

Production d'une protéine recombinante

Lic. B. PIRSON
Lic. J-M. SERONT
ISICHt - Mons

Production de la protéine recombinante GFP (Green Fluorescent Protein d'Aequoria victoria) par une bactérie (E. coli) contenant le plasmide pGlo. Mise au point d'un protocole pour la production, l'extraction et la purification de la protéine recombinante.

Mots-clés : GFP, pGlo, protéine recombinante.

Production of recombinant protein GFP (Fluorescent Green Protein of Aequoria Victoria) by a bacteria (E. coli) transfected by the pGlo plasmid. Development of a protocol for the production, the extraction and the purification of recombinant protein.

Key words: GFP, pGlo, recombinant protein.

1. Introduction

Actuellement, la production de protéines recombinantes grâce à l'utilisation de bactéries et plasmides est une des techniques les plus utilisées. Pour rentabiliser au maximum cette production, il est nécessaire d'optimiser les conditions de production, d'extraction et de purification de la protéine d'intérêt.

2. Objectifs généraux

Le premier objectif poursuivi est de simuler le fonctionnement d'un « team » de laboratoire de recherche. Les enseignants choisissent le sujet de recherche afin que, grâce à celui-ci, une majorité des techniques fondamentales de la biologie moléculaire puissent être expérimentées.

Le sujet est présenté aux étudiants à partir d'un article de base. Ceux-ci complètent ensuite les informations en consultant la littérature existante sur le thème proposé via les sources classiques (bibliothèques) ou Internet. Une première discussion enseignants-étudiants permet de choisir le matériel ainsi que les techniques adaptées au sujet de recherche ("en ligne" avec le matériel dont dispose le laboratoire).

Les étudiants construisent leur propre plan expérimental en tenant compte des limites de temps imposées (6 semaines à raison de 3 demi-journées par semaine). Chaque semaine, une séance d'information réunit l'ensemble du groupe et les enseignants. Les idées et stratégies sont discutées. Les expériences sont soigneusement consignées dans un cahier de laboratoire. Une évaluation de l'état d'avancement du projet est soumise à l'appréciation critique du groupe et de nouvelles orientations sont proposées.

Le second objectif confronte les étudiants avec la technologie de production d'une protéine recombinante.

Le projet est conclu par un rapport final, une présentation orale par le groupe d'étudiants et éventuellement par la réalisation d'un poster.

Cette année académique 2005-2006, notre choix s'est porté sur la production de la protéine GFP par la bactérie *Escherichia coli*, protéine habituellement produite par *Aequoria victoria* (méduse). En effet, cette protéine GFP a la propriété de produire de la fluorescence, ce qui facilite sa mise en évidence.

3. Etat de la question

3.1 La protéine GFP

Initialement identifié chez la méduse *Aequoria victoria*, le gène *gfp* (*green fluorescent protein*) a été cloné et utilisé comme gène rapporteur.

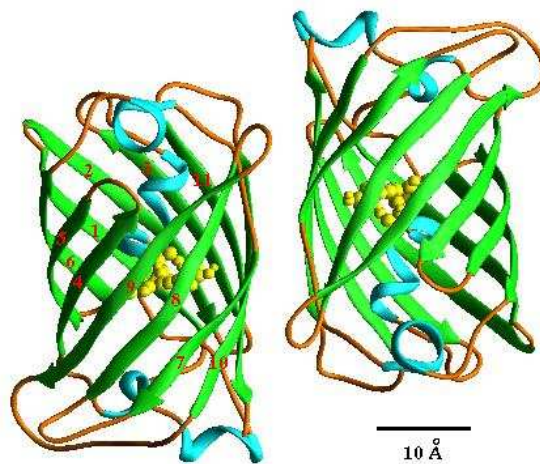


Figure 1 : Structure cristallographique de la protéine GFP.

La protéine GFP a une masse moléculaire de 27000 et possède une structure remarquable : le fluorophore, composé d'une séquence de trois acides aminés sérine-tyrosine-glycine, est enchâssé dans une structure cylindrique de 300 nm de diamètre et 400 nm de hauteur. Cette structure est formée de 11 feuillets β et d'une hélice α portant le fluorophore. Les feuillets β sont disposés en parallèle et forment une cage dans laquelle est enfermée l'hélice α qui contient le chromophore. Les boucles entre les feuillets β forment le plancher de la cage.

