

EXPRESSION *IN VIVO* DU GÈNE GFP SOUS LA FORME D'UNE PROTÉINE DE FUSION

LUCIAN NEGURĂ^{1*}, ANCA HUMĂ¹

Mots-clef : GFP, clonage, phagemide, expression in vivo, protéine de fusion.

Résumé: La protéine GFP (Green Fluorescent Protein) se caractérise par sa capacité à émettre sous l'action des rayons UV une lumière verte de basse énergie, ce qui la rend particulièrement utile en biotechnologies. Nous avons procédé à la production de la GFP in vivo par *Escherichia coli* sous forme d'une protéine chimère de fusion avec le peptide α de la β -Galactosidase. Pour ce, nous avons cloné un fragment d'ADN comportant la phase codante de la GFP à partir d'un vecteur dans lequel il ne peut être exprimé vers un autre vecteur permettant lui cette expression.

INTRODUCTION

La protéine fluorescente verte (Green Fluorescent Protein ou GFP) est une protéine monomérique de 28 kDa produite par une espèce de méduse (*Aequorea victoria*) qui vit dans les eaux froides de l'océan Pacifique [Prasher 1992, Yang 1996]. Le rôle de cette protéine spontanément fluorescente est de catalyser une réaction d'oxydation accompagnée d'une production de photons, de telle manière qu'elle puisse absorber les rayonnements solaires ultraviolets en émettant une lumière verte de basse énergie. Une propriété essentielle de la GFP est représentée par sa capacité de tolérer des fusions N- ou C- terminales avec une grande variété d'autres protéines, fusions qui conservent les propriétés natives de ces protéines. Cette énorme flexibilité d'un marqueur non invasif dans les cellules vivantes permet à l'heure actuelle de nombreuses applications comme un traceur de lignées cellulaires, un rapporteur de l'expression des gènes ou encore la mesure des interactions protéine-protéine. Dès lors, la GFP s'est montrée d'une grande importance en biotechnologies par son utilisation dans le but de visualiser in vivo d'autres molécules qui peuvent y être attachées [Chalfie 1994]. Par la réalisation de constructions GFP-protéine cible ou même GFP-virus, on peut suivre l'évolution et les mouvements intracellulaires des molécules, en utilisant des radiations UV que la GFP transforme en lumière verte, facile à suivre. Ces buts justifient l'ampleur qu'a pris la production à échelle industrielle de la GFP, par le clonage et l'expression de son gène dans divers microorganismes ou cellules des organismes supérieurs [Cubitt 1995].

Nous avons essayé d'obtenir l'expression in vivo de la protéine GFP sous la forme d'une protéine de fusion avec le peptide α de la β -Galactosidase, par l'insertion du gène de la GFP dans un vecteur d'expression et la transformation de colonies d'*E. coli* avec le vecteur phagemidique ainsi obtenu, en vue de la synthèse protéique par la souche bactérienne.

STRATÉGIE ET MÉTHODOLOGIE DE TRAVAIL

Le point de départ des manipulations est représenté par une construction appelée pGFP, dans laquelle un fragment ADN de 0,73 kbases aux extrémités nettes, correspondant à la phase codante de la protéine GFP, a été inséré dans le site EcoRV de la cassette de clonage du vecteur pBluescript II/Sk+ (Figures 1). Le vecteur phagemidique pBluescript II/Sk+, utilisé fréquemment en génie génétique pour clonage ou expression des gènes [Altin-Mees 1992], est constitué de 2961 paires de bases incluant les éléments essentiels suivants :

