

République Algérienne Démocratique et Populaire  
Université Frères Mentouri Constantine 1  
Faculté des sciences de la nature et de la vie  
Département de Biochimie et de Biologie Moléculaire et Cellulaire  
Biochimie Option Biochimie appliquée

# *Protéines de fusion*

Ce sont des protéines ou des peptides qui permettent la production, la détection, la stabilité et la purification aisée des protéines recombinantes lorsqu'elles sont ajoutées en fusion à la protéine d'intérêt.

l'expression d'ARN messenger eucaryote dans E. coli se heurte à de nombreux problèmes dus en particulier à la mauvaise initiation de la traduction. Si la protéine de fusion est placée en amont de la protéine d'intérêt, on aura normalement une bonne initiation de la traduction.

La surexpression de protéines dans E. coli s'accompagne souvent de formation de corps d'inclusion dans lesquels la protéine surexprimée est mal repliée. La fusion avec une autre protéine qui se replie correctement aide souvent à avoir une protéine bien repliée.

### **La production**

Ce sont des protéines ou des peptides qui permettent la production, la détection, la stabilité et la purification aisée des protéines recombinantes lorsqu'elles sont ajoutées en fusion à la protéine d'intérêt.

### **La détection**

→ En effet on dispose d'anticorps pour les protéines de fusion mais de plus certaines d'entre elles sont des enzymes qui peuvent être détectées par leur activité.

### **La stabilité**

→ Les petits peptides sont rapidement dégradés chez E. coli. Un des moyens pour éviter cette dégradation est donc de les produire en fusion.

La principale justification de l'utilisation des protéines de fusion est la purification.

Cette technique rend possible la purification de la protéine hétérologue en une seule étape grâce à une chromatographie d'affinité.

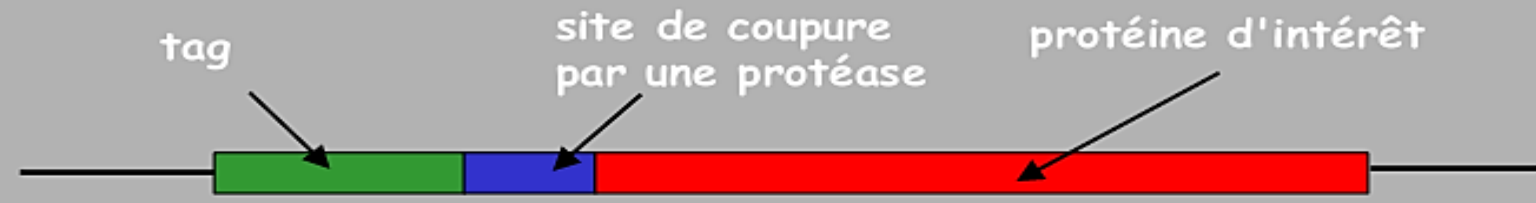
d'autres avantages tels que la protection de la protéine d'intérêt contre les protéases de l'hôte, l'amélioration du taux d'expression et enfin l'augmentation de la solubilité de la protéine désirée en réduisant la formation des corps d'inclusion

**Tags d'affinité**

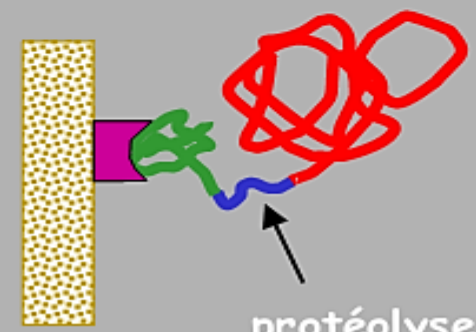
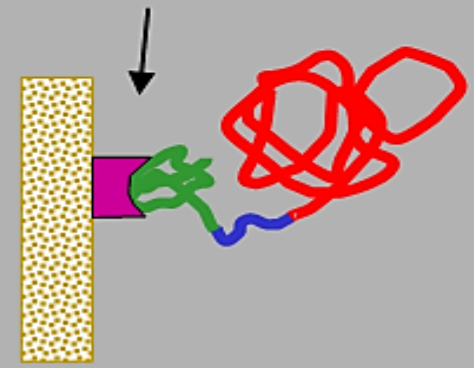
**Tags de coloration**

**Principe**

Construction du gène codant la protéine de fusion par PCR  
ou insertion directe du gène dans un vecteur commercial



chromatographie  
d'affinité



protéine  
d'intérêt  
purifiée



protéolyse  
spécifique

## 1. Tags d'affinité




La technologie des tags de purification se base sur différents types d'interaction :

- ➔ enzyme-substrat,
- ➔ récepteur bactérien-sérum protéique,
- ➔ polyhistidine-ions métalliques
- ➔ anticorps-antigène.

Le choix d'un système de purification dépend de la protéine elle-même et de ses applications futures, ce qui rend impossible l'élaboration d'une procédure de purification générale valide dans tous les cas.

**1. Tags d'affinité**

**Les différents systèmes**

 TAG		 ligand de chromatographie
6 His (polyhistidine) <i>0,84 kDa</i>	_____	métaux chélatés
Glutathion S Transférase (GST) <i>26 kDa</i>	_____	glutathion
Maltose Binding Protein (MBP) <i>40 kDa</i>	_____	amylose
Calmodulin Binding Peptide (CBP) <i>4 kDa</i>	_____	calmoduline
Chitine Binding Domain (CBD) <i>5 kDa</i>	_____	chitine
Protéine A	_____	IgG
Thiorédoxine	_____	oxyde de phénylarsine

## 1. Tags d'affinité

### Caractéristiques

Les tags d'affinité permettent une purification rapide de complexes protéiques en une seule étape, en conditions natives ou dénaturantes et sans avoir besoin de connaître la composition ou la fonction de ces protéines.

Le degré final de purification est très élevé (90 à 99 %) avec un bon rendement de protéines purifiées

Les tags d'affinité sont largement utilisés pour la purification et la détection des protéines recombinantes, particulièrement à partir de mélanges complexes comme les lysats cellulaires.

