes les plus puissants développés dans le but de cloner des gènes d'intérêt et d'exprimer des protéines recombinantes dans *E. coli* est le système de vecteur pET.
- Ce système de vecteurs est fondamentalement un dérivé de la série de plasmides pBR322.
- Il a été conçu pour tirer le meilleur parti du bactériophage T7 et de ses caractéristiques.
- Le gène cible cloné dans ce système de vecteurs est sous le contrôle transcriptionnel et traductionnel des signaux d'expression de l'ARN polymérase du bactériophage T7.

# Vecteurs d'expression pET

- Le plus grand avantage de l'utilisation de l'ARN polymérase T7 est sa spécificité extrêmement élevée pour le promoteur T7.
- De plus, la polymérase possède un niveau élevé d'activité et d'efficacité de traduction grâce aux signaux d'initiation de la traduction.
- Étant donné que le promoteur T7 n'est pas reconnu par l'ARN polymérase de l'hôte, l'expression basale du gène d'insertion peut être évitée.
- L'expression est en outre contrôlée par l'ajout pratique d'un inducteur (IPTG) lorsque la culture bactérienne atteint sa phase logarithmique de croissance.



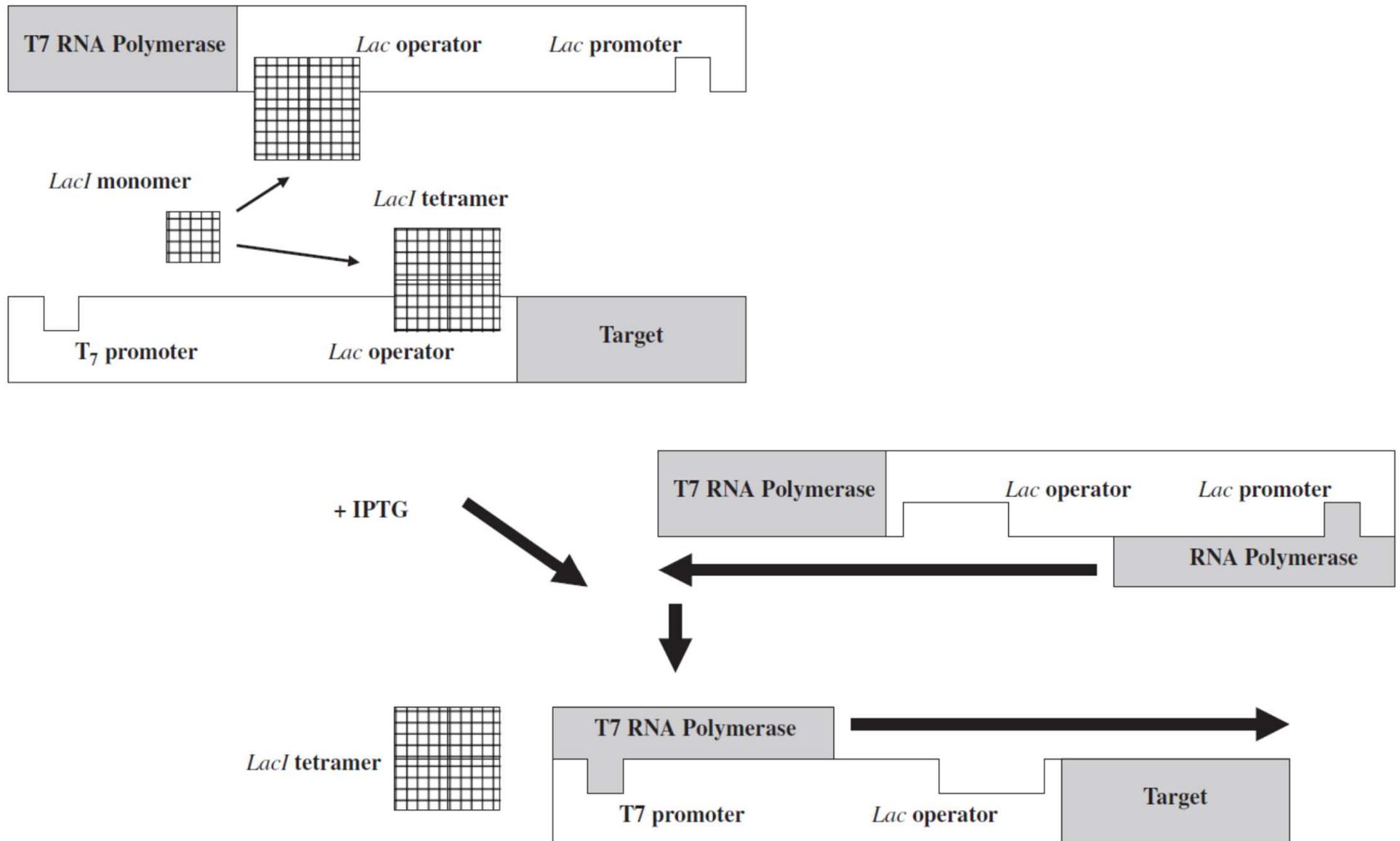
# Vecteurs d'expression pET

- LacO : site de liaison pour LacI. Cet élément inhibe l'activité du promoteur T7 lorsque la protéine LacI est présente, empêchant l'expression du gène d'intérêt.
- RBS : site de liaison du ribosome et élément d'initiation de la traduction du bactériophage T7. Cela permet une production efficace de la protéine d'intérêt.
- ORF : le cadre de lecture ouvert de votre gène d'intérêt est placé ici.
- T7 Terminator : séquence signal pour terminer la transcription réalisée à partir du gène d'intérêt, empêchant la transcription continue. Ampicilline : gène de résistance à l'ampicilline. Il permet au plasmide d'être maintenu par sélection à l'ampicilline dans *E. coli*.
- pBR322 ori : origine de réplication pBR322. Les plasmides porteurs de cette origine ainsi que du gène Rop existent en faible nombre de copies chez *E. coli*.

# Vecteurs d'expression pET

- Rop : Répresseur d'amorce. Il code une petite protéine qui régule le nombre de copies du plasmide. La présence de la protéine Rop, en combinaison avec l'origine de réplication pBR322 sur le plasmide, entraîne un faible nombre de copies du plasmide.
- LacI : Le promoteur naturel *d'E. coli* et la séquence codante du répresseur lac. En l'absence d'induction du système (c'est-à-dire sans IPTG), la protéine LacI réprime la transcription du gène d'intérêt à partir du promoteur T7lac, ainsi que la transcription de l'ARN polymérase T7 à partir du promoteur LacUV5 dans les souches hôtes utilisées pour la production de protéines recombinantes.

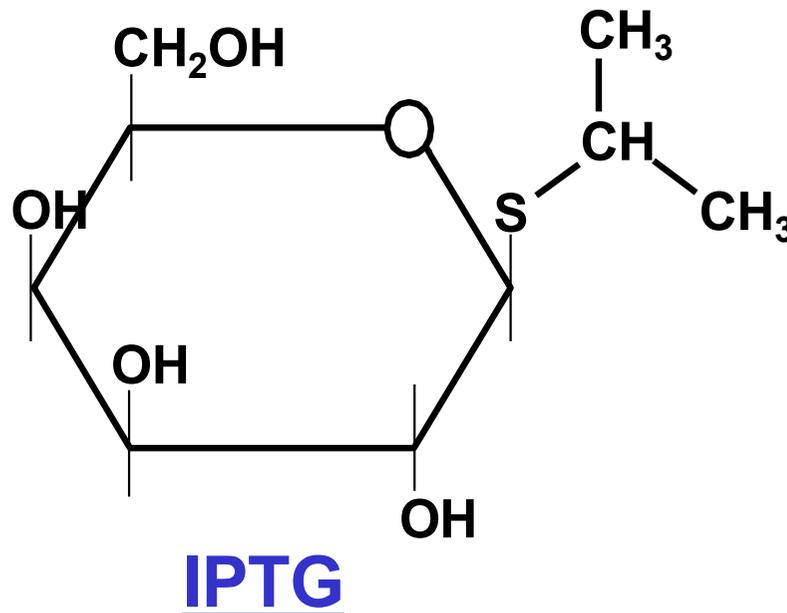
# Control du systeme pET



# IPTG : Structure - Utilisation

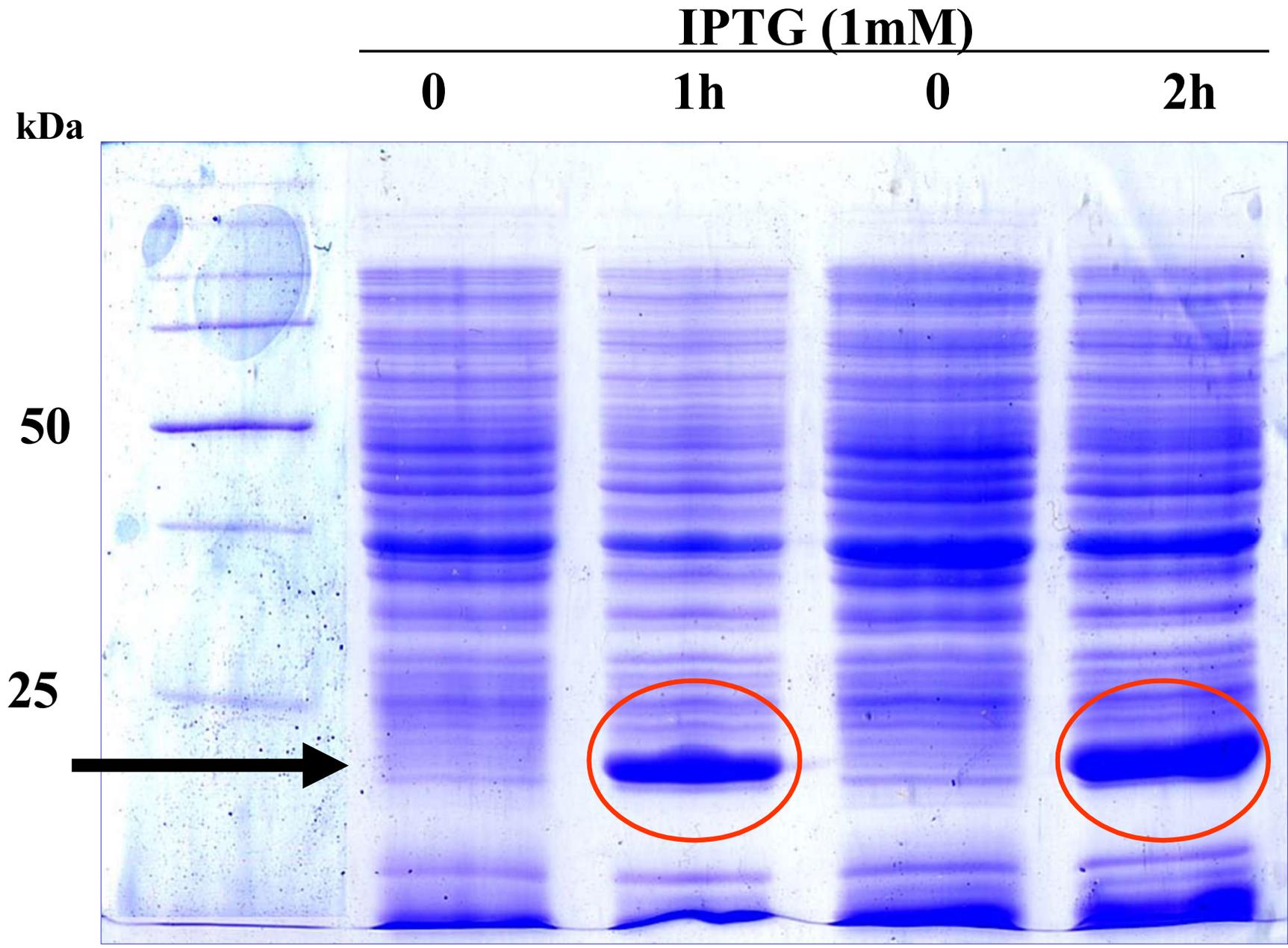
La présence d'IPTG permet d'induire l'activité du gène *LacZ*, codant la  $\beta$ -galactosidase permettant l'hydrolyse du lactose (Gal + Glu) et qui favorise son utilisation, en se fixant et en inhibant le répresseur *lac*.

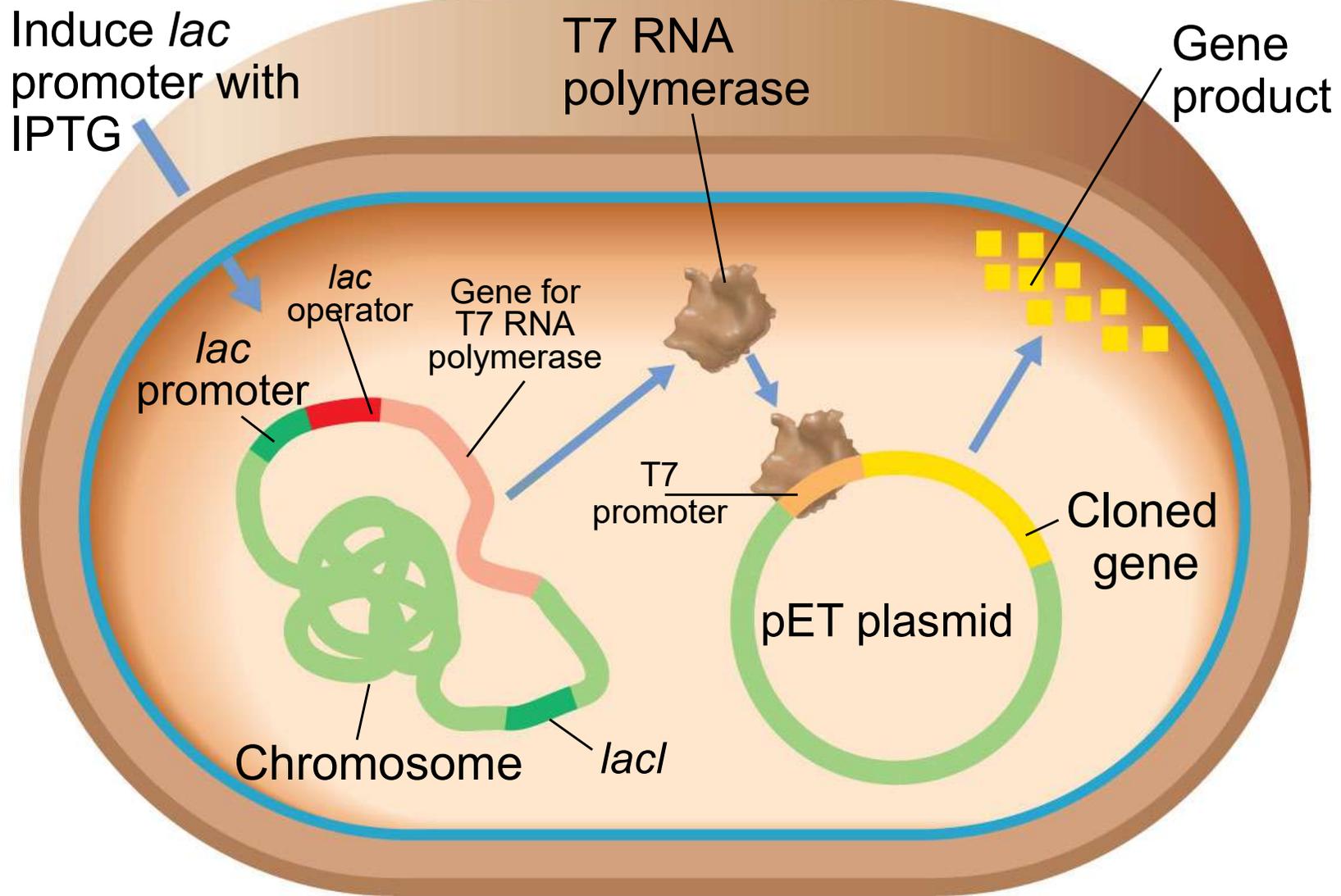
Lorsqu'un gène d'intérêt est placé sous le contrôle du promoteur du gène *lacZ*, la présence d'IPTG permet à ce gène d'être exprimé