- Bla (Ap^R) : Gène de résistance à l'ampicilline, nécessaire pour la sélection positive des colonies bactériennes transformées par pBSK ayant acquis par cette transformation une résistance à l'antibiotique respectif ;
- MCS (Multiple Cloning Site) : Cassette de clonage contenant des sites uniques de restriction pour de nombreuses enzymes de restriction ; cette cassette est flanquée par les promoteurs spécifiques des ARN-Polymérases des phages T3 et T7 (P_{T3} et P_{T7} respectivement), promoteurs nécessaires pour le processus de la transcription. La nomenclature du vecteur (SK ou KS) provient de l'ordre de succession à l'intérieur de cette cassette des sites de restriction SacI et KpnI, dans le sens de la réplication plasmidique ;
- f1 : Origine de réplication du phage filamentueux f1 permettant la synthèse d'ADN simple brin de pBSK lorsque la bactérie hôte contenant le phagemide est infectée par un phage helper. Cet ADN simple brin peut correspondre au brin sens (+) si l'origine est f1+ ou antisens (-) si l'origine est f1-. Le repère sens/antisens est donné par le sens de lecture du gène lacZ', la nomenclature du vecteur (pBSK+/-) tenant également compte de la position de cette origine +/- ;
- rep (pMB1) : Origine de réplication du plasmide Col E1. En absence du phage helper, le plasmide pBluescript est répliqué uniquement sous sa forme double brin ;

GFP sera préparé par digestion de la pGFP par l'enzyme de restriction BamHI et sera ensuite inséré dans le site BamHI du vecteur pBluescript II/SK+. Quoique cette opération permet l'insertion bidirectionnelle du fragment (extrémités BamHI cohésives identiques, Figure 3), une des deux orientations permettra aux régions codantes de lacZ' et de GFP de se trouver en phase. Cette nouvelle construction, appelée pSKGFP, permettra l'expression in vivo par E coli de la GFP sous la forme d'une protéine de fusion avec le début du peptide α .

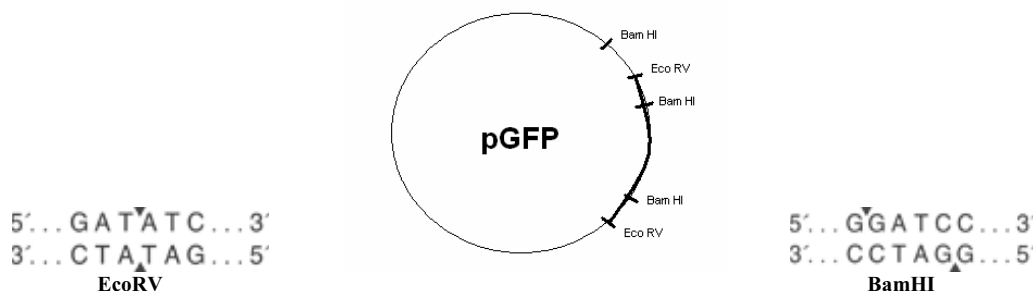


Figure 3. Les sites de restriction EcoRV et BamHI

Le travail a consisté en trois grandes étapes :

1) Clonage du fragment ADN portant la phase codante de la protéine GFP

Afin d'insérer le fragment dans la cassette de clonage du vecteur, pGFP et pBSK+ ont été soumis à l'action de l'enzyme de restriction BamHI. Alors que le vecteur pBSK fournit par cette digestion un seul fragment linéaire d'environ 3 kbases, pGFP génère quant à lui le fragment souhaité de 0,73 kbases, un fragment plasmidique de 3 kbases et un court fragment provenant de la cassette de clonage (Figure 4). Les produits des deux digestions ont été déposés en gel d'agarose 0,6%, ensemble avec un échantillon de pBSK natif non digéré (témoin d'efficacité de l'hydrolyse) et avec un marqueur de taille BRL (MassRuler™ DNA Ladder). Après localisation dans le gel du fragment de 0,73 kbases, ce dernier sera élué de l'agarose par extraction au phénol/chloroforme.

L'estimation de la quantité d'ADN obtenue après purification, aussi bien pour pGFP que pour pBSK+, a été effectuée d'une manière relative, par comparaison de l'intensité de fluorescence du BET (Bromure d'ethidium) d'un échantillon d'ADN récupéré par rapport à la fluorescence de quantités d'ADN connues et ayant la même origine. Pour ce, une gamme étalon de différentes quantités d'ADN a été réalisée par dilutions en cascade. Après migration en gel d'agarose, la quantité d'ADN a été estimée par intensité des bandes.