Ils peuvent aussi bien être fusionnés en N comme en C terminal dans plusieurs systèmes d'expression de protéines.



**1. Tags d'affinité****Matrices et conditions d'élution de quelques tags d'affinité**

<b>Tag d'affinité</b>	<b>Matrice</b>	<b>Condition d'élution</b>
Poly-Arg	Résine échangeuse de cation (DEAE)	Gradient de NaCl de 0 à 400 mM à pH >8,0
Poly-His	Ni <sup>2+</sup> -NTA, Co <sup>2+</sup> -CMA (Talon)	20-250 mM d'Imidazole ou pH bas
FLAG	Anticorps monoclonal Anti-FLAG	pH 3,0 ou 2-5 mM EDTA
Streptag II	Strep-Tactin (streptavidine modifiée)	2,5 mM desthiobiotine
c-myc	Anticorps monoclonal anti-c-myc	Low pH
S <sup>-</sup>	Fragment S de la RNaseA	3 M guanidine thiocyanate 0,2 M citrate pH 2, 3 M magnésium chloride
HAT (natural histidine affinity tag)	Co <sup>2+</sup> -CMA (Talon)	150 mM imidazole ou pH bas
Cellulose-binding domain	Cellulose	Famille I : guanidine Hcl ou urea > 4 M Famille II / III : éthylène glycol
SBP	Streptavidine	2 mM Biotine
Glutathione S-transférase	Glutathion	5 – 10 mM glutathion réduit
Maltose-binding protein	Amylose	10 mM maltose

## 1. Tags d'affinité

Il existe deux approches d'utilisation des tags d'affinité.

### La première approche

consiste à utiliser des tags de petite taille, qui peuvent être facilement fusionnés à la protéine d'intérêt par le biais d'amorces et qui n'interfèrent pas avec la structure et la fonction de la protéine d'intérêt.



Dans certains cas, le clivage de ces partenaires de fusion n'est pas nécessaire. Cependant, leur élimination est obligatoire pour des applications en cristallisation, en production d'anticorps ou encore lorsqu'ils affectent la structure tertiaire et/ou l'activité biologique de la protéine

## 1. Tags d'affinité

Il existe deux approches d'utilisation des tags d'affinité.

### La seconde approche

consiste à employer un peptide de plus grande taille ou une protéine comme partenaire de fusion.



L'avantage de cette méthode est l'augmentation de la solubilité des protéines de fusion

mais dans ce cas et pour différentes applications de la protéine d'intérêt,  
le ligand doit obligatoirement être éliminé

**1. Tags d'affinité**

**Séquences et tailles de quelques tags d'affinité**

Tag	Résidus	Séquence	Taille (KDa)
Poly-Arg	5-6	RRRRR (En général 5)	0,80
Poly-His	2-10	HHHHHH (En général 6)	0,84
FLAG	8	DYKDDDDK	1,01
Streptag II	8	WSHPQFEK	1,06
c-myc	11	EQKLISEEDL	1,20
S-	15	KETAAAKFERQHMDs	1,75
HAT-	19	KDHLIHNVHKEFHAAHANK	2,31
3x FLAG	22	DYKDHDGDYKDHDIDYKDDDDK	2,73
CBDs	27-189	Domaine	3 - 20
SBP	38	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP	4,03
GST	211	Protéine	26,00
MBP	396	Protéine	40,00

La première approche

La seconde approche

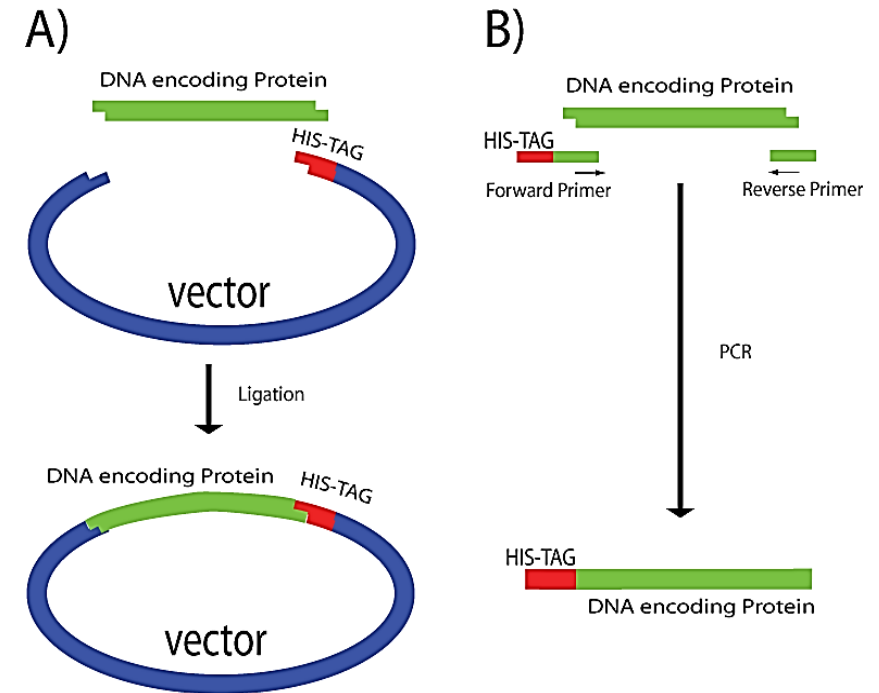
## 1. Tags d'affinité

### Tag polyhistidine

Le polyhistidine-tag, une répétition de 6 à 10 histidines, est le tag le plus utilisé pour la purification de protéines recombinantes.