(\*) **IPTG** = Isopropyl- $\beta$ -D-1-thiogalactopyranoside  
(analogue structural de l'allolactose)

# Induction par l'IPTG de la synthèse d'une protéine recombinante





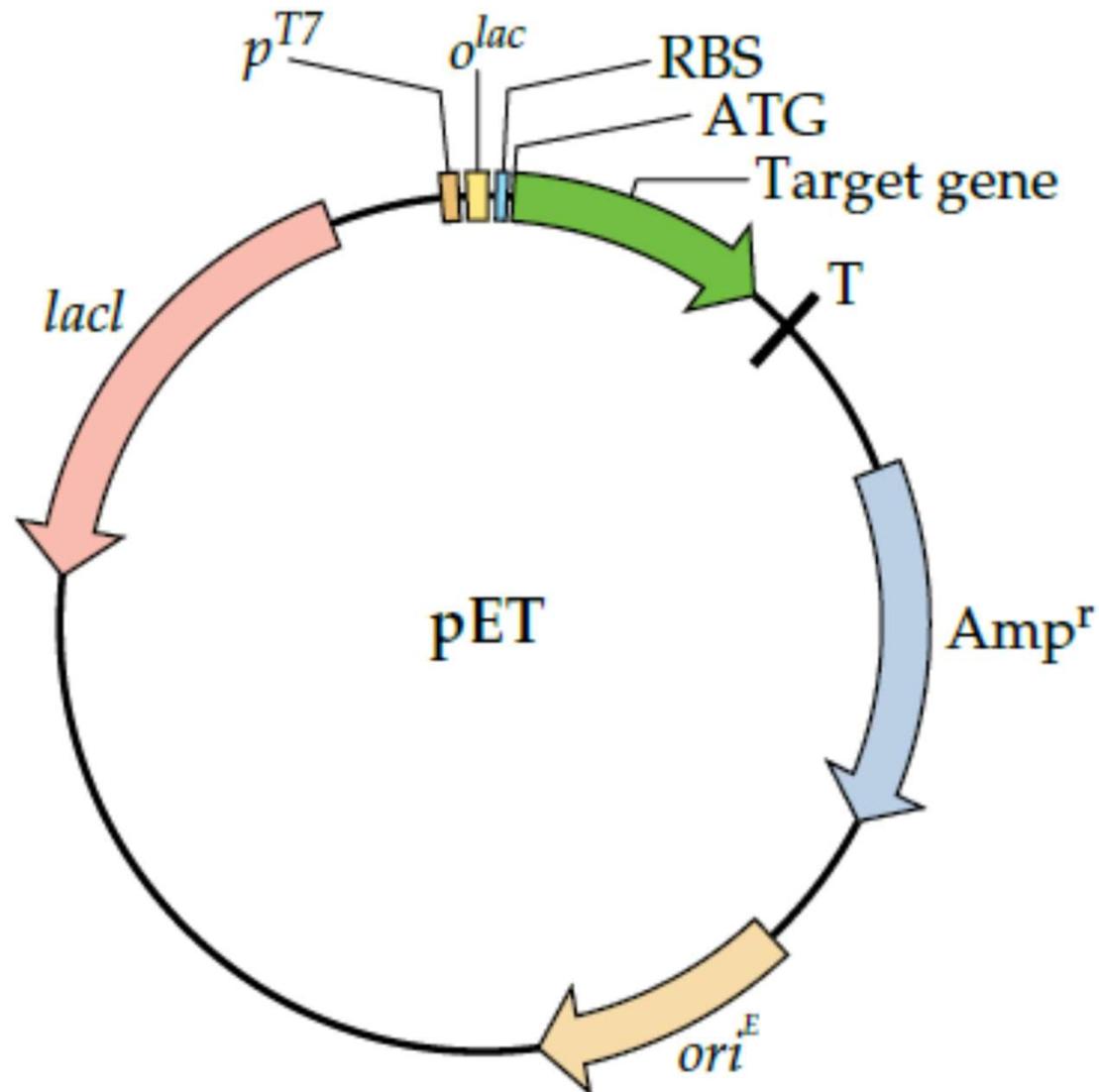
# Vecteurs d'expression pET

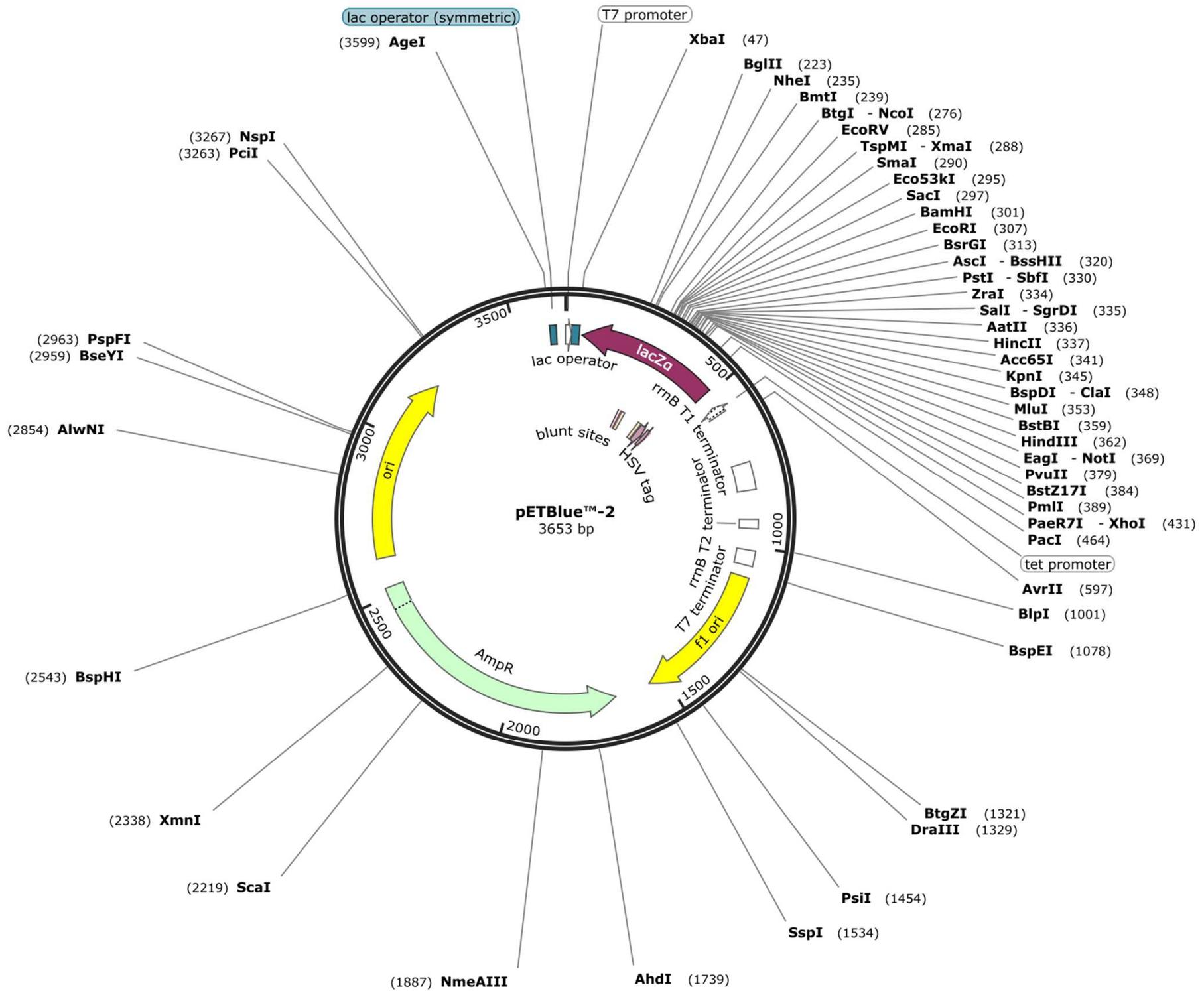
2 types:

- Deux variantes différentes du système pET sont disponibles : les **vecteurs transcriptionnels** et les **vecteurs traductionnels**.
- Les gènes qui portent leur propre site de liaison au ribosome (RBS) et le codon de départ ATG sont souvent clonés dans les vecteurs transcriptionnels. Ces types de vecteurs sont généralement utilisés pour le clonage et l'expression de gènes procaryotes.
- Les vecteurs qui portent un site de liaison au ribosome (RBS) du phage T7 sont appelés vecteurs de traduction. En général, ces vecteurs sont utilisés pour le clonage et l'expression de gènes eucaryotes

# Vecteurs d'expression pET

2 types:







# Vecteurs d'expression pET

- **Avantages :**
- Forte expression : le système de régulation de la transcription et de la traduction T7 permet une production de très haut niveau de protéines d'intérêt, dans de nombreux cas proche de 50 % de la protéine totale de la culture.
- Expression étroitement contrôlée : l'expression du gène d'intérêt est généralement très fortement réprimée en l'absence d'IPTG ajouté, et cet état « off » est très robuste pour la plupart des gènes d'intérêt dans la plupart des souches hôtes.

# Vecteurs d'expression pET

- **Inconvénients**
- Exigences relatives à l'hôte : les vecteurs pET doivent être conservés dans une souche *d'E. coli* dépourvue du gène de l'ARN polymérase T7 et doivent être transférés vers une souche hôte distincte contenant le gène de l'ARN polymérase T7 avant l'induction de l'expression protéique.
- Expression potentiellement faible chez certains hôtes : même en l'absence d'IPTG, il peut y avoir une expression de faible niveau de l'ARN polymérase T7 à partir du promoteur LacUV5 dans certaines souches hôtes d'expression, ce qui pourrait entraîner une toxicité bactérienne pour certains gènes d'intérêt dans certaines souches hôtes.

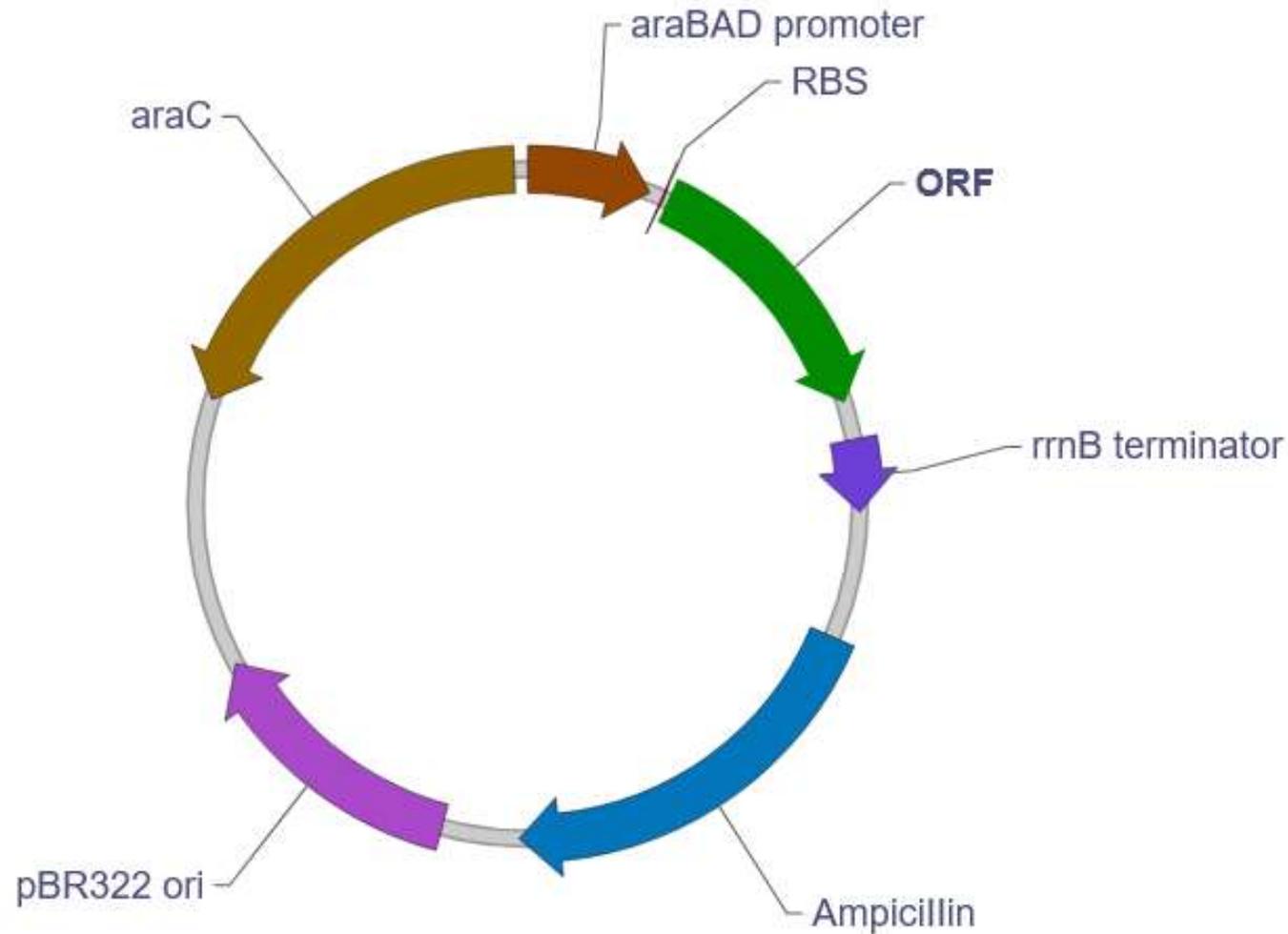
# Vecteurs d'expression pBAD

- Le système vectoriel pBAD est un système fiable et contrôlable pour l'expression de protéines recombinantes dans les bactéries.
- Ce système est basé sur l'opéron araBAD, qui contrôle le métabolisme de la L-arabinose d'*E. coli*.
- Le gène d'intérêt est placé dans le vecteur pBAD en aval du promoteur araBAD, qui pilote l'expression du gène d'intérêt en réponse à la L-arabinose, et est inhibé par le glucose.
- Le contrôle précis des niveaux d'expression rend ce système idéal pour la production de protéines problématiques, telles que des protéines présentant des problèmes de toxicité ou d'insolubilité.