Il est très important de connaître les quantités exactes de l'ADN afin d'établir le rapport optimum vecteur/insert lors de la réaction de ligation. Cette dernière a été effectuée par l'ADN ligase du bactériophage T7, qui catalyse la formation d'une liaison phosphodiester entre les extrémités 3'OH et 5'phosphate de fragments d'ADN double brin ; la réaction nécessite du Mg²⁺ comme co-facteur et de l'ATP comme source d'énergie. L'enzyme joint les extrémités franches ou cohésives (cas échéant) et répare des coupures simple brin dans l'ADN. Pour favoriser l'insertion de l'ADN dans le fragment linéarisé, on utilise un rapport molaire Insert/Vecteur de 3/1, compte tenu de la taille et de la quantité de chaque fragment. La ligation a également été réalisée en absence d'insert, le vecteur recircularisé sur lui-même représentant le témoin de ligation nécessaire pour la suite des manipulations.

Le plasmide recombinant a été introduit dans des cellules d'Escherichia coli. Les colonies utilisées ont été du type NM522 présentant le génotype suivant :

sup E thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5 (rk-, mk-) [F' proAB lacI^qZΔM15]

Pour la transformation, les cellules ont été rendues compétentes par leur traitement au TSB (Tryptic soy broth), PEG (polyéthylène glycol) et DMSO (diméthylsulfoxyde). Pendant la transformation ont été réalisés un témoin de viabilité, un témoin de transformation, un témoin de digestion, ainsi qu'un témoin de ligation.

2) Vérification relative et absolue de l'orientation de l'insert

Étant donné que les extrémités du vecteur et du fragment sont cohésives BamHI, l'insertion peut s'effectuer dans les deux sens, une vérification de l'orientation du fragment GFP à l'intérieur du nouveau vecteur s'avérant nécessaire. La simple présence de l'insert dans les colonies bactériennes a été vérifiée par une réaction PCR en utilisant comme amorces des séquences complémentaires aux promoteurs T3 et T7. L'analyse de l'orientation de l'insert a été effectuée par trois méthodes distinctes :

- Par une réaction PCR en utilisant l'amorce T3, ainsi qu'une amorce G1 interne au fragment GFP. L'amplification peut avoir lieu uniquement si les deux amorces se trouvent en sens contraire, c'est-à-dire dans un seul cas d'orientation de l'insert, celui de la transcription GFP.
- Une détermination de l'orientation relative de l'insert par hybridation d'ADN simple brin. Nous avons procédé à la biosynthèse *in vivo* d'ADN simple brin + par transformation des colonies avec un phage helper (M13K07) qui permet la réplique monocaténaire à partir de l'origine f1. La séquence du brin ADN obtenu à partir du plasmide sera variable sur la portion de l'insert fonction de l'orientation de ce dernier (on obtient l'un ou l'autre des brins du fragment). Deux brins d'ADN provenant de colonies différentes peuvent s'hybrider entre eux uniquement si l'insert a une orientation différente sur les deux plasmides ; deux brins provenant de la même construction ne possèdent pas de séquences complémentaires et ne peuvent s'hybrider. On parle lors d'une détermination relative de l'orientation, d'une colonie par rapport à une autre.
- L'orientation absolue de l'insert sera déterminée par hybridation du simple brin obtenu au point précédent avec l'ARN issu de la transcription *in vitro* par la T3 ARN polymérase. Cet ARN possède une séquence complémentaire avec celle du brin matrice sens et peut s'y hybrider. Dans l'autre orientation, le transcrit n'aura pas de séquence complémentaire et il n'y aura pas d'hybridation.

3) Expression *in vivo* de la GFP

La présence de la protéine GFP (sous forme chimère, mais conservant ses propriétés) a été analysée par électrophorèse en gel de polyacrylamide dénaturant et par coloration au bleu de Coomassie.