Cette méthode a été décrite pour la première fois en 1987

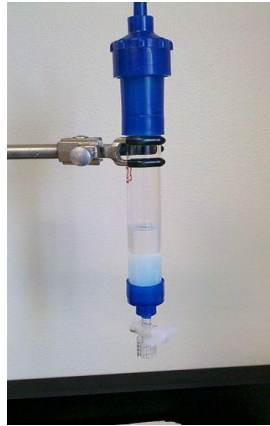
Actuellement, plus de 60% de protéines produites pour étude structurale sont taguées de cette manière



## 1. Tags d'affinité

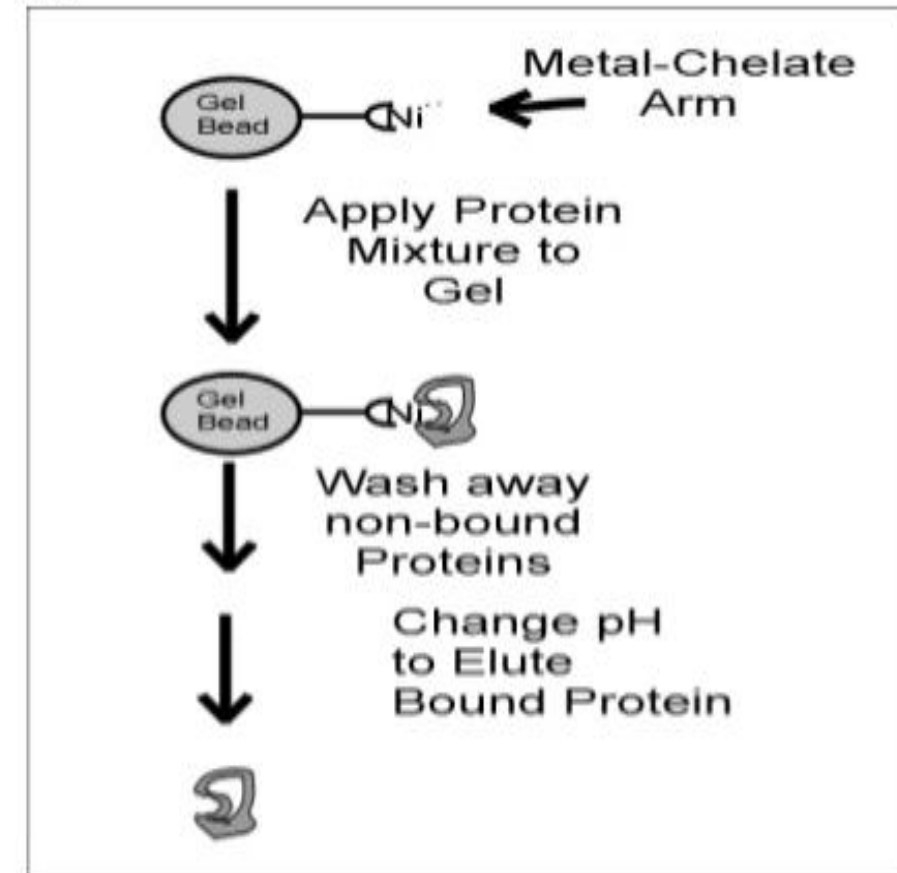
### Tag polyhistidine

L'IMAC est basée sur l'interaction entre certains acides aminés exposés à la surface des protéines, particulièrement l'histidine, et les ions métalliques de transition qui sont immobilisés en utilisant un agent chélateur.



Les protéines étiquetées sont efficacement retenues sur la colonne IMAC à pH neutre ou légèrement alcalin, puis sont éluées par une solution d'imidazole à 250 mM

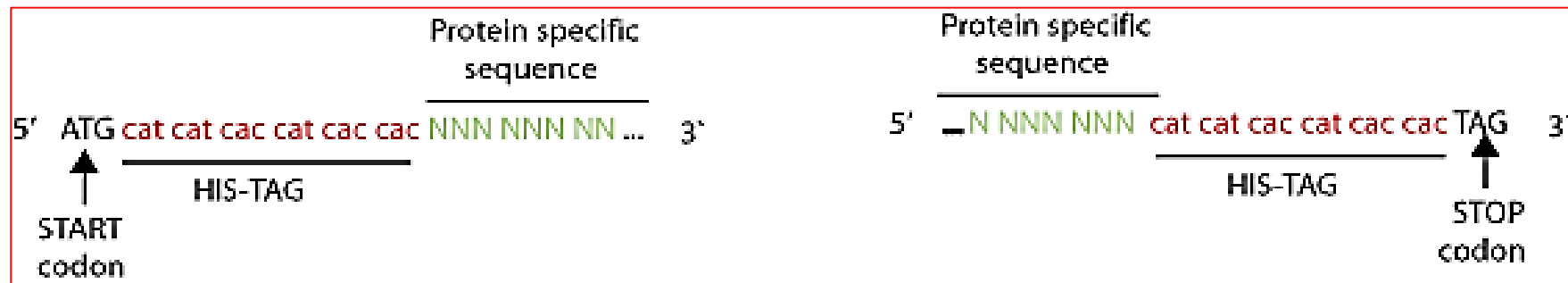
### Immobilized Metal Affinity Chromatography



## 1. Tags d'affinité

### Tag polyhistidine

Les tags d'affinité poly-histidine sont communément fusionnés en N ou C terminal des protéines d'intérêt, et l'emplacement optimal diffère d'une protéine à l'autre