Mécanisme de fluorescence : la protéine GFP joue le rôle d'accepteur dans un mécanisme de transfert d'énergie par voie lumineuse. Chez *Aequoria victoria*, la source de lumière provient de l'aécorine, une protéine activée à partir de fluctuations du taux de calcium intracellulaire. L'aécorine émet

une lumière bleue (470 nm), absorbée par la protéine GFP qui elle-même réémet à une longueur d'onde plus élevée (508 nm).

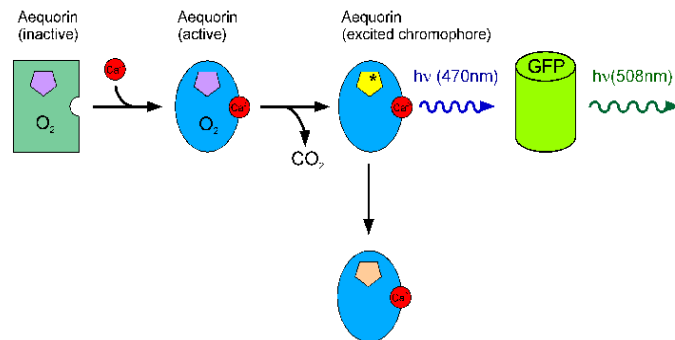


Figure 2 : Mécanisme du processus de fluorescence naturelle.

La photo-activation de la protéine GFP entraîne une réaction autocatalytique, provoquant l'émission, par la protéine, d'une fluorescence. Une cyclisation de la protéine est observée suite à la création d'une liaison chimique entre un atome d'azote de la glycine et un atome de carbone de la sérine; le cycle formé se déshydrate spontanément. L'oxygène environnant attaque une liaison de la tyrosine, ce qui a pour effet de former une double liaison entre celle-ci et la glycine. La molécule formée est le fluorophore produisant une émission de lumière.

3.2 Les plasmides

Le plasmide¹ pGLO sera utilisé comme « transporteur » (vecteur) du gène *gfp*.

Ce plasmide contient :

- le gène *bla* codant pour les β -lactamases qui détruisent l'ampicilline et qui confère à la bactérie une résistance à celle-ci.

¹ Les plasmides sont des molécules d'ADN circulaire, simple ou double brin, se reproduisant de manière indépendante du génome et ne codant pas des fonctions essentielles des bactéries.

- le gène *araC* : codant pour la protéine ARAC, facteur de régulation de la transcription de l'opéron arabinose, et ici, de la protéine GFP.
- P_{BAD} : gène promoteur de l'opéron arabinose et du gène *gfp*.
- *ori* : origine de réplication du plasmide dans *E.coli*.



Figure 3 : Cartographie du plasmide pGLO.

Le plasmide pNIL004 est semblable au plasmide pGLO à l'exception du promoteur qui est un « promoteur fort » permettant une expression plus importante de la protéine GFP dans la bactérie.

3.3 Production de la protéine GFP

La production de la protéine GFP dans *E. coli* va dépendre de l'efficacité du vecteur d'expression contenant la séquence codante pour celle-ci.

Les avantages de la production de protéines par des bactéries sont :

- un procédé moyennement complexe;
- une expression rapide;
- une production élevée;
- un milieu de culture bon marché;
- une purification simple;
- peu d'interférence avec l'hôte.

Par contre la production de protéines se fait souvent sous forme d'agrégats insolubles appelés "corps d'inclusion". Ceux-ci peuvent être isolés par centrifugation, puis solubilisés dans une solution d'urée 8M; les protéines se

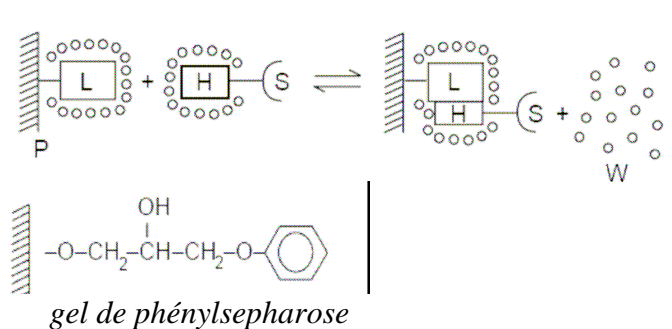
trouvent alors sous forme dénaturée. Une étape supplémentaire de renaturation permettra de retrouver la structure initiale de la protéine.