# Vecteurs d'expression pBAD

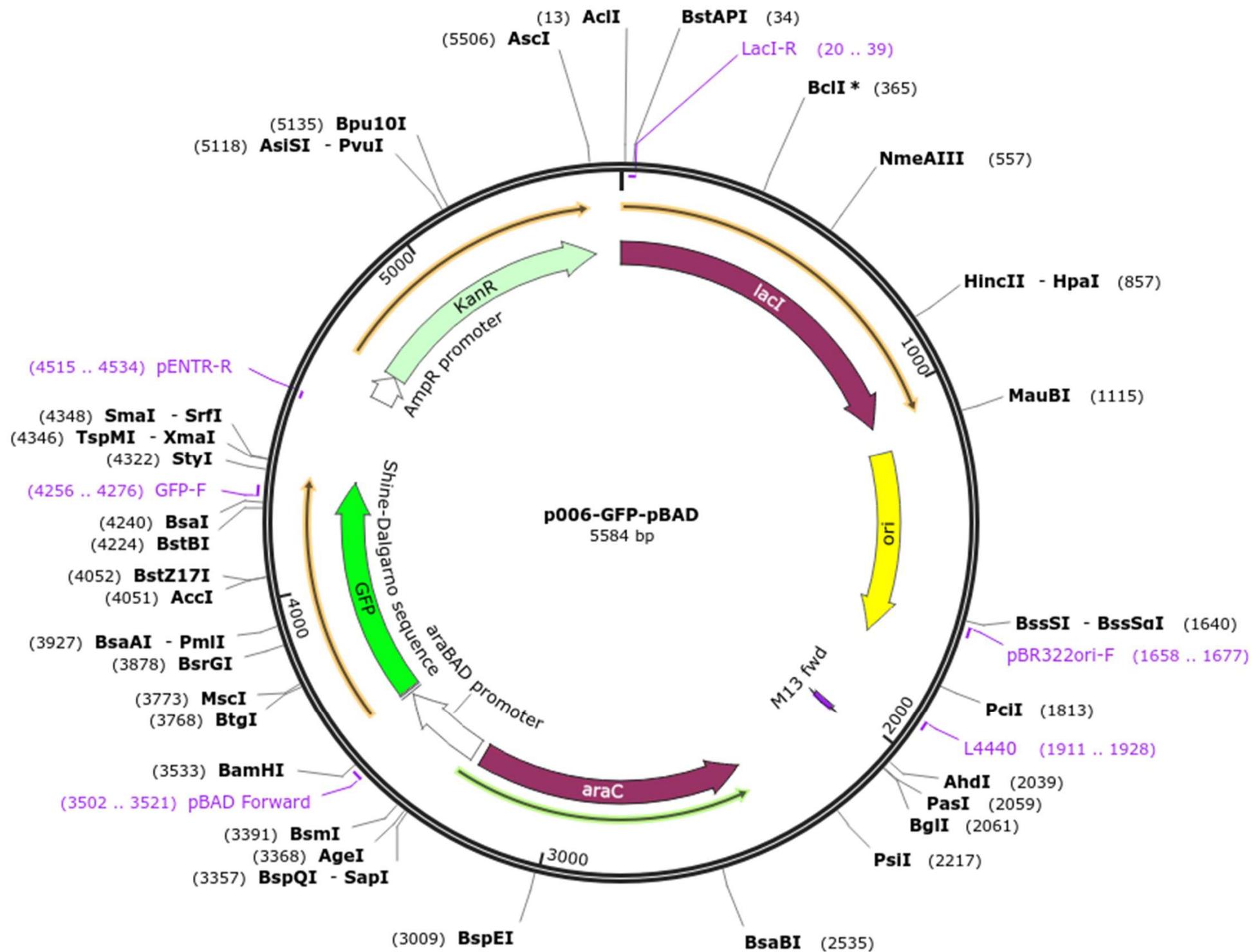
- Le gène d'intérêt est cloné dans le vecteur pBAD et maintenu dans un milieu de croissance dépourvu de L-arabinose. Cela réduit l'instabilité du plasmide qui pourrait résulter de l'expression de protéines d'intérêt qui peuvent être nocives pour les cellules hôtes.
- Par la suite, l'expression du gène d'intérêt peut être induite par l'ajout de L-arabinose au milieu. Les vecteurs pBAD contiennent le promoteur bidirectionnel araBAD, qui pilote l'expression à la fois du gène d'intérêt et de la protéine régulatrice AraC.
- En l'absence de L-arabinose, AraC se dimérise pour former une boucle dans la région promotrice, bloquant la transcription.
- Lorsque de la L-arabinose est ajoutée, AraC change de conformation et se lie à des sites alternatifs dans le promoteur, activant ainsi la transcription.

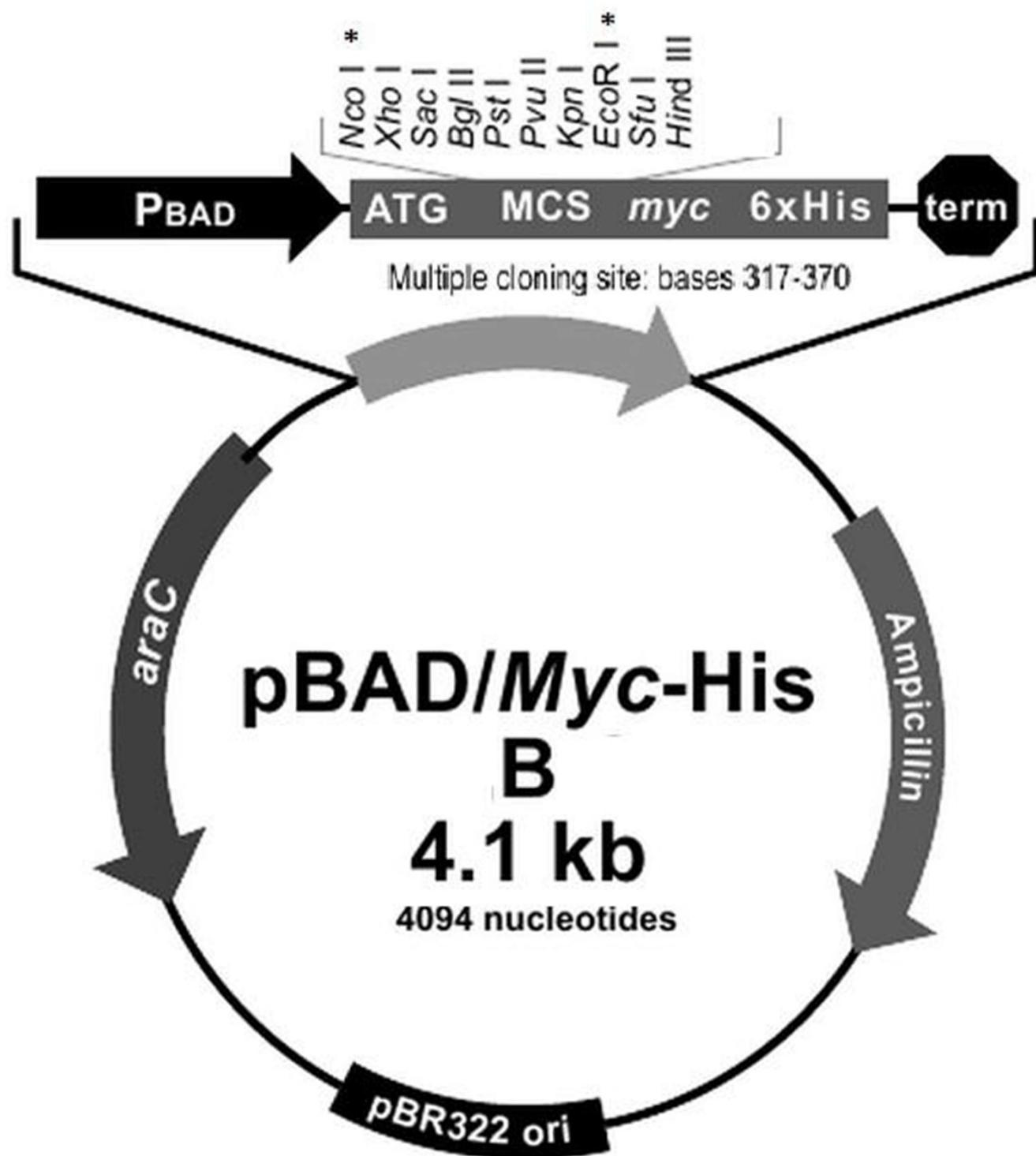
# Vecteurs d'expression pBAD



# Vecteurs d'expression pBAD

- Promoteur araBAD : pilote la transcription du gène d'intérêt lorsque la L-arabinose est présente et que le glucose est absent. Ce promoteur contrôle également l'expression d'AraC.
- RBS : site de liaison du ribosome et élément d'initiation de la traduction du bactériophage T7. Cela permet une production efficace de la protéine d'intérêt.
- ORF : le cadre de lecture ouvert de votre gène d'intérêt est placé ici.
- Termineur rrnB : séquence signal pour terminer la transcription réalisée à partir du gène d'intérêt, empêchant la transcription continue.
- Ampicilline : gène de résistance à l'ampicilline. Il permet au plasmide d'être maintenu par sélection à l'ampicilline dans *E. coli*.
- pBR322 ori : origine de réplication de pBR322. Les plasmides portant cette origine existent en nombre de copies moyen dans *E. coli*.
- araC : code la protéine régulatrice de l'opéron araBAD d'*E. coli*. AraC inhibe l'expression du promoteur araBAD en l'absence de L-arabinose ou en présence de glucose, et active la transcription en présence de L-arabinose et en l'absence de glucose.





# Le vecteur d'expression pBAD

- Avantages
- Expression très étroitement contrôlée : l'expression à partir des vecteurs pBAD est plus étroitement contrôlée que celle à partir des vecteurs pET. L'expression du gène d'intérêt est à des niveaux de base très faibles en l'absence de L-arabinose et est encore plus réprimée en présence de glucose.
- Forte induction : le système opéron araBAD est hautement inductible. Il est possible d'obtenir une induction  $>1000$  fois supérieure après élimination du glucose et ajout de L-arabinose.
- Induction peu coûteuse : la L-arabinose est peu coûteuse, ce qui rend l'expression protéique à grande échelle plus économique.

# Le vecteur d'expression pBAD

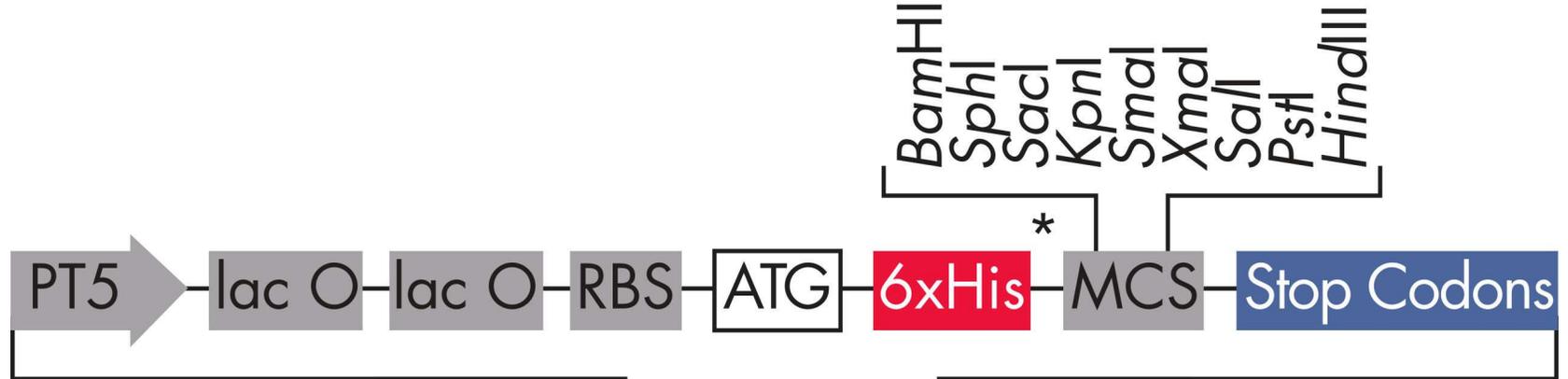
- Inconvénients
- Expression sous-maximale : les vecteurs pBAD ne sont généralement pas capables d'atteindre le niveau d'expression très élevé possible avec les vecteurs pET.
- Métabolisme de l'inducteur (le L-arabinose) : les souches *E. coli* de type sauvage peuvent cataboliser la L-arabinose. Lors de l'expression de la protéine d'intérêt, des souches hôtes mutantes pour le catabolisme de L-arabinose (telles que TOP10 ou LMG194) doivent être utilisées pour éviter une expression incohérente due à l'épuisement de L-arabinose dans le milieu au fil du temps.

# Le vecteur d'expression pQE

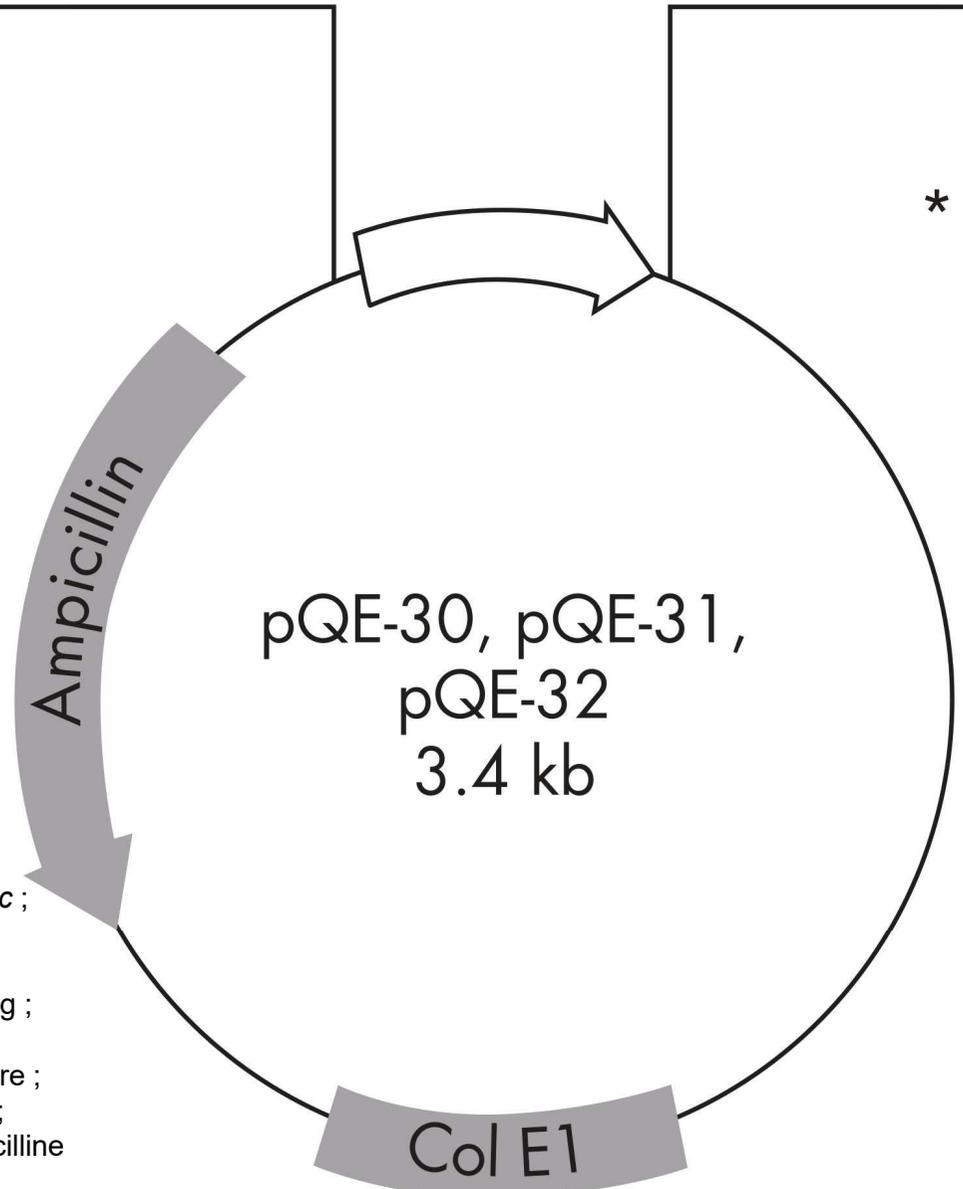
- L'expression des protéines à partir d'un vecteur pQE est basée sur le système de transcription/traduction du promoteur T5.
- Le promoteur T5 en amont de l'insert est reconnu par l'ARN polymérase naturelle de la bactérie et permet la surexpression de la protéine d'intérêt.
- Les souches d'expression utilisées avec ce vecteur sont les souches M15 qui possèdent le plasmide pREP4. Celui-ci confère la résistance à la kanamycine et exprime constitutivement la protéine répresseur lac codé par le gène lacI.
- Le promoteur T5 du phage T5 est reconnu par l'ARN polymérase d'*E. coli* : les vecteurs utilisant le promoteur T5-lacO comprennent un site opérateur lac pour le contrôle de l'expression par le répresseur lac.