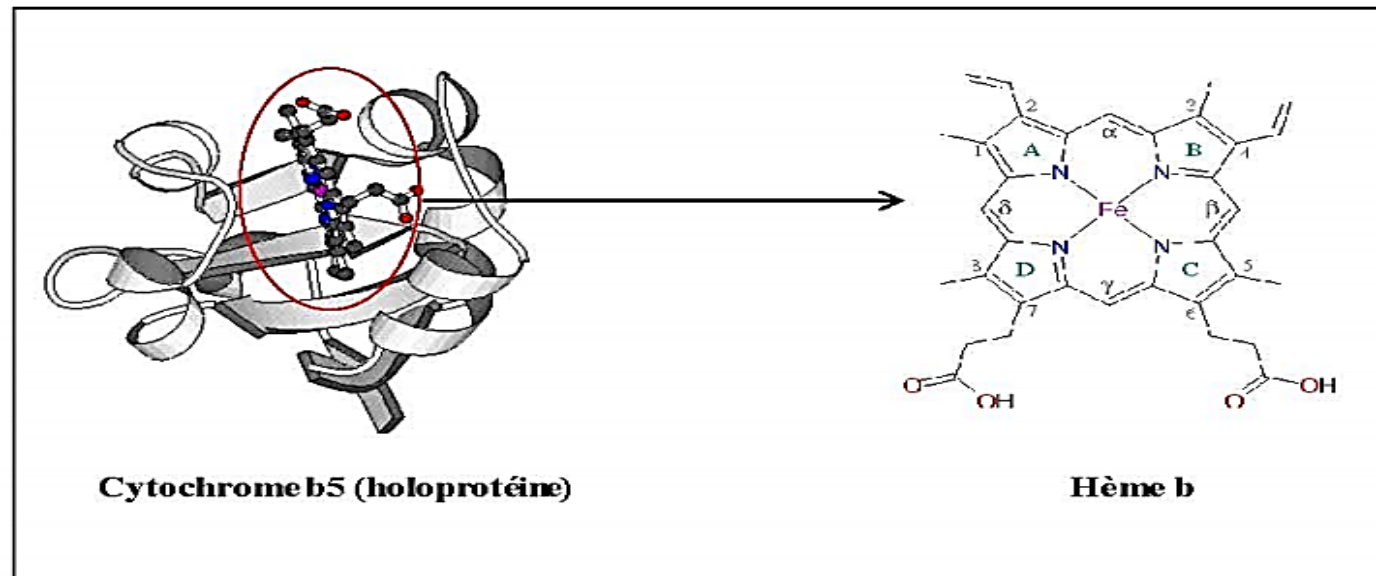
L'utilisation du tag poly-His a été décrite dans un grand nombre de systèmes d'expression : la bactérie (Krahulec et al. 2010; Liu and Pan 2010), la levure (Luo et al. 2008; Jin et al. 2009), les cellules mammifères (Reinke et al. 2009) et les cellules d'insectes infectées par le baculovirus (Michel et al. 2008; Zhou et al. 2010).

**2. Tags de coloration**

**Domaine de fixation de l'hème du cytochrome b5**

Le cytochrome b5 peut être exprimé sous deux formes :

- la forme apoprotéique qui est incolore
- la forme holoprotéique rouge. L'holoprotéine acquiert cette couleur en fixant un cofacteur « hème b » sur le domaine de fixation de l'hème qui ne dépasse pas 15 kDa.



**Structure de la forme holoprotéique du cytochrome b5**



## 2. Tags de coloration

### **Domaine de fixation de l'hème du cytochrome b5**

l'efficacité de ce domaine dans la détection de l'expression de la protéine GST, où il permet de détecter à l'œil nu jusqu'à 0,1 mg de protéine de fusion cytb5-GST par mL

le cytochrome b5 présente une solubilité de 95%. Il est ainsi hautement exprimé sous forme soluble chez *E. coli*. Ces caractéristiques lui permettent de satisfaire tous les critères d'un bon partenaire de fusion.

En plus des caractéristiques de coloration et de solubilité, le cytochrome b5 permet la purification de la protéine de fusion par chromatographie d'affinité aux métaux lourds en une seule étape. En effet, le domaine de fixation de l'hème du cytochrome b5 possède dans sa séquence six histidines.

## **2. Tags de coloration**

### **Domaine de fixation de l'hème du cytochrome b5**

La nature extrêmement acide du cytochrome b5 permet également sa purification par chromatographie échangeuse d'anions comme la chromatographie DEAE-sépharose

En résumé, les avantages de l'utilisation de cytochrome b5 résident dans sa couleur, son haut niveau d'expression, sa haute solubilité, sa stabilité, sa facilité de purification et sa taille réduite.

## 2. Tags de coloration

### **Domaine FMN (« Flavine mononucléotide ») du cytochrome p450**

Au niveau de la structure de la flavoprotéine cytochrome p450 réductase, le domaine FMN joue le rôle de coenzyme qui assure le transfert des électrons lors des réactions catalysées par cette enzyme. Il se fixe sur la flavoprotéine grâce au domaine de fixation du FMN et lui confère la coloration.

Ainsi ce domaine de 21 kDa présente un grand intérêt dans le suivi de la production et de la purification de protéines recombinantes. Finn et al ont montré la possibilité d'utilisation de ce domaine en tant que marqueur et ont discuté ses avantages de surexpression sous forme soluble chez *E. coli* (Finn et al. 2005). Des niveaux d'expression de 30-40 mg/L ont ainsi été décrits.

## 2. Tags de coloration

### **Domaine FMN (« Flavine mononucléotide ») du cytochrome p450**

En comparant ce marqueur protéique au domaine de fixation de l'hème du cytochrome b5, ils présentent tous les deux les avantages de surexpression, de solubilité et de faible poids moléculaire.

Mais au niveau de la coloration, le domaine FMN est 10 fois moins sensible que celui de cytochrome b5. De plus le coenzyme FMN n'est pas stable, il change de couleur suivant son état d'oxydation où il peut être verdâtre ou jaune

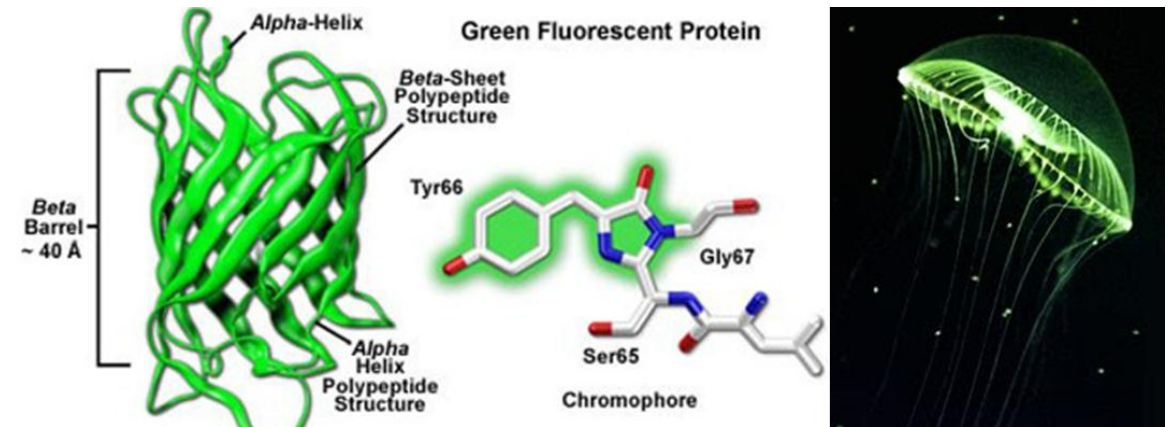
## 2. Tags de coloration

## Protéines fluorescentes

### La protéine GFP (« Green Fluorescent Protein »)

La **GFP** est une protéine autofluorescente de 30 Kda, découverte et isolée pour la première fois en 1962 à partir de la méduse *Aequorea victoria*,