3.4 Extraction et purification de la protéine GFP

La protéine GFP n'est pas excrétée par *E. coli* : on aura donc production de la protéine dans le cytoplasme. Il faudra dès lors envisager une technique de rupture cellulaire afin d'isoler la protéine. De nombreuses techniques sont disponibles. Dans notre cas nous avons choisi de combiner diverses méthodes pour optimiser la rupture des cellules d'*E. coli* (voir partie expérimentale).

La protéine GFP formant des "corps d'inclusion", elle sera isolée, séparée, solubilisée dans l'urée 8M, et enfin renaturée. Une séparation, par chromatographie d'interaction hydrophobe, achèvera la purification de la protéine.

Cette méthode repose sur une interaction entre les molécules à fractionner et le support chromatographique. **La chromatographie en phase inverse** utilise un support greffé, porteur de chaînes hydrophobes, et un éluant de polarité variable. A ce procédé s'apparente **la chromatographie d'interaction hydrophobe (HIC)** dans laquelle l'éluant s'effectue en utilisant un gradient décroissant de force ionique.



| | |
|---|---|
| P : matrice de polymère | S : molécule soluble |
| H : partie hydrophobe à la surface des molécules solubles | L : ligand attaché à la matrice de polymère |
| W : molécules d'eau de la solution | |

Figure 4 : Principe de la chromatographie d'interaction hydrophobe.

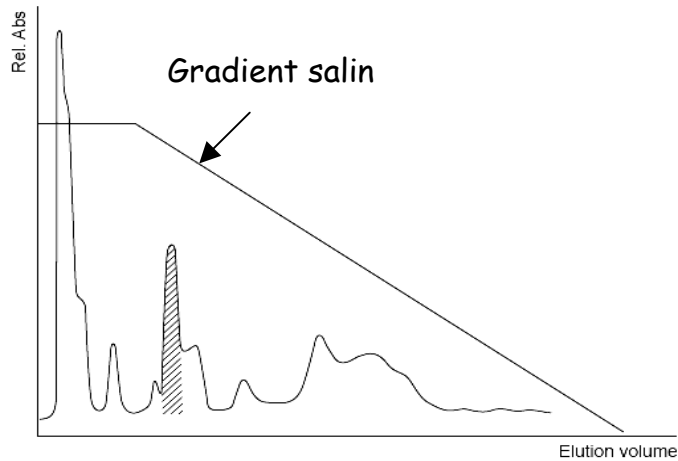


Figure 5 : Chromatogramme obtenu (absorbance = $f(\text{volume d'élution})$) suite à l'élution, par un gradient salin, d'une colonne HIC.

4. Partie expérimentale

4.1 Transformation de bactéries Escherichia coli-TG1 par le plasmide pGLO

La transformation des bactéries E.coli-TG1 par le plasmide pGLO est réalisée par un traitement des bactéries au chlorure de calcium (neutralisation des charges négatives de l'ADN plasmidique ainsi que de celles portées par la membrane des bactéries) associé à un choc thermique; une incubation sur glace qui rigidifie les membranes cellulaires est suivie d'une élévation brutale de la température qui augmente la perméabilité de ces mêmes membranes, favorisant ainsi la pénétration du plasmide dans la bactérie.

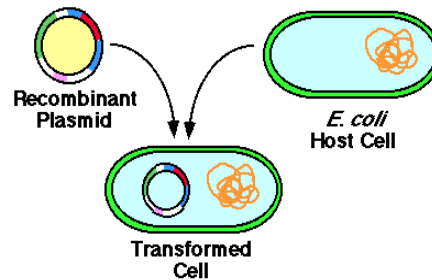


Figure 6 : Transformation de *E.coli-TG1* par le plasmide pGLO.