# Le vecteur d'expression pQE

- En absence d'IPTG, le promoteur T5, et donc l'expression de la protéine d'intérêt, est inhibé par le répresseur lac. En présence d'IPTG, le répresseur est inhibé ce qui permet la reconnaissance du promoteur T5 par l'ARN polymérase, et la surexpression de la protéine.
- En aval du site de clonage, le vecteur possède une séquence codant pour une extrémité polyhistidine.
- séquence du gène de la résistance à l'ampicilline permet la sélection des bactéries transformées.

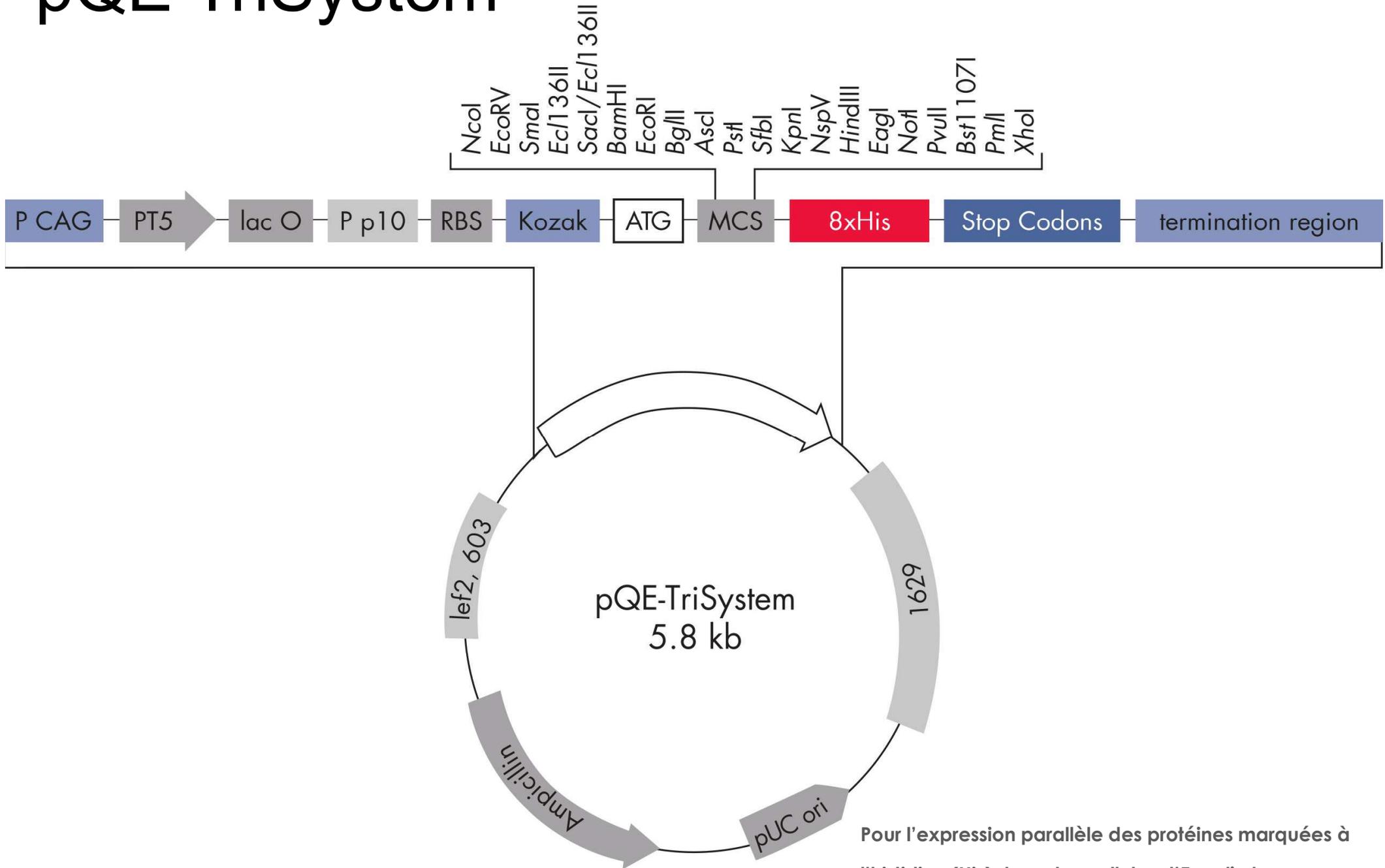


\* pQE-30 --  
 pQE-31 AC  
 pQE-32 - G



PT5 : promoteur T5 ; lac O : opérateur *lac* ;  
 RBS : site de liaison du ribosome ;  
 ATG : codon d'initiation ;  
 6xHis : séquence de marqueurs 6xHis-tag ;  
 MCS : site de clonage multiple ;  
 Codons stop : dans les 3 cadres de lecture ;  
 Col E1 : Col E1, origine de la réplication ;  
 Ampicilline : gène de résistance à l'ampicilline

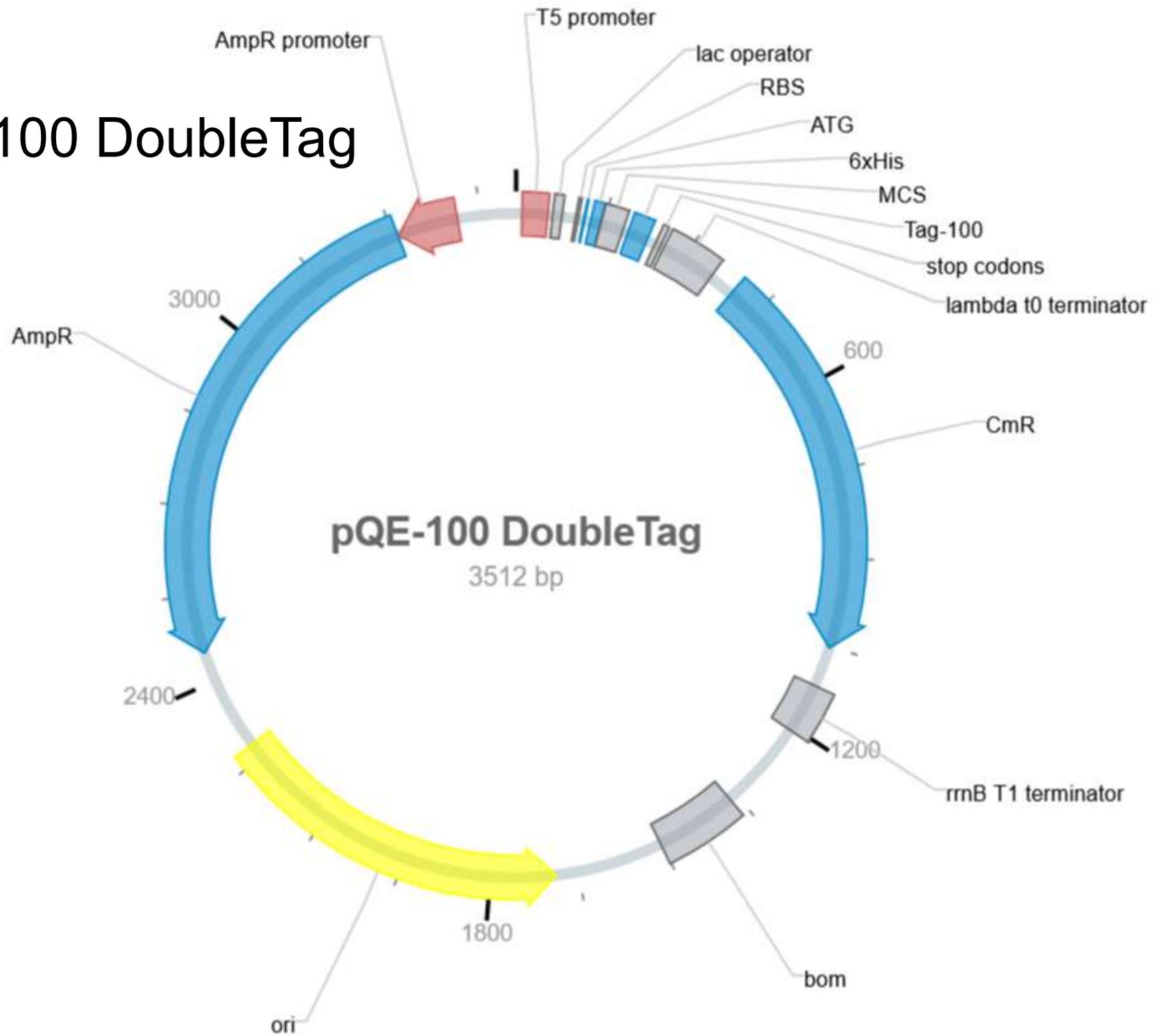
# pQE TriSystem



MCS : site de clonage multiple ;  
 Codons stop : dans les 3 cadres de lecture ;  
 Col E1 : Col E1, origine de la réplique ;  
 Ampicilline : gène de résistance à l'ampicilline

Pour l'expression parallèle des protéines marquées à l'histidine (His) dans des cellules d'*E. coli*, de mammifère et d'insecte infectées au baculovirus, avec une seule construction

# pQE-100 DoubleTag

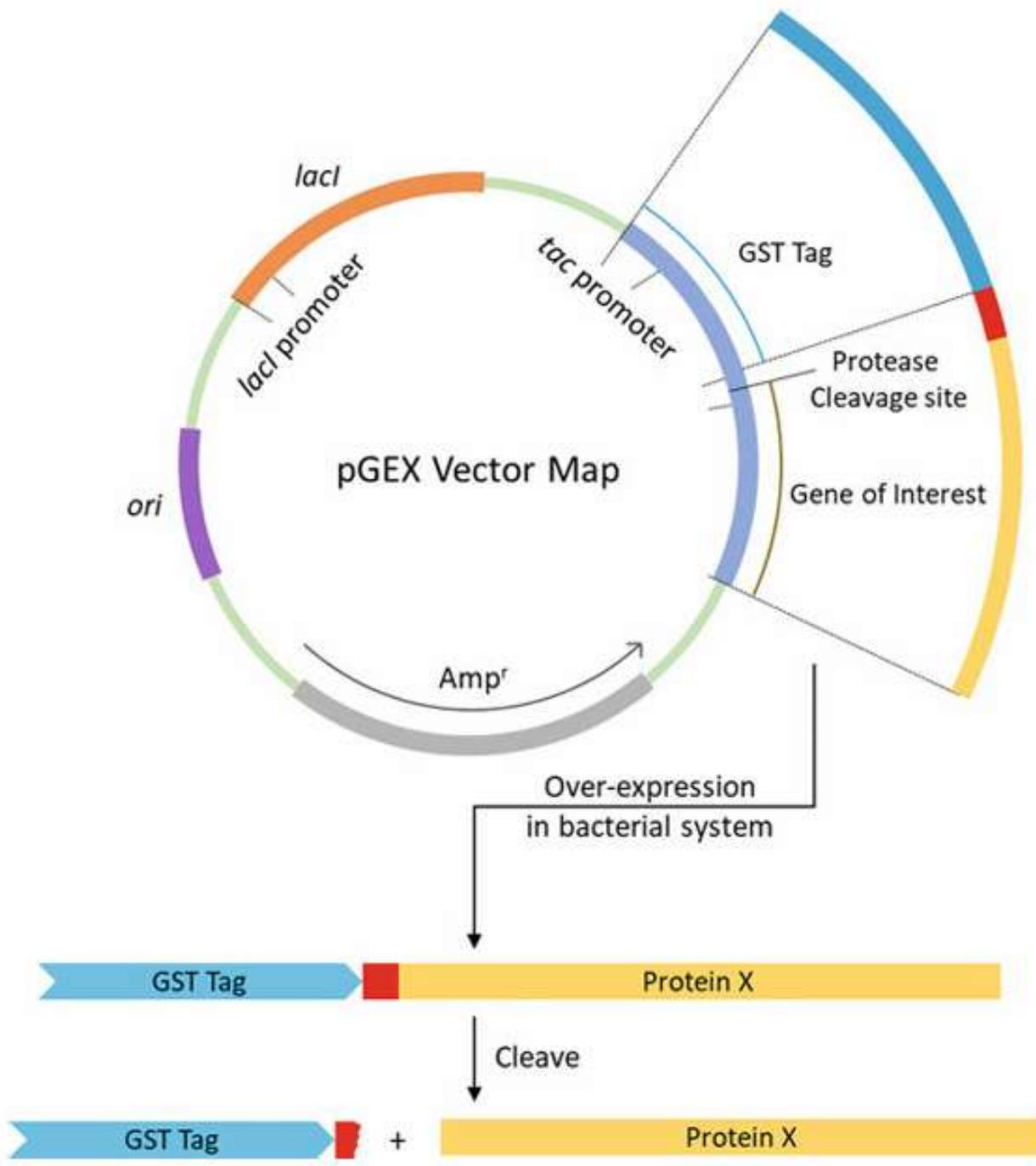


# Le vecteur d'expression pGEX

- Le gène cloné dans ce système plasmidique est marqué avec GST et exprimé sous forme de protéine de fusion
- L'expression de la protéine de fusion est sous le contrôle du promoteur tac (Ptac), qui est un promoteur inductible par IPTG qui maintient un contrôle plus strict sur son expression de protéine recombinante.
- Il existe 13 vecteurs pGEX disponibles dans le commerce.
- Ces vecteurs plasmidiques diffèrent subtilement les uns des autres en termes de choix de sites de restriction ainsi que de clivage des protéases.
- Neuf de ces vecteurs plasmidiques ont des sites de reconnaissance pour différentes protéases (par exemple, la thrombine, la présépsion, le facteur Xa) entre le gène de la protéine et le tag GST. Ce site permet de séparer le tag GST de la protéine après la purification.