elle émet de la lumière verte lorsqu'elle est excitée en UV



En 1998, Albano et al ont montré pour la première fois l'efficacité de la GFP dans le développement de bioprocédés de contrôle de la production de protéines recombinantes.

ils ont produit la protéine CAT (chloramphénicol acétyl transférase) en fusion avec la GFP et ont montré que cette dernière était plus fiable pour la quantification de la CAT que l'activité de la CAT elle-même. De plus, la méthode de fluorescence était dix fois plus rapide que la méthode enzymatique

**2. Tags de coloration**

**Protéines fluorescentes**

**La protéine GFP (« Green Fluorescent Protein »)**

Différents variants plus efficace ou émettant à différentes longueur d'onde ont été générés par mutagenèse.

- **GFPuv** se replie plus facilement que la GFP et a un rendement de fluorescence 16 fois plus importante que la GFP
- **GFPmut1** contient deux mutations (F64L et S65T) son rendement de fluorescence est 35 fois supérieur à celui de la GFP
- **GFP+** combine les mutations de GFPuv et de GFPmut1 ce qui lui donne une augmentation de 130 fois par rapport à la GFP
- **L'EGFP** est un variant optimisée pour l'expression en cellules eucaryotes.

## 2. Tags de coloration

### Autres Protéines fluorescentes

La **HcRed** a été générée par mutagenèse à partir d'une protéine non fluorescente du corail, *Heteractis crispa*.

Elle émet dans le rouge.

Ces protéines gardent leurs propriétés de fluorescence en fusion N ou C terminale.

En fusion, elle permet donc la détection et la localisation de la protéine dans les cellules.



*Heteractis crispa*

## 2. Tags de coloration

### Autres Protéines fluorescentes

**La DsRed** a été isolée de *Discosoma* sp.

elle peut être excitée à 488 nm comme la GFP mais émet dans le rouge à 583 nm (la GFP émet à 500 nm).

Son expression a été améliorée par optimisation des codons pour donner la DsRed1.

Elle présentait un défaut important, sa maturation prenait trop longtemps elle avait une demie vie supérieure à 24 heures.

La vitesse de maturation a été améliorée et la DsRedT4 ainsi obtenue a une demie vie de maturation inférieure à 45 minutes



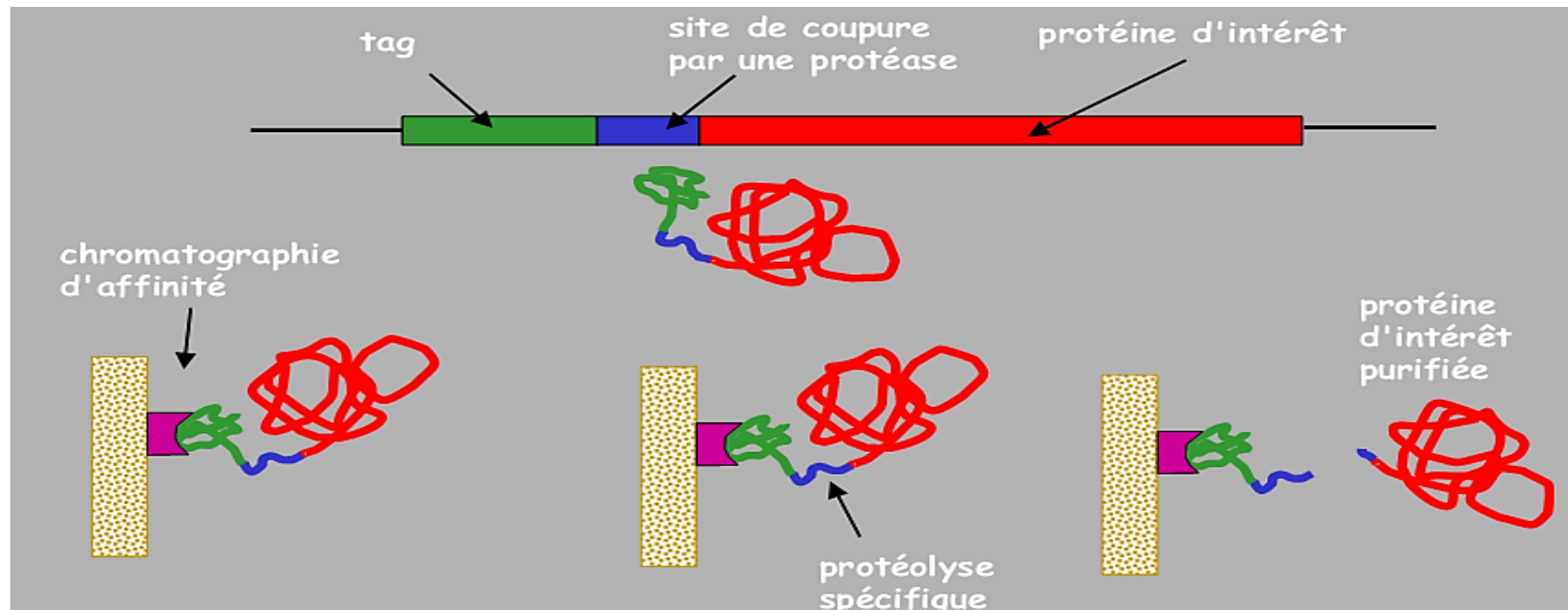
*Discosoma* sp.