Les bactéries ainsi transformées sont déposées sur un milieu de culture gélosé Luria Bertani (LB) contenant notamment de l'ampicilline (incubation à 37°C pendant 24 h). Le plasmide pGLO contenant un gène de résistance à cet antibiotique, seules les bactéries ayant incorporé le pGLO se développeront sur ce milieu de culture.

4.2 Vérification de l'incorporation du plasmide pGLO par les bactéries (mise en évidence de la production de la protéine fluorescente GFP)

Une colonie isolée est prélevée à la surface du milieu de culture LB/Ampicilline et est réensemencée sur un milieu gélosé LB/Ampicilline contenant de plus du L-Arabinose; celui-ci étant l'inducteur d'expression de la protéine GFP (le gène codant pour cette protéine se trouve dans le plasmide pGLO).

Après une période d'incubation à 37°C pendant 24 heures, la boîte de Pétri, soumise à une exposition aux rayonnements UV, présente une fluorescence verte, caractéristique de la protéine GFP produite.

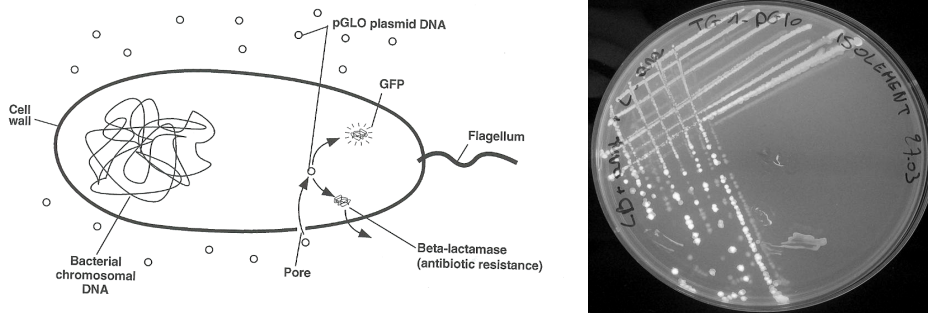


Figure 7 : Colonies de bactéries *E.coli-TG1* transformées par le plasmide pGLO et produisant la protéine fluorescente GFP.

4.3 Vérification de l'incorporation du plasmide pGLO par les bactéries (mise en évidence de la présence du gène *gfp*)

Une colonie isolée est prélevée à la surface du milieu de culture LB/Ampicilline et est réensemencée dans 50 ml de bouillon LB/Ampicilline. Celui-ci est alors incubé, sous agitation permanente, à 37°C pendant 24 heures. L'ADN plasmidique est ensuite extrait des bactéries à l'aide d'un kit commercialisé par la firme SIGMA (GenElute™ High Performance Plasmid Midiprep Kit).

La quantité d'ADN plasmidique n'étant pas suffisante pour permettre sa visualisation sur gel d'agarose, le plasmide pGLO a été soumis à une réaction en chaîne par polymérase ou PCR. Cette technique de biologie moléculaire permet d'amplifier énormément le nombre de copies d'une séquence spécifique d'ADN et ce, même si la quantité initiale d'ADN est très faible. Dans notre cas, la séquence amplifiée est celle codant pour la protéine GFP (en utilisant des amorces à la polymérisation « bordant » le gène *gfp*). La vérification de l'incorporation du plasmide pGLO par les bactéries revient donc à mettre en évidence l'existence du gène *gfp*.

Le résultat de l'amplification PCR est analysé par électrophorèse sur gel d'agarose : sous l'action d'un champ électrique, les molécules d'ADN (toutes chargées négativement) migrent dans le gel, vers une électrode positive, en parcourant une distance qui dépend de leur charge et donc de

leur taille (exprimée en nombre de paires de bases ; pb). Un marqueur de poids moléculaire (composé de dix étalons standards) est analysé en parallèle avec les échantillons ; celui-ci permet, après coloration du gel au SYBR Green et visualisation sous lampe UV, d'évaluer la taille des molécules d'ADN ainsi séparées.

L'étude du gel réalisé par les étudiants atteste de la présence d'une bande située entre 700 et 800 pb ; il s'agit là vraisemblablement du gène *gfp* (750 pb).