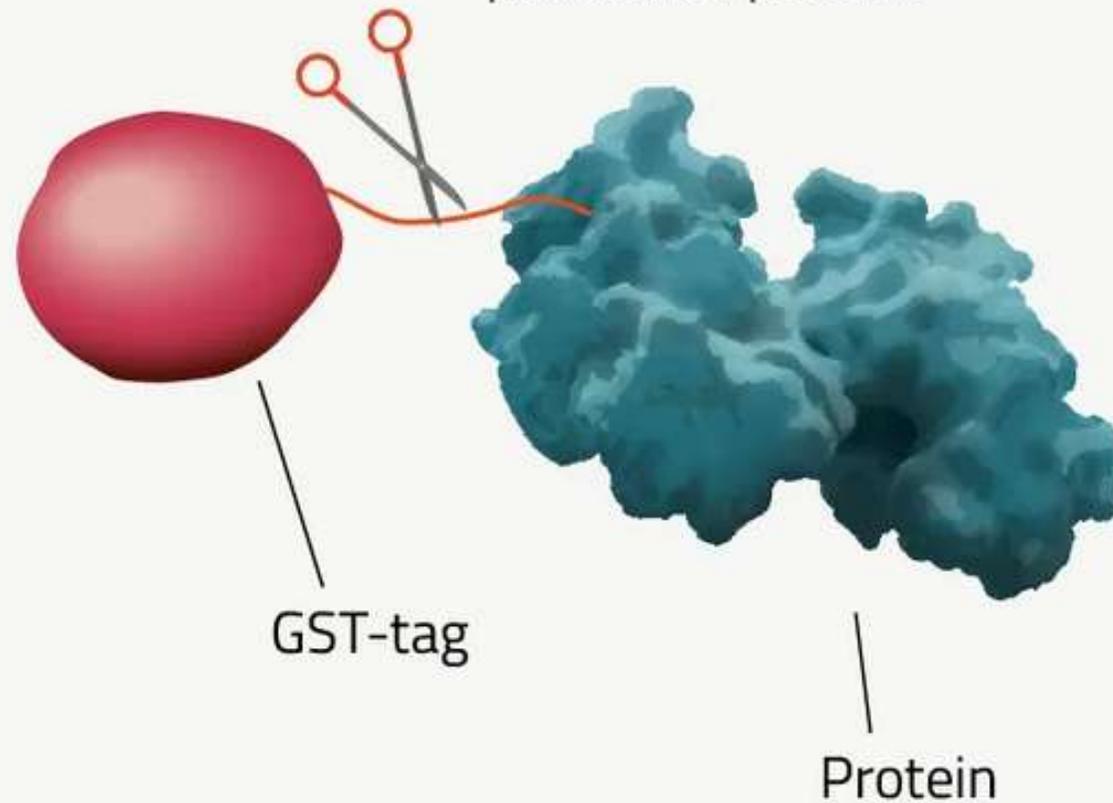
# Le vecteur d'expression pGEX

- Qu'est-ce que le tag GST ?
- Il existe plusieurs tags d'affinité protéique utilisés pour la purification des protéines et le tag GST en fait partie.
- GST est l'abréviation de Glutathion-S-transférase et fait référence à une protéine entière, pas seulement à quelques acides aminés comme le tag Rho1D4.
- Le tag enzymatique GST est utilisé pour la purification des protéines depuis la fin des années 1980 (Smith & Johnson 1988) et a été établi comme un moyen fiable pour les tests de purification des protéines qui nécessitent le plus haut niveau de pureté



# Glutathione S-transferase (GST) tag

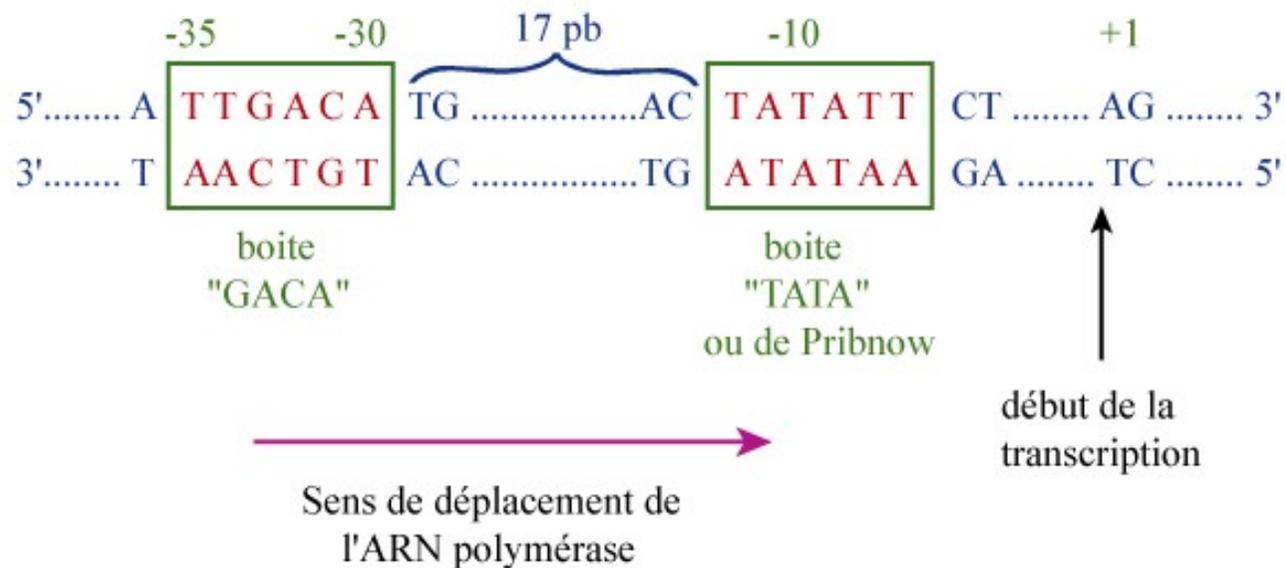
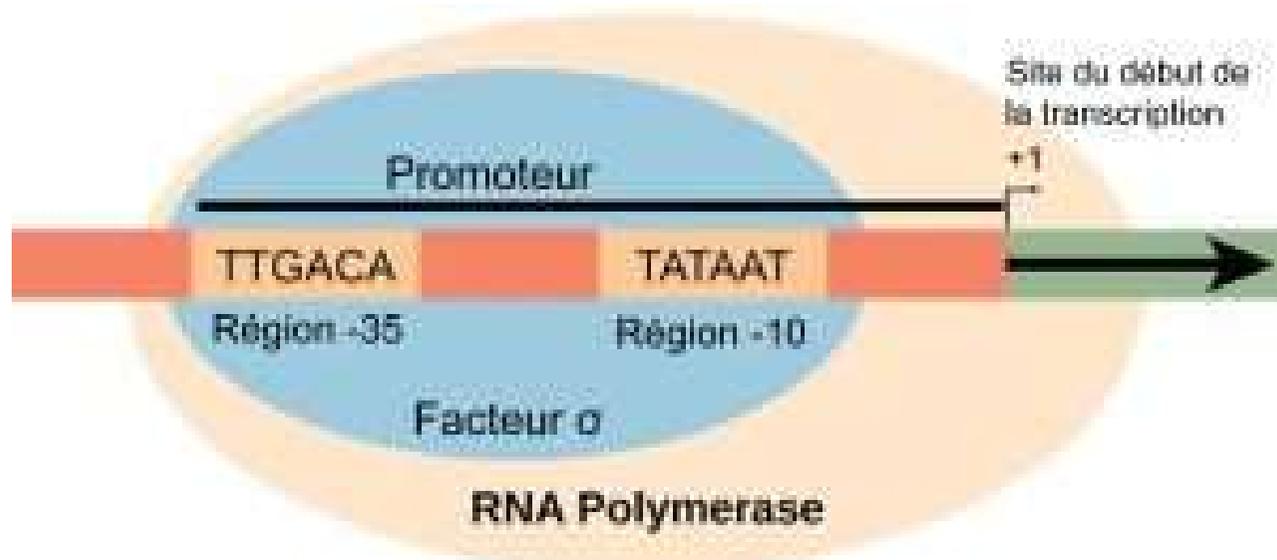
Removal of tag after purification possible



# Promoteurs d'expression

- Un gène est exprimé, c'est-à-dire quand il est lu et copié en ARN.
- Le **promoteur** d'un gène est un segment d'ADN qui contrôle l'expression du gène. Chez les procaryotes, le promoteur se lie à une enzyme, **l'ARN polymérase**, qui lit la séquence d'ADN et génère la molécule d'ARN correspondante.
- Le promoteur définit si un gène doit être transcrit et à quel taux.
- Certains promoteurs sont constamment actifs alors que d'autres ne sont actifs que dans certaines cellules et selon les besoins.

# Promoteurs d'expression



# Promoteurs d'expression

- Les promoteurs bactériens canoniques sont composés de deux "boîtes" de séquences conservées : la boîte -35 et la boîte de Pribnow ou boîte -10, située respectivement environ 35 et 10 nucléotides en amont du premier nucléotide transcrit de l'ARN.

Probabilité d'observation pour chaque nucléotide

séquence en -10

T	A	T	A	A	T
77 %	76 %	60 %	61 %	56 %	82 %

séquence en -35

T	T	G	A	C	A
69 %	79 %	61 %	56 %	54 %	54 %

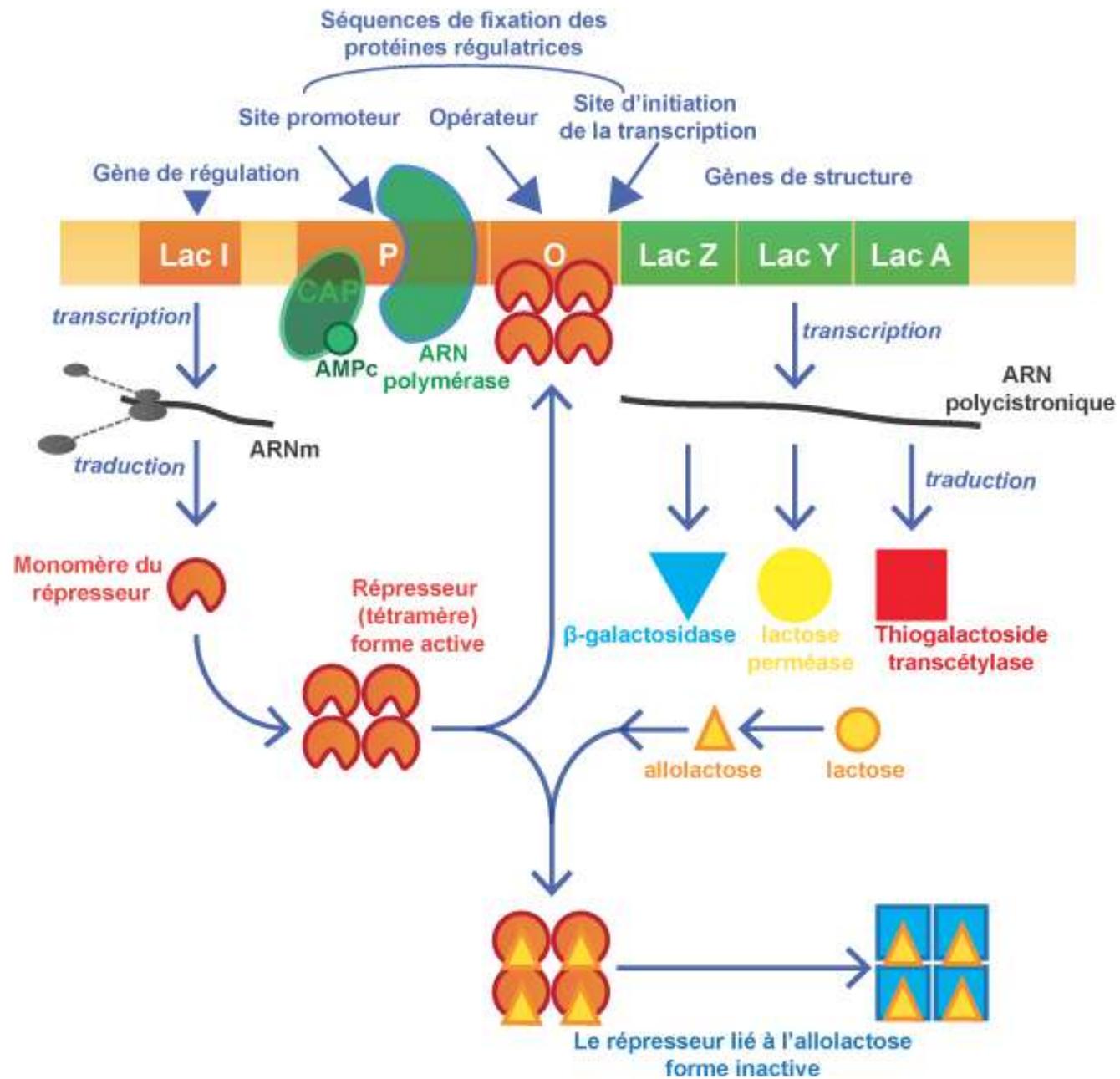
# Promoteurs d'expression

- Promoteurs Plac / PlacUV5
- Promoteur tac/trc ;
- Promoteur T7 ;
- Promoteur arapBAD;
- Promoteur pL

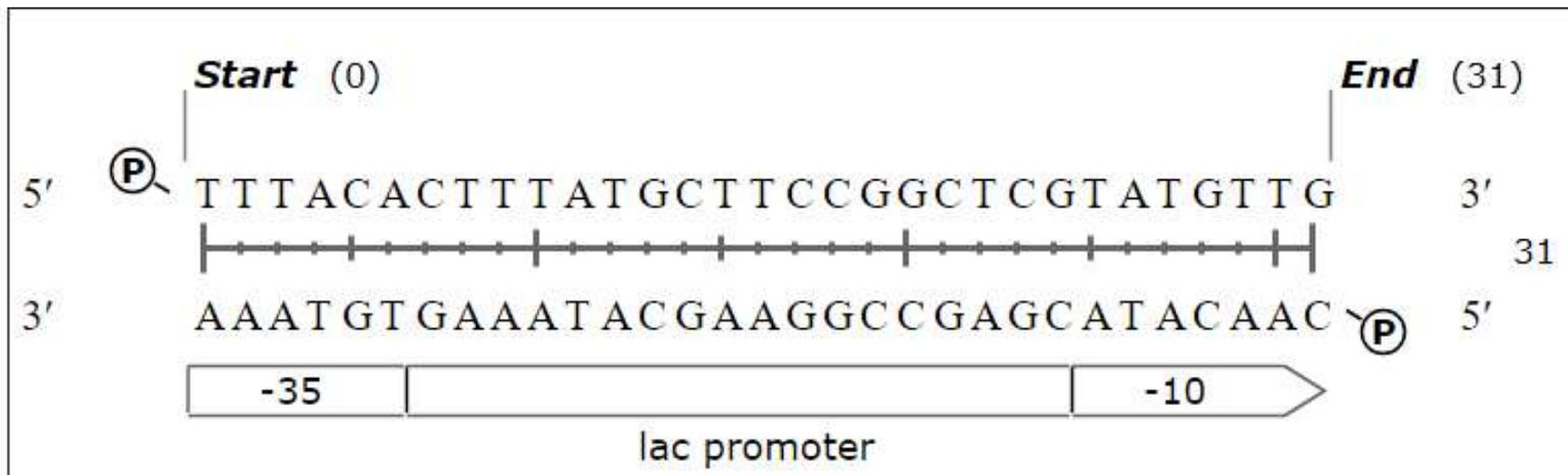
# Promoteur Plac

- L'un des promoteurs procaryotes les plus couramment utilisés est le promoteur pLac inductible négatif. Ce promoteur nécessite l'élimination du répresseur lac (protéine lacI) pour que la transcription soit activée.
- En présence de lactose ou d'IPTG analogue du lactose, le répresseur lac subit un changement conformationnel qui l'élimine des sites lacO au sein du promoteur et cesse la répression du gène cible.
- Un système inductible lac simplifié se retrouve dans de nombreux vecteurs d'expression bactérienne.