## Clivage des protéines de fusion

Souvent on veut enlever des fusions entre deux protéines après une expression in vitro et purification.

Dans ce cas, on intercale entre les deux protéines une séquence qui va coder pour un site de coupure.



**Clivage des protéines de fusion**

L'optimisation de cette étape a préoccupé diverses équipes de recherche qui ont développé plusieurs méthodes de clivage

**Clivage chimique**

**Self-clivage**

**Protéases spécifiques**

**Exopeptidases**

## **Clivage des protéines de fusion**

### **Clivage chimique**

Le clivage de protéines de fusion par voie chimique est devenu possible grâce à la reconnaissance et l'action de certains produits chimiques au niveau de séquences de un à deux résidus.

Le bromure de cyanogène, l'hydroxylamine et l'acide formique sont les réactifs les plus utilisés

- ➔ Le bromure de cyanogène reconnaît les méthionines et coupe au niveau de leur extrémité C-terminale
- ➔ l'hydroxylamine et l'acide formique rompent respectivement les liaisons asparagine/glycine et aspartate/proline

**Clivage des protéines de fusion**

**Clivage chimique**

le clivage chimique se limite à quelques cas de peptides et de protéines de petite taille.

Il nécessite une étude préalable de la séquence de la protéine de fusion

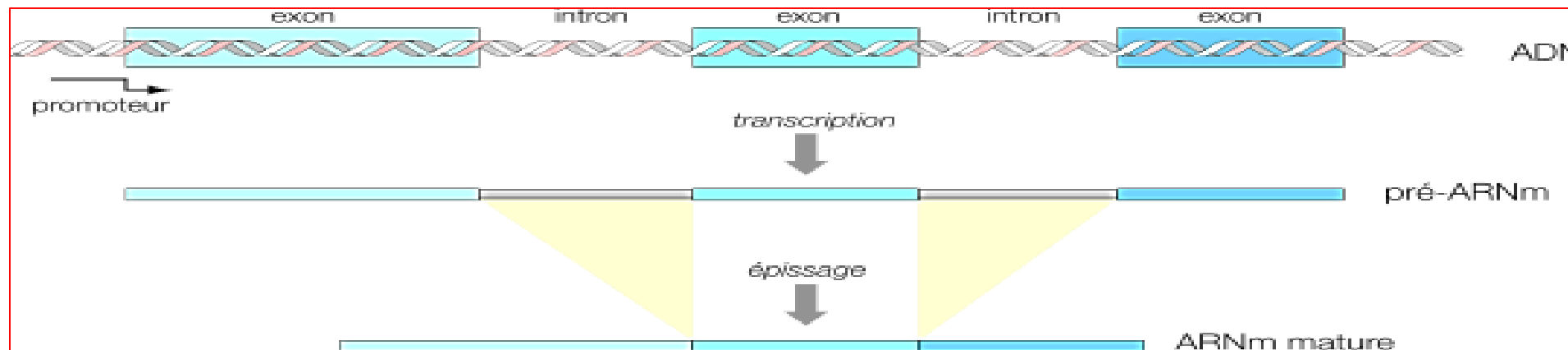
et aussi des caractéristiques physico-chimiques des deux partenaires de fusion.

## Clivage des protéines de fusion

### Self-clivage

Naturellement, l'épissage protéique est un processus auto-catalytique permettant la maturation de certaines protéines.

Lors de l'épissage de ces protéines, les segments internes dans le produit de traduction (intéines) s'excisent eux-mêmes de la protéine pendant que les polypeptides qui les encadrent (extéines) se réunissent pour former la protéine finale