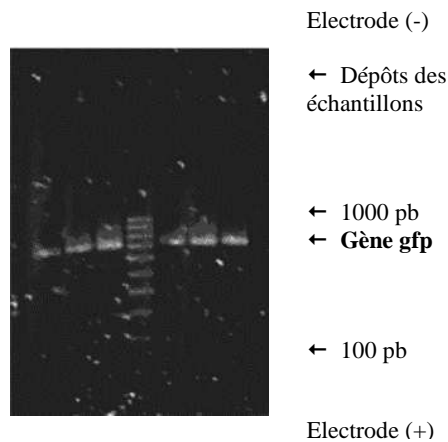


Figure 8 : Mise en évidence de la présence du gène *gfp* par électrophorèse sur gel d'agarose.

4.4 Production de la protéine GFP par les bactéries E.coli-TG1 transformées (ensemencement en milieu liquide)

Une colonie isolée est prélevée à la surface du milieu de culture LB/Ampicilline et est réensemencée dans 5 ml d'un bouillon de culture LB/Ampicilline/L-Arabinose (37°C pendant 24 hrs) afin d'obtenir une pré-culture bactérienne. Celle-ci est ensuite introduite Une colonie isolée est prélevée à la surface du milieu de culture LB/Ampicilline et est réensemencée dans 5 ml d'un bouillon de culture LB/Ampicilline/L-Arabinose (37°C pendant 24 hrs) afin d'obtenir une pré-culture bactérienne. Celle-ci est ensuite introduite dans 50 ml de bouillon LB/Ampicilline/L-Arabinose, et est incubée, sous agitation permanente, à 37°C pendant 48 à 72 heures; l'objectif visé étant une production protéique la plus importante possible.

Dès lors, l'estimation de la concentration de la suspension bactérienne est obtenue par une mesure de densité optique à 620 nm.
 dans 50 ml de bouillon LB/Ampicilline/L-Arabinose, et est incubée, sous agitation permanente, à 37°C pendant 48 à 72 heures; l'objectif visé étant une production protéique la plus importante possible.
 Dès lors, l'estimation de la concentration de la suspension bactérienne est obtenue par une mesure de densité optique à 620 nm.

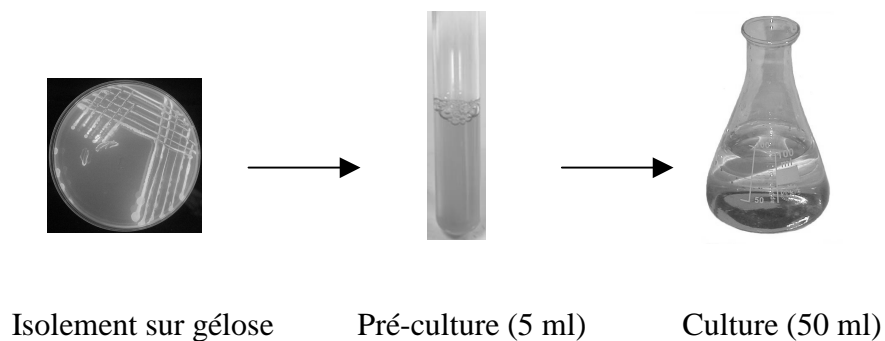


Figure 9 : Production de la protéine GFP.

Cette expérience a été reproduite à l'identique avec des bactéries E.coli–TG1 ayant incorporé un autre plasmide nommé pNIL004. Celui-ci est semblable au pGLO, mais possède cependant un "promoteur fort" permettant une expression plus importante de la protéine GFP.

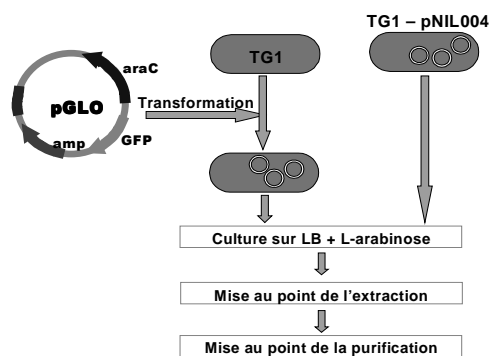


Figure 10 : Schéma opérationnel en vue de la production de la protéine GFP.