# Exp : opéron lactose



# Séquence du promoteur lac



# Promoteur lacUV5

- Le promoteur lacUV5 est un promoteur muté de l'opéron lac *Escherichia coli* qui est utilisé en biologie moléculaire pour piloter l'expression génique sur un plasmide.
- lacUV5 est très similaire au promoteur lac classique, ne contenant que 2 mutations de paires de bases dans la région de l'hexamère -10, par rapport au promoteur lac.
- LacUV5 est l'un des promoteurs les plus couramment utilisés en biologie moléculaire car il ne nécessite aucun activateur supplémentaire et il entraîne des niveaux élevés d'expression génique.

# Promoteur lacUV5

- Le promoteur muté lacUV5 varie de la séquence du promoteur lac de 2 bases dans le consensus -10 (TATGTT dans le promoteur lac ; TATAAT en lacUV5) ce qui le rend beaucoup plus fort.

*LacUV5*

```
TCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAAATGTGTGGAATTGTGAGCGGATAACAATTT  
CACACAGGAAACAGCT
```

*LacZ*

```
TCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGAAATTGTGAGCGGATAACAATTT  
CACACAGGAAACAGCT
```

position

^-35

^-10

^+1

# Promoteur lacUV5

- La séquence du promoteur lacUV5 est plus conforme à la séquence consensus reconnue par les facteurs sigma bactériens que le promoteur lac-traditionnel que le promoteur lac.
- Pour cette raison, lacUV5 recrute plus efficacement l'ARN polymérase, conduisant ainsi à une transcription plus élevée des gènes cibles.
- De plus, contrairement au promoteur lac, lacUV5 fonctionne indépendamment des protéines activatrices ou d'autres éléments régulateurs cis (à l'exception des régions promotrices -10 et -35).

# Promoteur lacUV5

- Bien qu'aucun activateur ne soit nécessaire, l'expression du promoteur lacUV5 peut être régulée par le répresseur LacI et peut être induite avec l'IPTG, qui est un inducteur efficace de l'expression des protéines lorsqu'il est utilisé dans la plage de concentration de 100  $\mu$ M à 1,5 mM.

# lacUV5 Séquence

*LacUV5*

TCACTCATTAGGCACCCCAGGCTTT**AC**ACTTTATGCTTCCGGCTCGTATAATTGTGTGGAATTGTGAGCGGATAACAATTT  
CACACAGGAAACAGCT|

*LacZ*

TCACTCATTAGGCACCCCAGGCTTT**AC**ACTTTATGCTTCCGGCTCGTATGTTTGTGTGAAATTGTGAGCGGATAACAATTT  
CACACAGGAAACAGCT|

position

<sup>-35</sup>

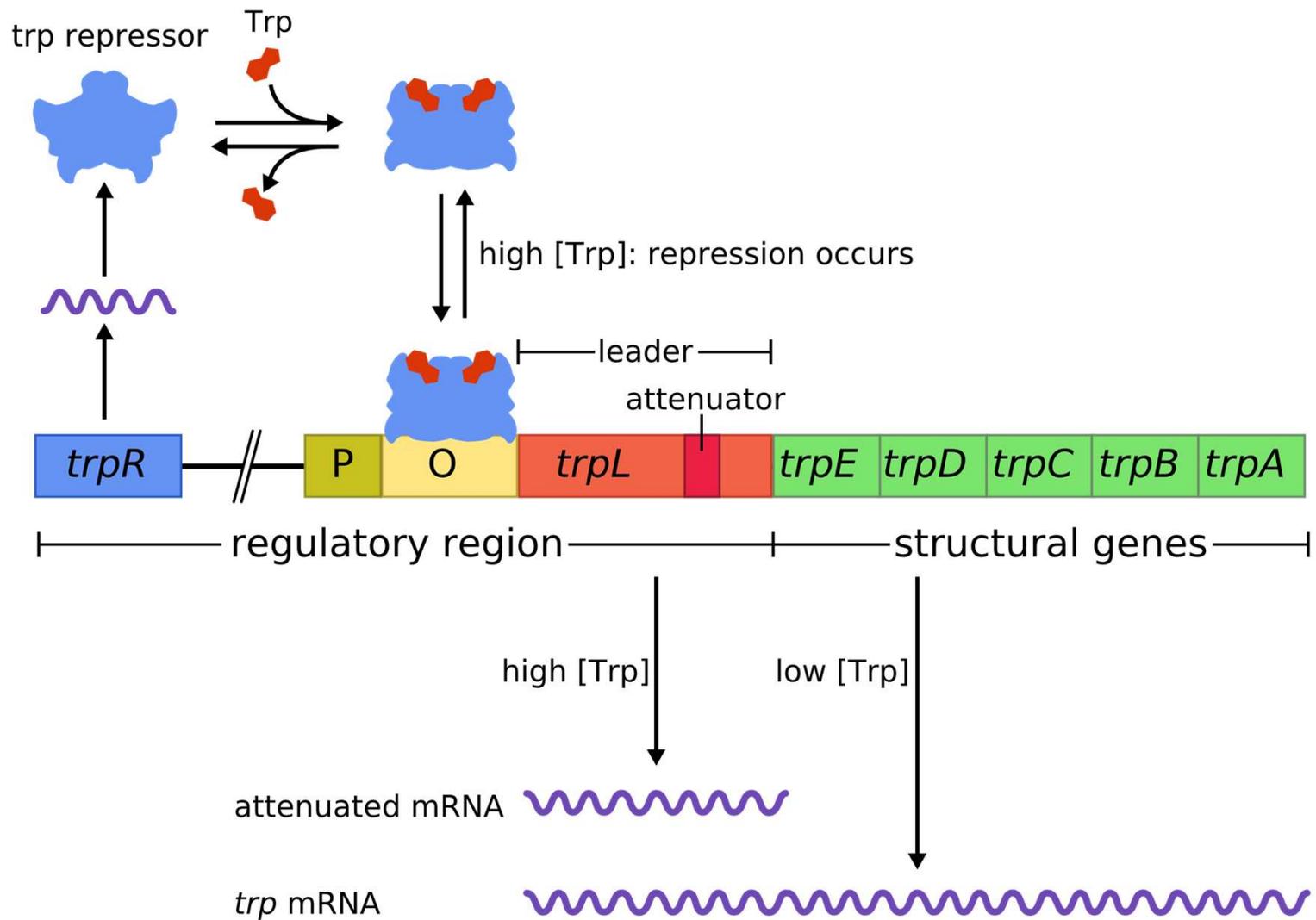
<sup>-10</sup>

<sup>+1</sup>

# Promoteur tac

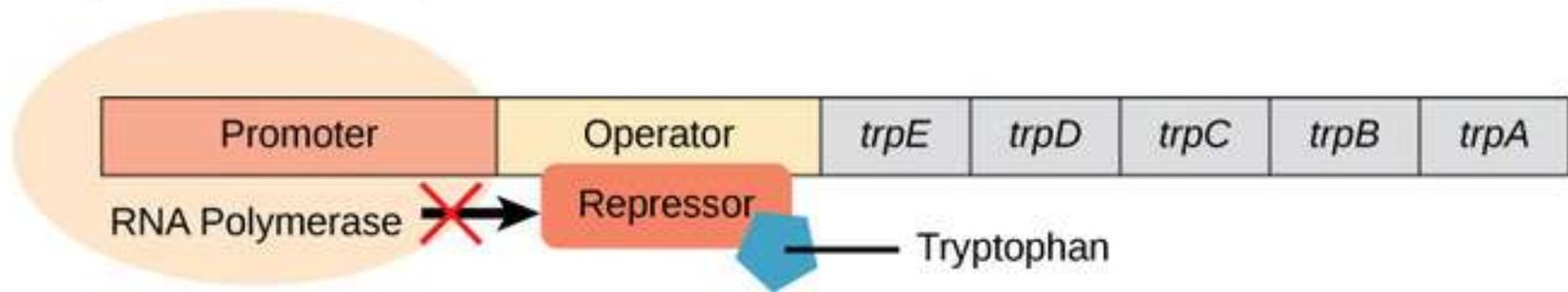
- Le promoteur Tac (abrégé en Ptac), est un promoteur d'ADN synthétique, issu de la combinaison des promoteurs des opérons trp et lac. Il est couramment utilisé pour la production de protéines chez *Escherichia coli*.
- Le promoteur tac est utilisé pour contrôler et augmenter l'expression d'un gène cible et participe à la surexpression de protéines recombinantes. Il doit son nom aux deux promoteurs qui composent sa séquence : les promoteurs « trp » et « lac ».

# Opéron trp

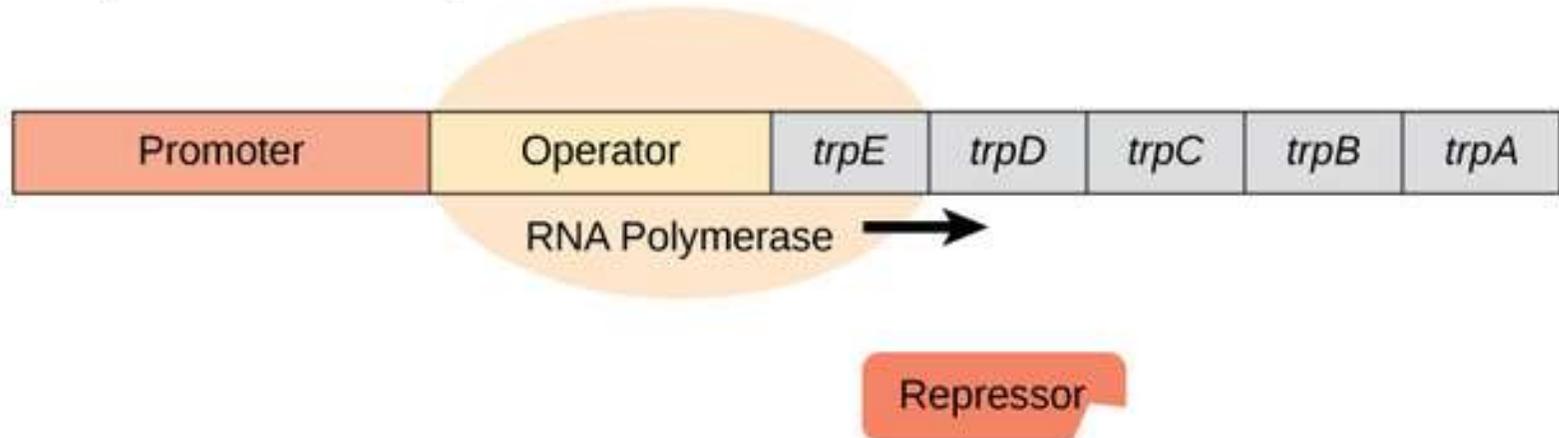


# Opéron trp

When tryptophan is present, the trp repressor binds the operator, and RNA synthesis is blocked.



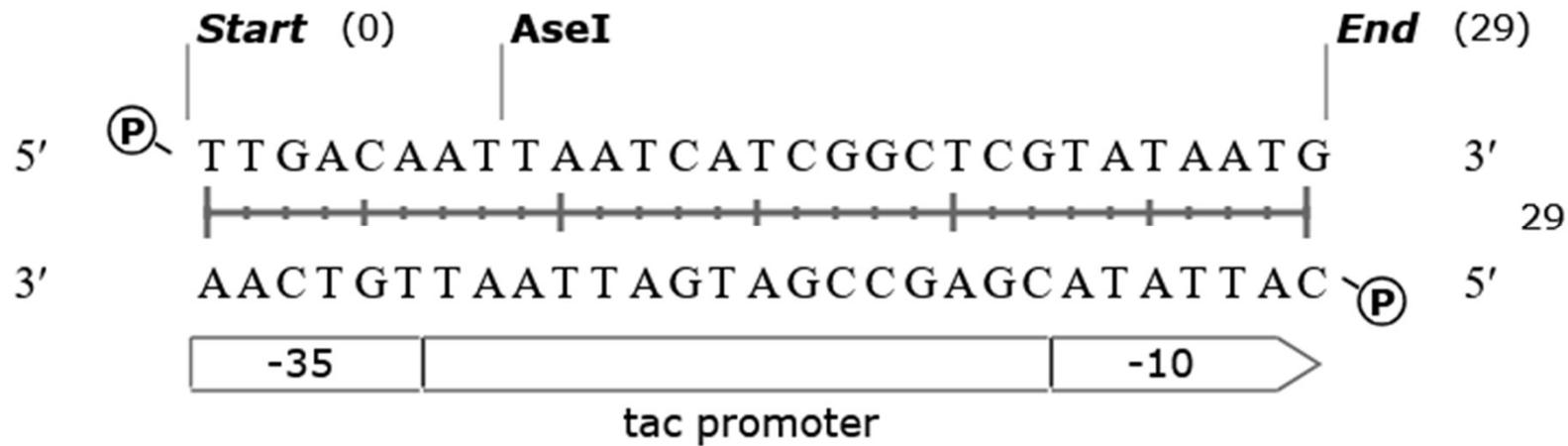
In the absence of tryptophan, the repressor dissociates from the operator, and RNA synthesis proceeds.



# Promoteur tac

- Les promoteurs bactériens sont constitués de deux parties : la région « -35 » et la région « -10 » (la boîte de Pribnow).
- Ces deux régions se lient au facteur sigma de l'ARN polymérase, qui initie ensuite la transcription du gène en aval.
- Le promoteur tac est constitué de la région « -35 » du promoteur trp et de la région « -10 » du promoteur lac (et diffère d'un **promoteur trc** apparenté d'une paire de base).
- Le promoteur tac est donc inductible par l'IPTG, tout en permettant une expression génique maximale plus élevée que les promoteurs lac ou trp.
- Cela le rend adapté à la production protéique hautement efficace d'une protéine recombinante.