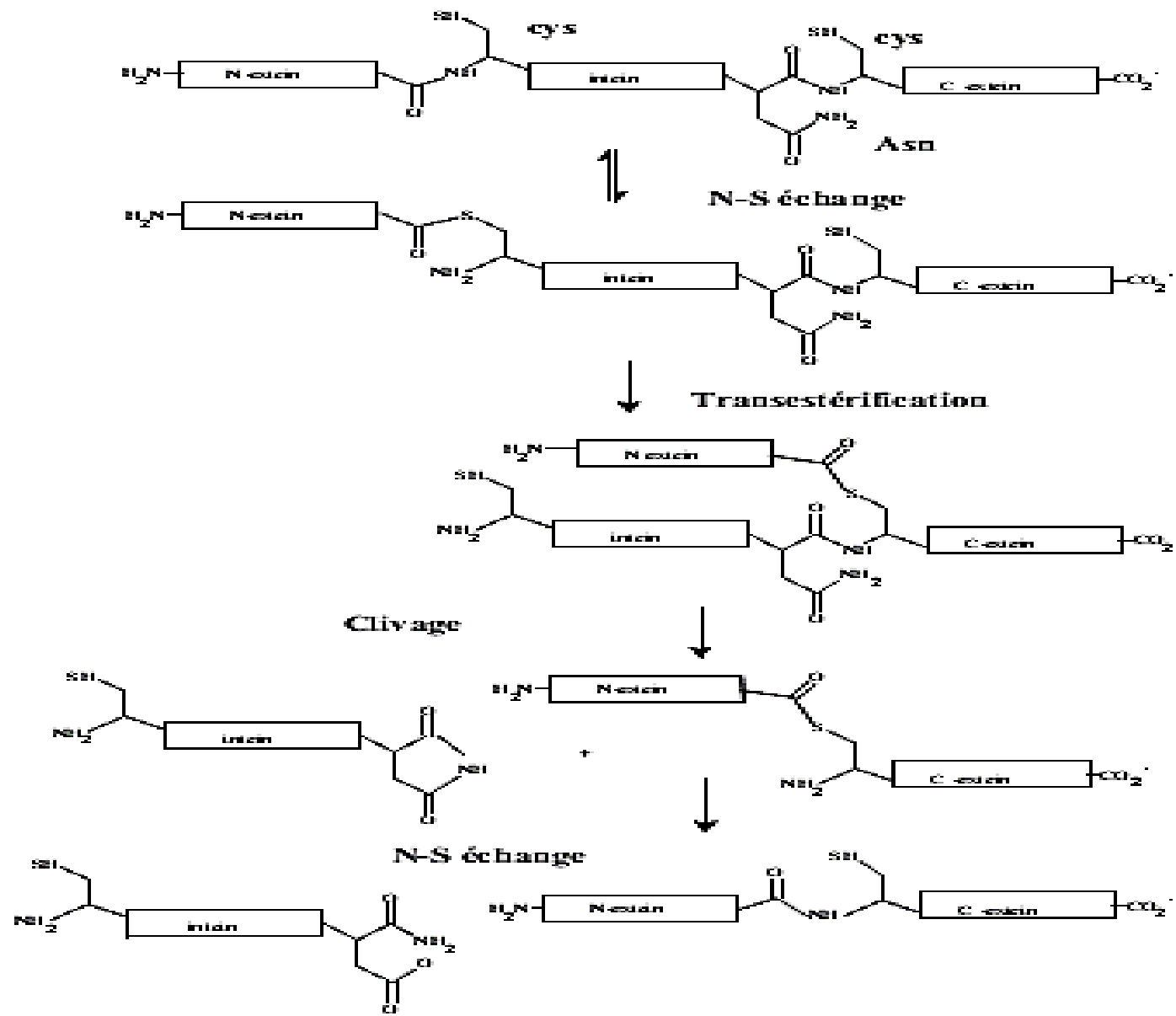


**Clivage des protéines de fusion**

**Self-clivage**

L'application biotechnologique d'une protéine intéine au niveau du clivage de protéines de fusion nécessite sa transformation d'une séquence d'auto-épissage à un simple élément d'auto-clivage en atténuant son potentiel de ligation.

Cette modification peut être effectuée en mutant l'extrémité N ou C-terminale de la jonction d'épissage avant son insertion entre les deux partenaires de fusion afin d'induire le processus d'auto-clivage dans des conditions contrôlables



Exemple de mécanisme d'épissage d'une intréine.

**Clivage des protéines de fusion**

**Self-clivage**

En fonction du mode de contrôle de leur réaction d'auto-clivage, les protéines intéines peuvent être répertoriées en deux catégories :

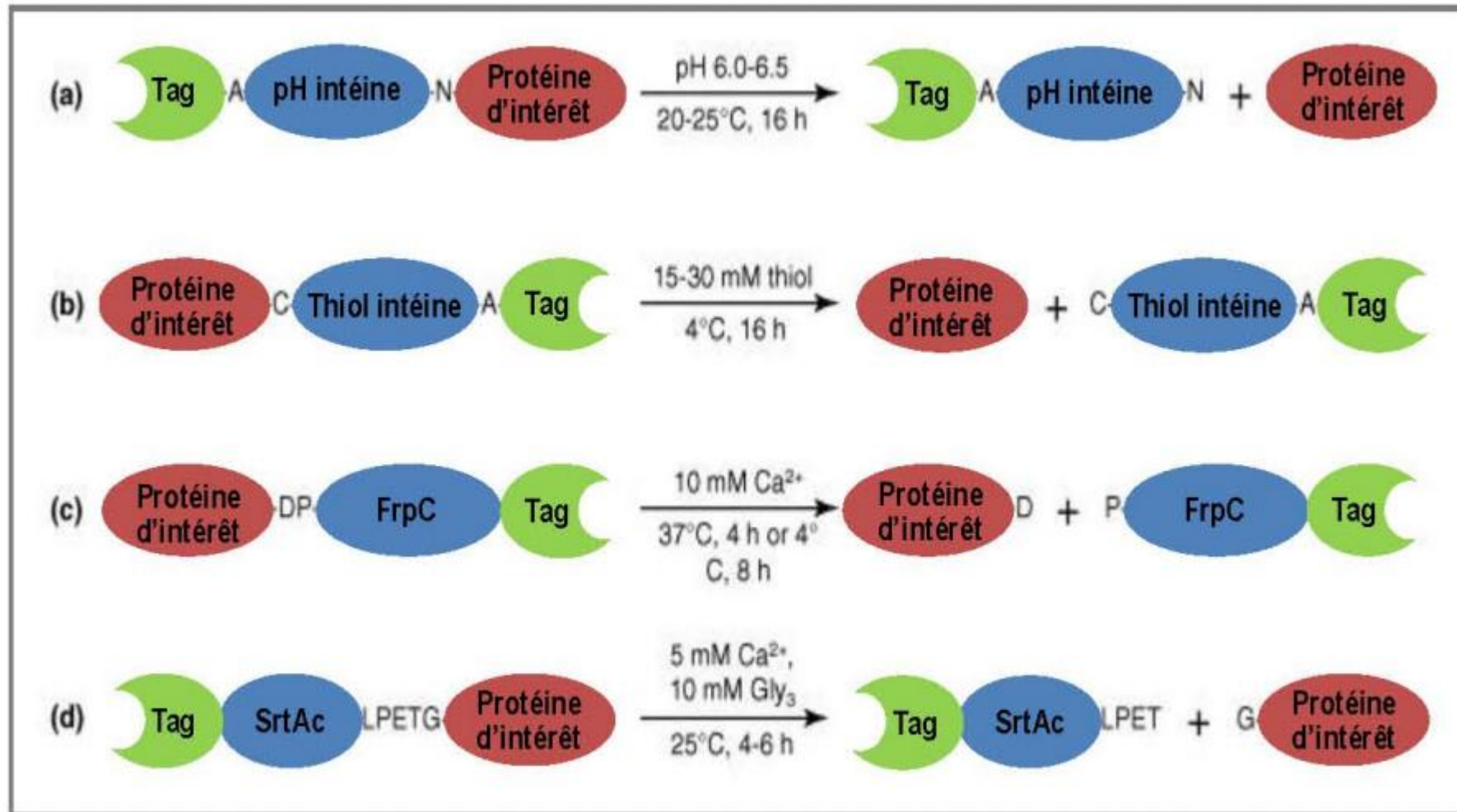
Les intéines inductibles au pH s'auto-clivent vers leur extrémité C-terminale

des intéines inductibles au thiol, qui présente les deux types de coupure N et C-terminale.