4.5 Extraction des protéines, et en particulier de la protéine GFP

Après de nombreux essais, les étudiants sont parvenus à extraire les protéines des cellules bactériennes (essais réalisés sur E.coli-TG1-pGLO et sur E.coli-TG1-pNIL004) ainsi qu'à les solubiliser selon le protocole d'extraction schématisé à la figure 11.

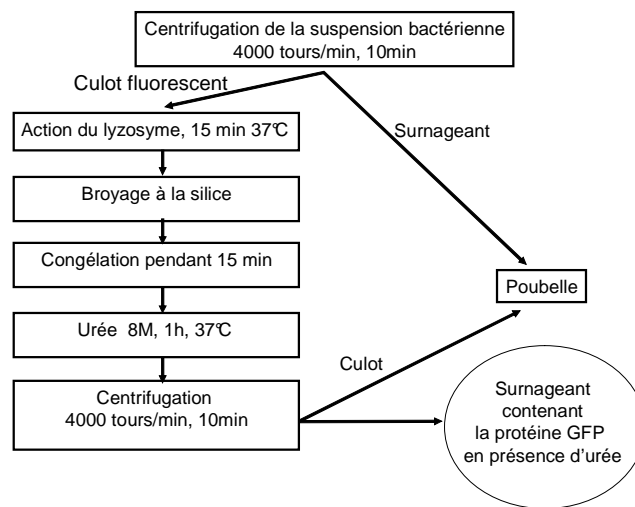
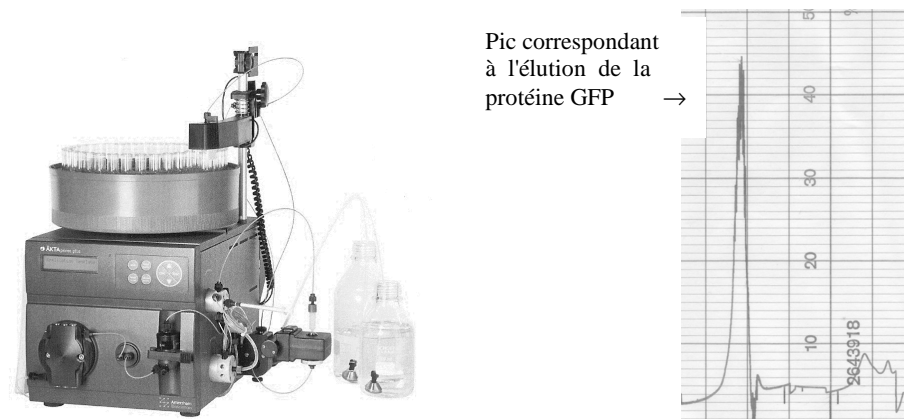


Figure 11 : Protocole d'extraction des protéines.

4.6 Purification de la protéine GFP par chromatographie hydrophobe

La colonne chromatographique est conditionnée à l'aide d'une solution de sulfate d'ammonium 2 M préparée dans un tampon Tris EDTA pH 7.4. L'échantillon protéique est introduit au sommet de la colonne et pénètre dans celle-ci. Une solution de sulfate d'ammonium 2 M dans le tampon Tris-EDTA traverse alors la colonne entraînant avec elle les protéines « indésirables » alors que les molécules protéiques de GFP restent accrochées au support hydrophobe. Ensuite, le tampon Tris-EDTA seul décrochera la protéine GFP. Une mesure continue de la densité optique à 280 nm de la solution éluée permet d'obtenir un chromatogramme attestant du "décrochage" des protéines. La totalité de la solution protéique est, quant à elle, récupérée dans divers tubes successifs.

(Remarque : un premier essai non concluant a été réalisé en effectuant un gradient de concentration de la solution de sulfate d'ammonium (2 M à 0 M) dans le Tris-EDTA).



*Figure 12 : - AKTAprime plus, un système compact pour chromatographie en phase liquide.
- Purification de la protéine GFP par HIC.*

4.7 Caractérisation de la protéine GFP

La caractérisation d'une protéine consiste notamment à estimer sa masse moléculaire à l'aide d'une technique appelée PAGE-SDS (électrophorèse sur gel de polyacrylamide en milieu dénaturant). Les étudiants n'ont malheureusement pas pu confirmer la masse moléculaire de la protéine GFP (27000), la quantité d'échantillon analysée ayant été trop faible.