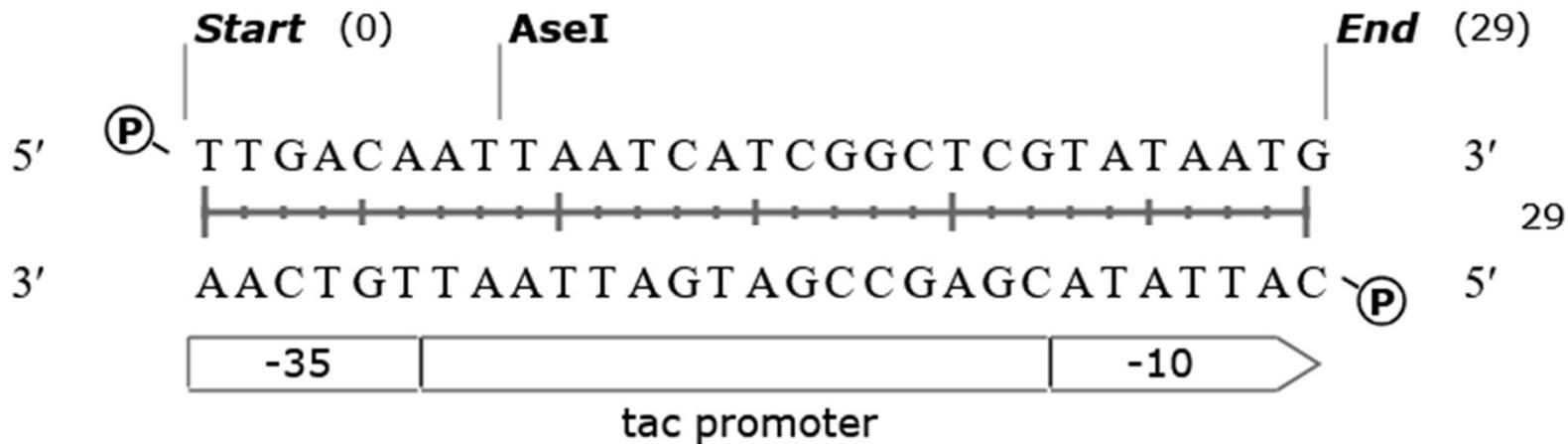
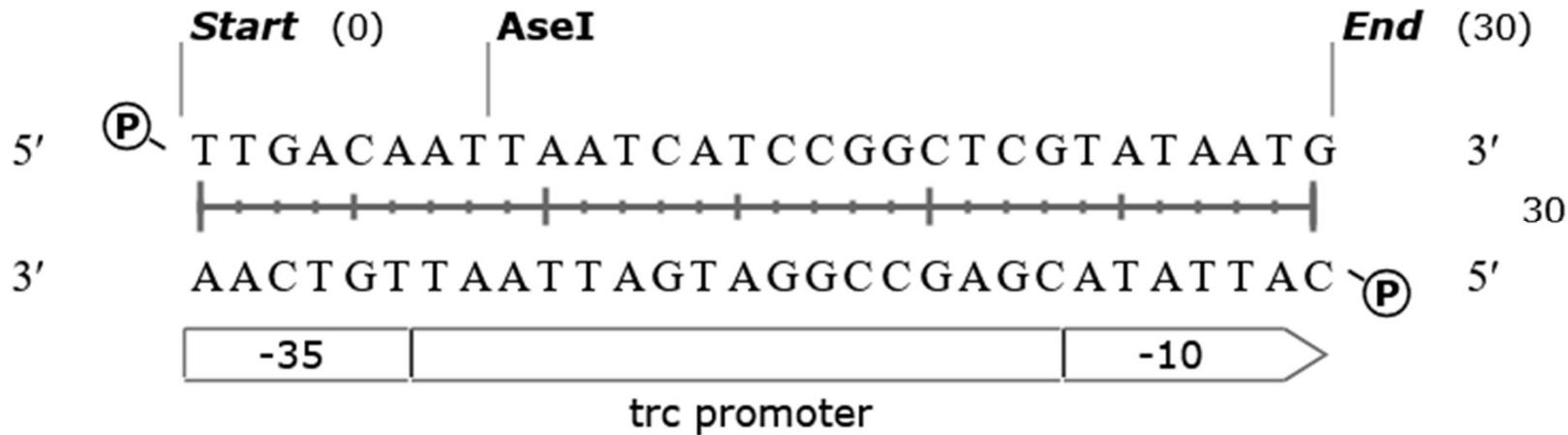
# Séquence promoteur tac



# Promoteur trc

- Le promoteur trc est un hybride des promoteurs lac et trp, plus puissant que le promoteur lac.
- Il a été créé par insertion d'une pb dans la séquence de 16 pb entre les séquences consensus -35 et -10 du promoteur tac, afin d'obtenir une distance de consensus optimale de 17 pb entre les signaux -35 et -10.
- Ce promoteur inclut la séquence opératrice lacO, qui peut être liée par le répresseur LacI.
- Pour que l'expression génique soit inductible par le promoteur trc, le répresseur LacI est nécessaire. Ce répresseur se lie à l'ADN opérateur avec une grande spécificité et inhibe l'expression génique jusqu'à l'ajout d'un inducteur comme l'allolactose ou l'IPTG, un analogue du promoteur.
- L'inducteur se lie au répresseur, provoquant une modification de sa forme. Ainsi modifié, le répresseur est incapable de se lier à l'opérateur, ce qui permet à l'ARN polymérase (RNAP) de transcrire les gènes suivants et conduit ainsi à des niveaux élevés de protéines codées.

# Séquence promoteur tac

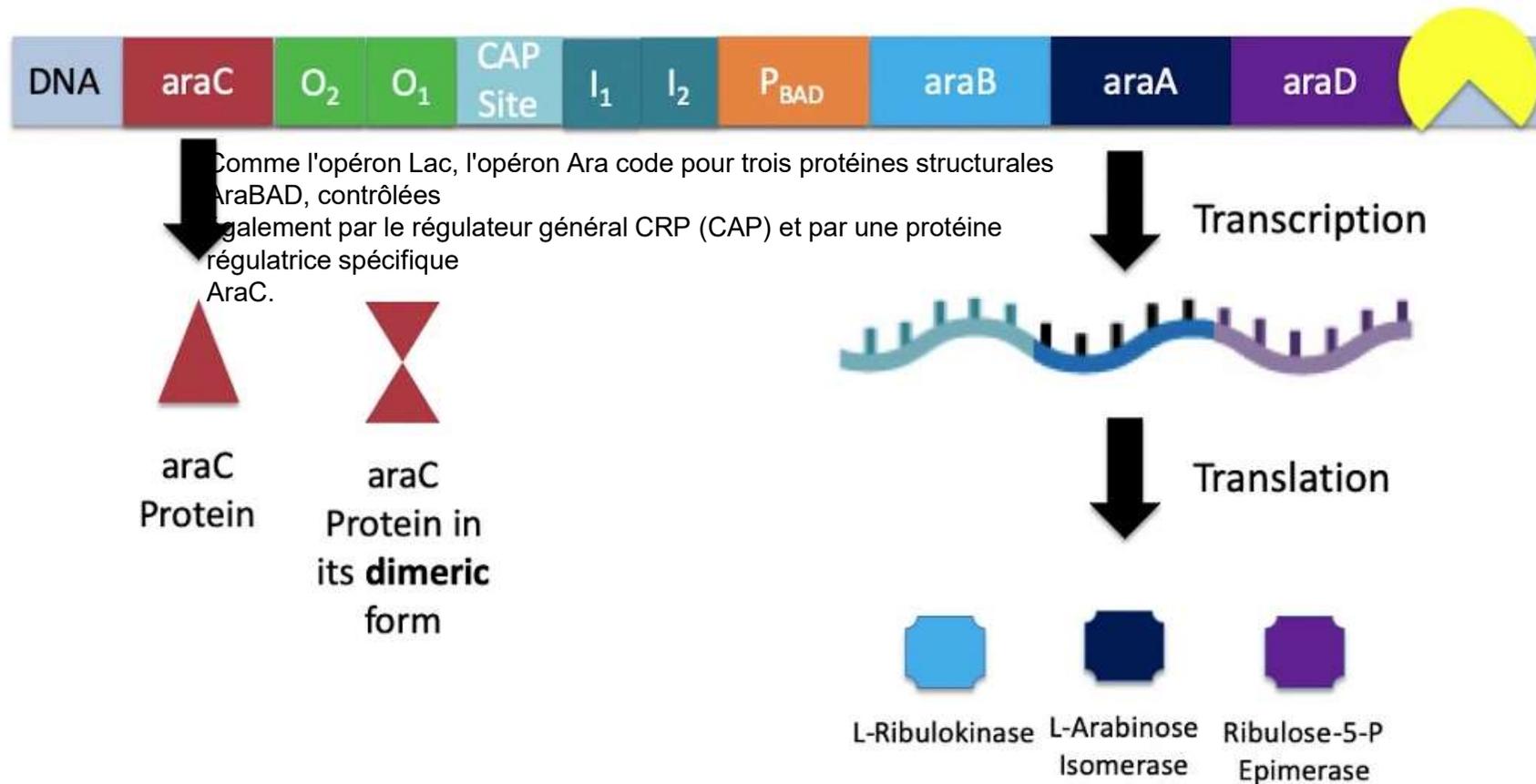


# Promoteur araBp

- pBad (systématiquement araBp) est un promoteur présent chez les bactéries, notamment dans les plasmides utilisés en laboratoire.
- Ce promoteur fait partie de l'opéron arabinose, dont le nom dérive des gènes dont il régule la transcription : araB, araA et araD. Chez E. coli, le promoteur PBAD est adjacent au promoteur PC (systématiquement appelé araCp), qui transcrit le gène araC en sens inverse. araC code pour la protéine AraC, qui régule l'activité des promoteurs PBAD et PC.
- La protéine CAP, récepteur de l'AMP cyclique, se lie entre les promoteurs PBAD et PC, stimulant leur transcription lorsqu'elle est liée à l'AMPc.

# Opéron arabinose

- L'opéron arabinose, ou opéron ara est un opéron servant au transport et au métabolisme de l'arabinose chez *Escherichia coli*



# Opéron arabinose

- *L'opéron arabinose est constitué de :*
  - 3 gènes structuraux :*
    - *araA: L-arabinose isomérase convertissant le L-arabinose en L-ribulose*
    - *araB: ribulokinase responsable du transfert d'un phosphate (P) sur le L-ribulose*
    - *araD: L-ribulose-5-phosphate 4-épimérase convertissant le L-ribulo*

# Opéron arabinose

- Promoteur pBAD : site de fixation de l'ARN polymérase pour la transcription des gènes araB, araA et araD
- Promoteur pC : site de fixation de l'ARN polymérase pour la transcription du gène araC
- Gène régulateur :
  - araC : code la protéine régulatrice AraC possédant un domaine de liaison de l'arabinose. La protéine AraC est impliquée dans la régulation positive et négative de l'opéron ara.
- 3 séquences régulatrices auxquelles se lie la protéine AraC:
  - araO1 : Site de liaison de la protéine AraC pour inhiber sa propre transcription par le promoteur pC.
  - araO2 : Site de liaison de la protéine AraC qui, en absence d'arabinose, permet l'inhibition de la transcription par le promoteur pBAD en formant un complexe avec la protéine AraC liée au site araI.
  - araI1 et 2 : Site de liaison de la protéine AraC qui, en absence d'arabinose, permet l'inhibition de la transcription par le promoteur pBAD en formant un complexe avec la protéine AraC liée au site araO2

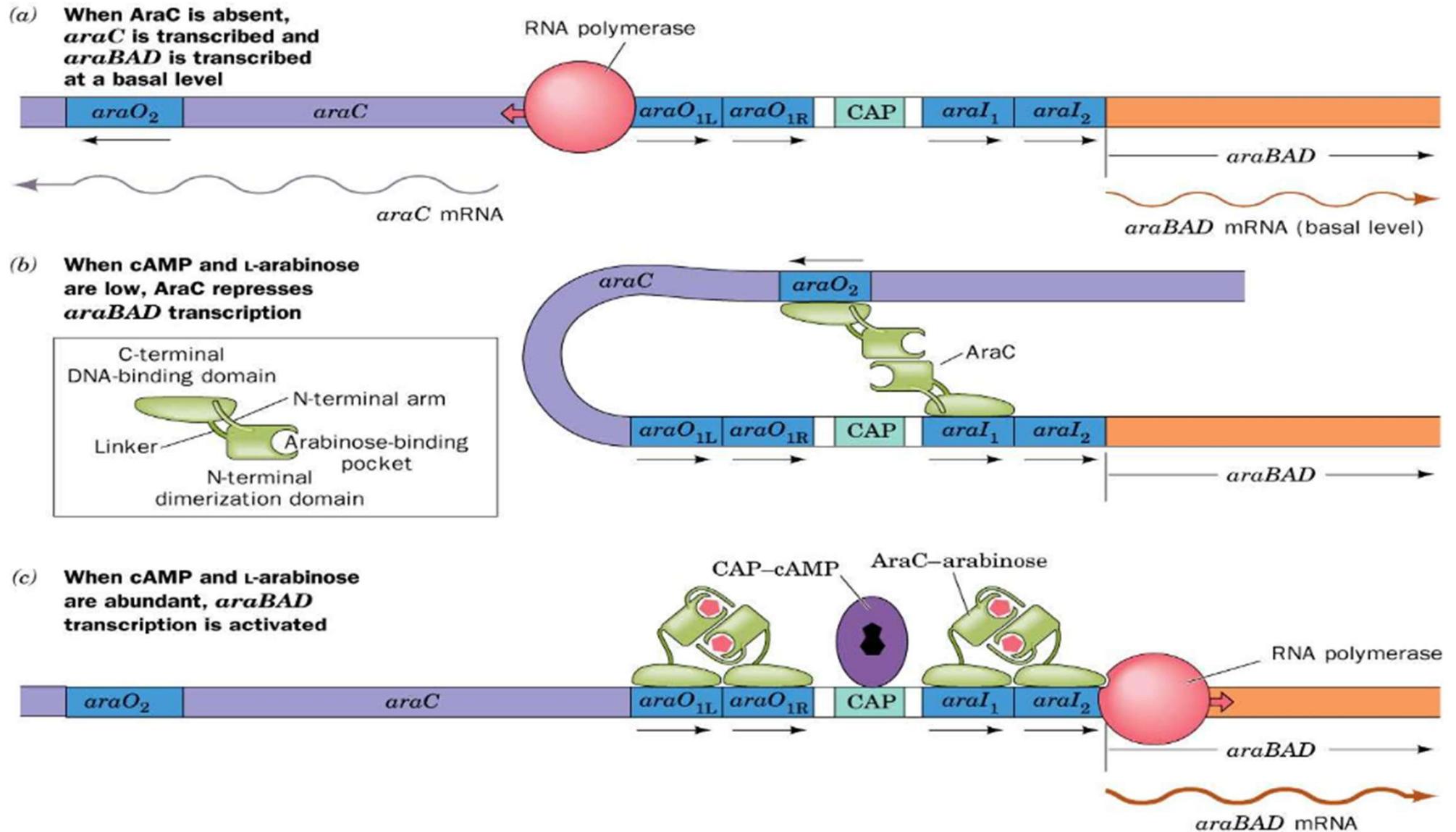
# Opéron arabinose

- **Mécanisme de régulation**
- L'opéron arabinose d'E. coli est régulé par une protéine AraC qui possède la propriété d'agir comme répresseur et activateur selon les conditions. La régulation de cet opéron fait intervenir la protéine AraC ainsi que trois séquences régulatrices différentes : araO2, araO1 et araI.
- **Régulation négative**
  - Lorsque l'arabinose est absent, la protéine AraC agit comme répresseur. La liaison de la protéine AraC au site araO2 et araI permet la répression de l'opéron et inhibe la transcription des gènes araB, araA et araD.
  - Une interaction entre les deux molécules d'AraC est responsable de cette inhibition puisqu'elle engendre une courbure de l'ADN, empêchant la fixation de l'ARN polymérase au promoteur pBAD.
  - En absence d'arabinose, AraC agit donc comme répresseur et inhibe la transcription des gènes de l'opéron.
  - Le taux d'AraC est autorégulé, lorsqu'un excès d'AraC se fixe en