**Clivage des protéines de fusion**

**Self-clivage**



## **Clivage des protéines de fusion**

### **Protéases spécifiques**

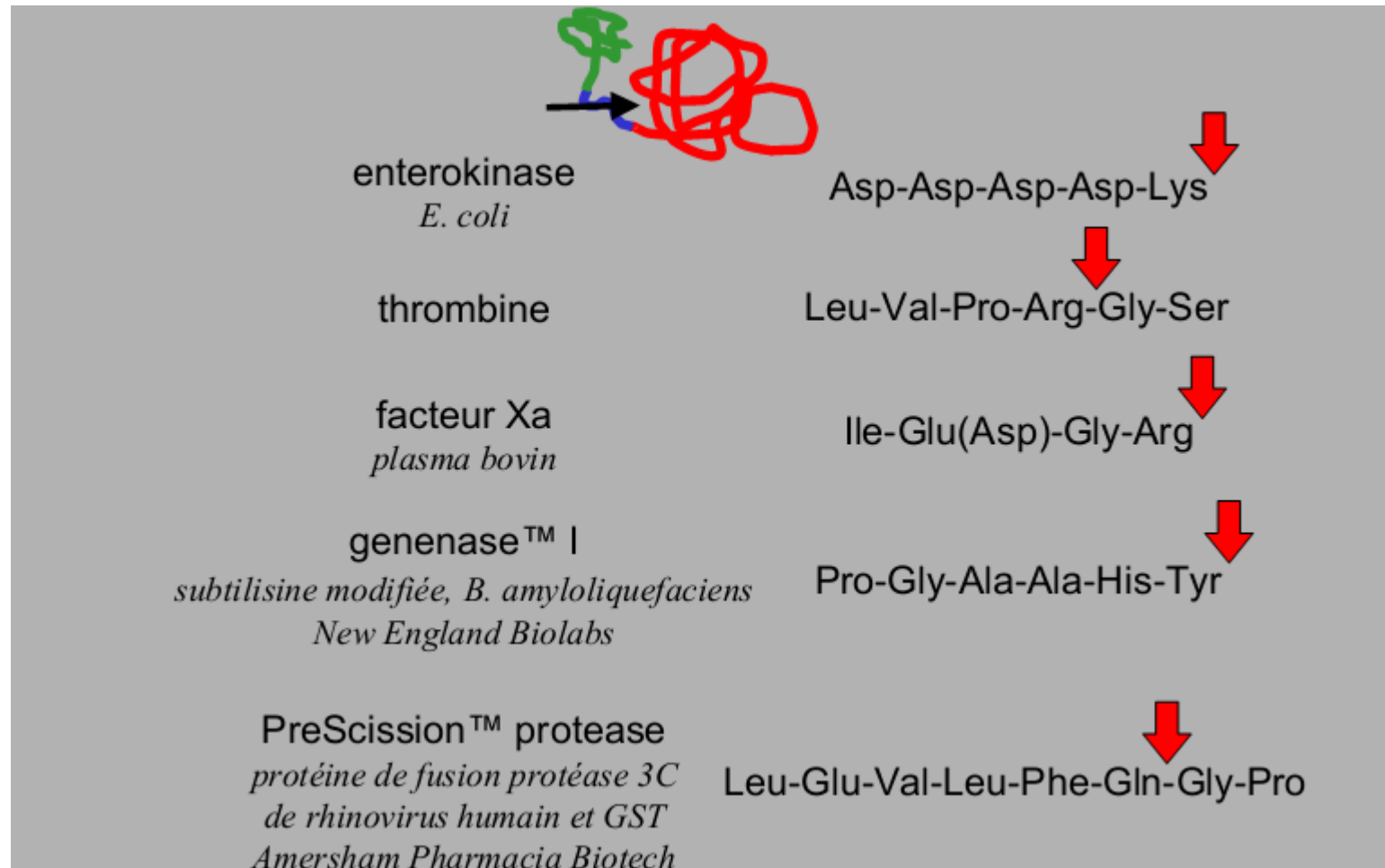
On peut utiliser une méthode enzymatique, des protéases spécifiques d'une séquence.

- **L'enterokinase** reconnaît et coupe la séquence Asp-Asp-Asp-Asp-Lys/
- **La thrombine** reconnaît et coupe la séquence Leu-Val-Pro-Arg-/-Gly-Ser. Selon le positionnement de la fusion à l'extrémité N ou en C terminale, on obtiendra deux ou quatre acide aminés.
- **Le facteur Xa** coupe en C-terminal de la séquence de reconnaissance Ile-Glu-Gly-Arg/
- **Les TEV protéinase** (Tobacco Etch Virus Proteinase) reconnaît la séquence Glu-X-X-Tyr-Gln/ avec X symbolisant n'importe quel acide aminé. Cette protéase a l'avantage d'être très spécifique de son site de coupure et de ne pas être inhibée par les inhibiteurs de protéase à sérine.

**Clivage des protéines de fusion**

**Protéases spécifiques**

**Les protéases spécifiques**



## **Clivage des protéines de fusion**

### **Protéases spécifiques**

Il y a plusieurs désavantages à l'utilisation de protéases.

- la coupure n'est pas toujours spécifique et on peut couper la protéine d'intérêt.
- La température assez haute nécessaire à la coupure peut affecter la stabilité de la protéine d'intérêt.
- La coupure est souvent inefficace du fait de l'inaccessibilité du site de coupure dans la protéine de fusion.
- Après la coupure, il faut retirer la protéine de fusion, qui va s'accrocher sur la colonne d'affinité ayant permis la purification. La protéase est retirée par exemple en utilisant une protéase biotinylée qui va s'accrocher sur une colonne de streptavidine agarose.

## Exemple : Purification de l'AS amylosaccharase

Purification par chromatographie d'affinité sur résine *Glutathione-Sepharose 4B*