Un dosage protéique par la technique de Bradford a également été réalisé mais il n'a cependant pas permis d'estimer la concentration totale en protéine GFP extraite des bactéries.

5. Conclusions

5.1 Conclusions

Grâce à ce projet, les étudiants ont pu se familiariser avec les techniques employées en biologie moléculaire : insertion du plasmide pGLO dans *E. coli*, amplification PCR pour la mise en évidence de la "réussite" de la transformation bactérienne (présence du gène *gfp* sur gel d'agarose), production de la protéine GFP par les bactéries, son extraction (mise au point du protocole) et sa purification (à l'aide d'une technique chromatographique).

5.1 Perspectives

Le remplacement du plasmide pGLO par le plasmide pBAD-6His, notamment, nous semble être une perspective intéressante. Le gène *gfp* produit par PCR à partir du plasmide pGLO (cfr partie expérimentale), puis isolé par électrophorèse, serait inséré au site de restriction Pst I/ EcoRI du plasmide pBAD.

La nouvelle construction aurait les mêmes caractéristiques que le pGLO, enrichies de la présence, à la fin de la protéine GFP, d'une terminaison peptidique de 6 Histidines.

Ce marquage nous permettrait de séparer la protéine GFP des autres molécules du mélange par une chromatographie d'affinité sur colonne « chelate de Ni » (His tag purification HisTrap; GE Healthcare).

Cette méthode devrait rendre plus rapides les étapes d'extraction - purification car l'on pourrait alors travailler avec un extrait cellulaire brut déposé directement sur la colonne « chelate de Ni ».

Dans le cas éventuel de la formation de corps d'inclusion, il serait possible, après redissolution de ceux-ci, de renaturer la protéine (re-folding On-column HisTrap; GE Healthcare) par une technique similaire.

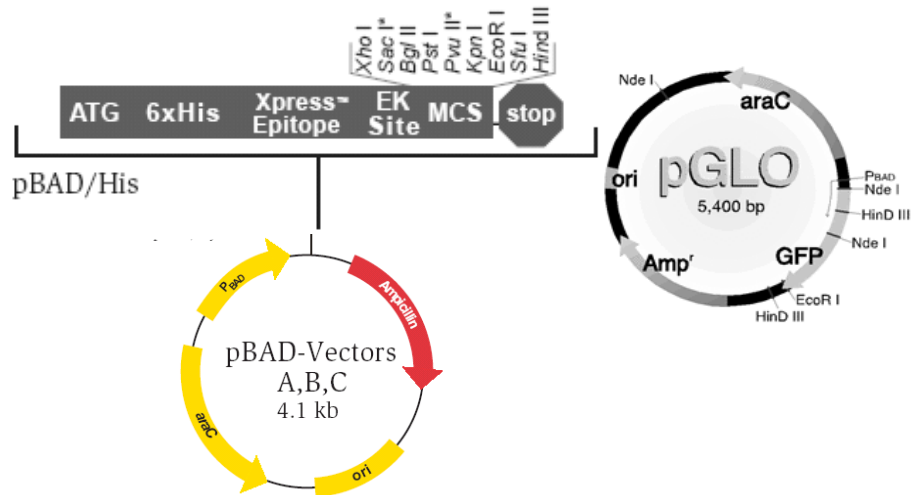


Figure 13 : Le plasmide pGLO produit la protéine GFP tandis que le plasmide pBAD produirait la protéine GFP-6His.

6. Bibliographie

- [1] RIGAUX P., *La Green Fluorescent Protein (GFP)*, adresse URL <http://www.scmbb.ulb.ac.be/Users/jean/BioInformatic/Session2001-02/GFP/>.
- [2] WATSON ET ALL, *ADN recombinant*, De Boeck Université.
- [3] HITOMI S. et BROWN K., *This little Light of Mine : Transform bacteria with a Jellyfish gene to make them glow*, adresse URL <http://www.panam.edu/dept/biotech/docs/Little%20Light.pdf>
- [4] N. AMEZIANE, M. BOGARD ET J. LAMORIL, *Principes de biologie moléculaire en biologie clinique*, Campus référence, Elsevier.