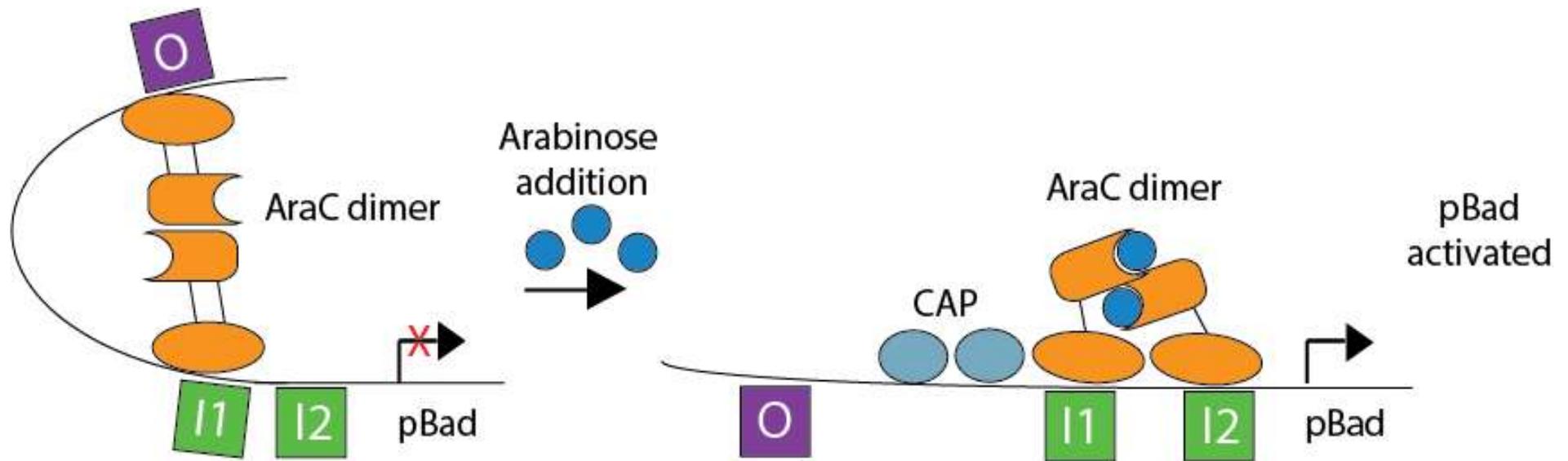
# Opéron arabinose

- **Régulation positive**
- En l'absence d'arabinose, la protéine régulatrice AraC se lie aux sites O et I1 en amont de pBad, bloquant la transcription.
- L'ajout d'arabinose fait que l'AraC se lie aux sites I1 et I2, ce qui permet au début de la transcription.
- En plus de l'arabinose, l'AMPc interagit avec la protéine activatrice de l'AMPc (CAP) peut également stimuler la liaison d'AraC aux sites I1 et I2.
- La supplémentation en glucose du milieu de culture diminue l'AMPc et réprime le promoteur pBad, diminuant ainsi la transcription.

# Opéron arabinose



# Promoteur araBp

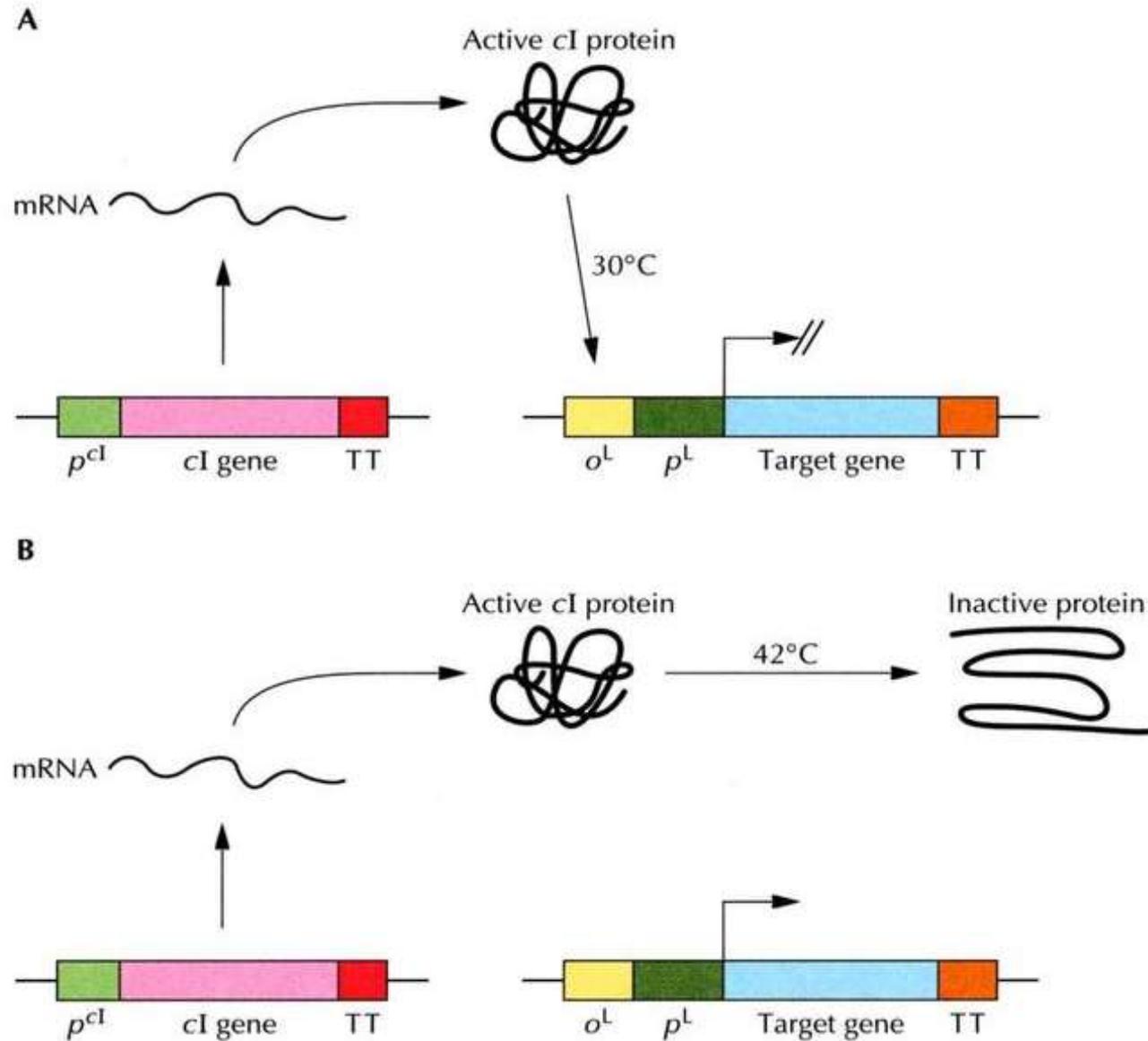


*CAP can also stimulate AraC to bind I1 and I2*

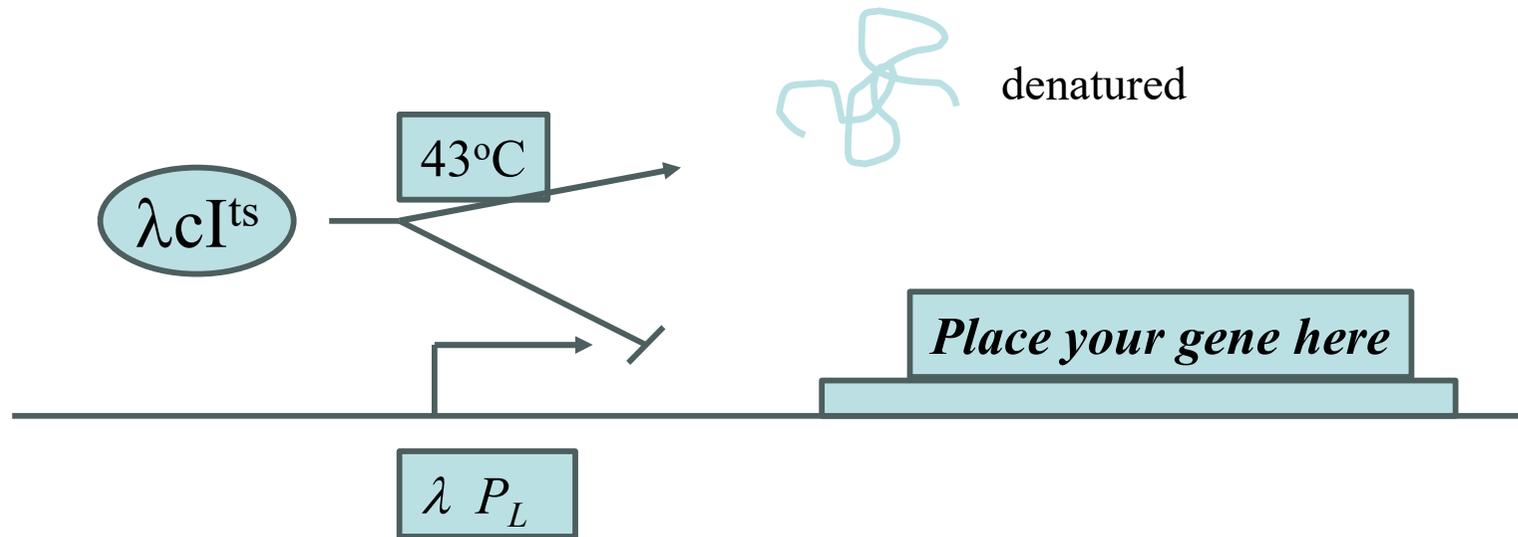
# Promoteur pL

- Promoteur PL du bactériophage  $\lambda$  est induit par un changement de température de  $\sim 30^\circ \text{ C}$  à  $\sim 42^\circ \text{ C}$
- La protéine répressive  $cI$  de phage lambda est sensible à la température qui entraîne sa dénaturation lorsque la température passe de 30 à 42 C, permettant ainsi l'expression du promoteur lambda pL.
- Ce répresseur régule normalement négativement l'expression des gènes des promoteurs pL et pR du bactériophage lambda.
- Cette action répressive est maximale à  $30^\circ \text{ C}$ .
- Cependant, lorsque la température augmente, généralement jusqu'à  $42^\circ \text{ C}$ , la fonctionnalité de la protéine est perdue et le répresseur  $cI$  ne peut plus se lier aux opérateurs de son promoteur.
- Par conséquent, l'expression du promoteur lambda augmente.

# Bacteriophage $\lambda$ $P^L$ Promoter



# Promoteur $\lambda P_L$

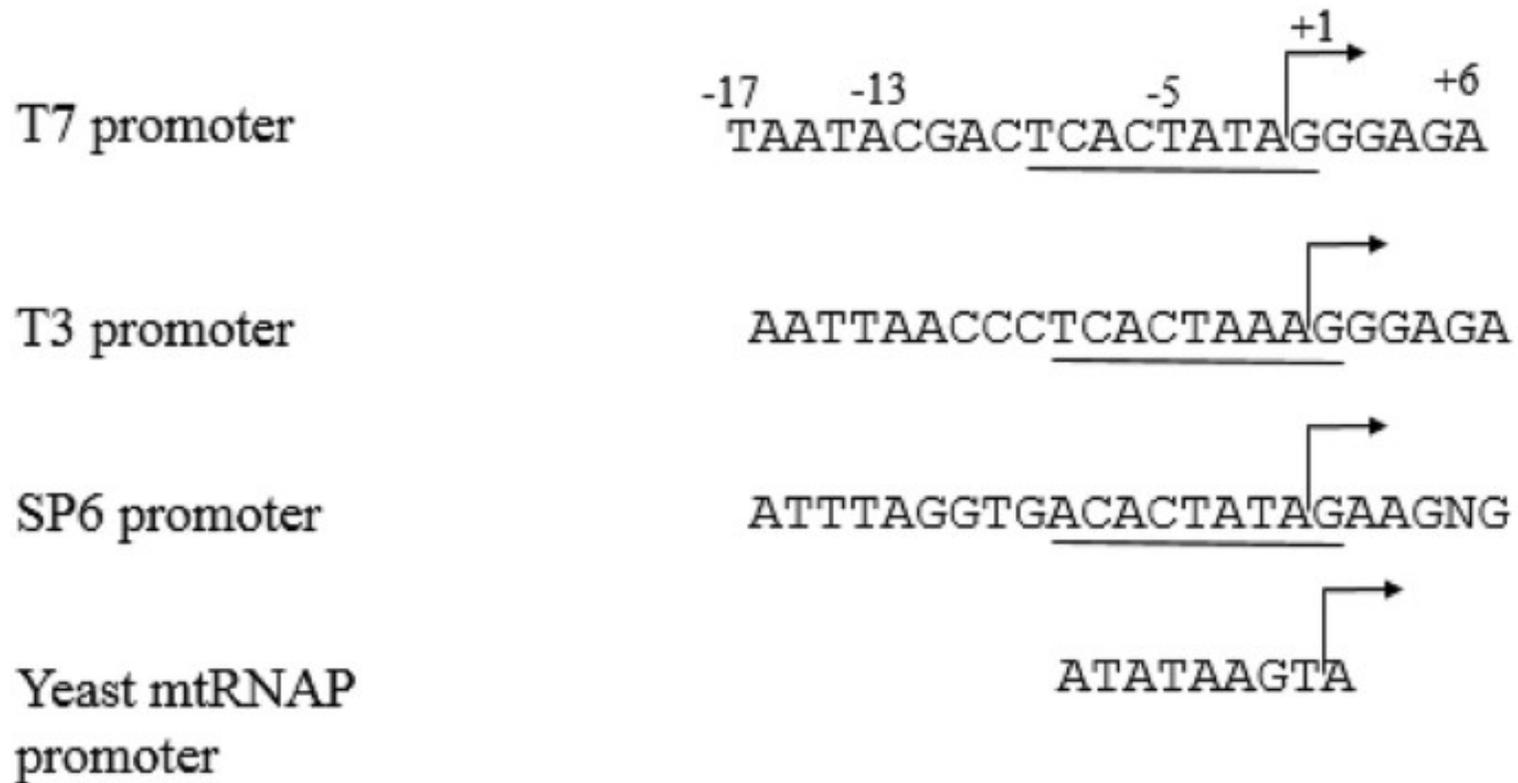


- ◆ « inducteur » n'est pas cher !
- ◆ Difficile d'obtenir des variations rapides de température à grande échelle
- ◆ Niveaux modérés de surexpression
- ◆ Moins perméable que le système lac

# Promoteur T7

- Le promoteur T7 est une séquence d'ADN de 18 paires de bases jusqu'au site d'initiation de la transcription en +1 (5' – TAATACGACTCACTATAG – 3'), reconnue par l'ARN polymérase T7.
- Le promoteur T7 est couramment utilisé pour réguler l'expression génétique des protéines recombinantes, et peut ensuite être utilisé pour diverses applications de recherche en aval.
- Le système d'expression T7 permet une expression de haut niveau du puissant promoteur du bactériophage T7. Il est idéal pour exprimer des protéines recombinantes solubles et non toxiques chez *E. coli*.
- En 2021, ce système avait été décrit dans plus de 220 000 publications de recherche.

# Promoteur T7



# Promoteur T7 ( $p^{T7}$ )

- Localisé dans bactériophage T7
- Induit par l'IPTG
- Souvent utilisé dans les vecteurs pET

# Promoteur T7

- Promoteur T7
  - Transcription par l'ARN polymérase T7
- Gene T7 ARN polymerase
  - Sur le chromosome d'E. coli, sur un lysogène du bactériophage L
  - Sous le contrôle du promoteur lac
  - Induit par l'IPTG