RÉSULTATS ET DISCUSSIONS

La figure 4 présente la migration en gel d'agarose 0,6% des produits de digestion de pSK et de pGFP par BamHI. On observe l'obtention d'un seul fragment linéarisé dans le cas du vecteur pSK et la séparation vecteur linéarisé-fragment GFP dans le cas de pGFP. Le petit fragment BamHI issu de la cassette de clonage de pGFP a une taille bien inférieure et ne peut être visible sur ce gel.

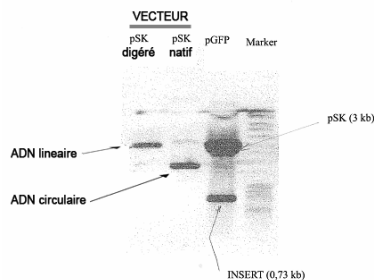


Figure 4. Efficacité de la digestion de restriction par BamHI

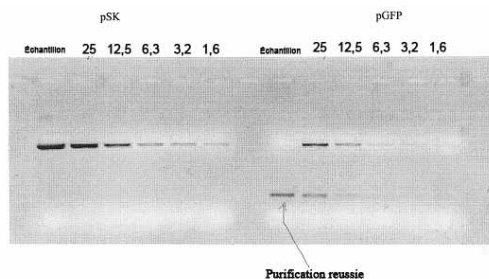


Figure 5. Estimation de la quantité d'ADN par intensité de bande sous UV

La quantité d'ADN obtenue dans le cas du vecteur et de l'insert a été estimée par intensité de bande par rapport à une gamme étalon. Nous avons réalisé des dilutions en cascade et déposés sur le gel des échantillons d'ADN de concentrations allant de 1,6 à 25 ng/μl en parallèle avec des échantillons des vecteurs digérés. Les résultats dans la figure 5 montrent que les concentrations de nos échantillons dépassent ceux de la gamme utilisée. Pour ce, nous avons dilué une partie de nos ADN afin de rentrer dans cette gamme (résultats non montrés). Finalement, nous avons conclu à une concentration de 150 ng/μl dans le cas du vecteur pSK et de 250 ng/μl pour pGFP. Cependant, dans le cas de l'insert GFP il a fallu considérer l'intensité relative par rapport à la taille de l'insert. Puisque 250 ng/μl correspond au vecteur entier de 3,72 kbases, pour l'insert seul (0,72 kbases) nous estimons une concentration environ 5 fois inférieure, c'est-à-dire de 50 ng/μl. Ces résultats correspondent à ceux obtenus en mesurant la densité optique des ADN à 260 et à 280 nm, ce qui prouve la relative précision de la méthode utilisée, quoiqu'elle soit plus élaborée que la simple méthode spectrophotométrique.

Le rapport massique pSK/GFP est fonction de la différence de taille (3 kbases pour le vecteur, 0,72 kbases pour le fragment), et devrait être d'environ 4,16. Cependant, en vue de la ligation nous devons également tenir compte du meilleur rapport Insert/Vecteur possible, avec un excès 3X pour l'insert, ce qui nous amène au résultat de 0,72 μg insert pour 1 μg vecteur. La ligation a été effectuée sur 100 ng vecteur pSK. Les bactéries rendues compétentes ont été transformées avec les ADN suivants en vue de vérifier l'efficacité du matériel et des opérations précédentes :

1. Le produit de ligation pSKGFP, afin de tester l'insertion ;
2. Le vecteur pSK reliqué sur lui-même, afin de tester le vecteur ;
3. Le vecteur natif pSK, non digéré par BamHI, afin de tester la compétence des bactéries ;
4. Le vecteur digéré par BamHI, afin de reconnaître les molécules circulaires d'origine de celles coupées puis reliquées ;
5. Sans ADN, culture sur milieu ampicilline, afin de vérifier la résistance portée par pSK vis-à-vis d'éventuelles sources de contamination ;
6. Sans ADN, culture sur milieu sans ampicilline, témoin de viabilité des bactéries.

En comptant dans chacun de ces cas le nombre de colonies transformées bleues (sans insert), blanches (présence d'insert) et blanches fluorescentes (insert GFP fonctionnel), nous avons pu déterminer le rendement et l'efficacité de chacune des manipulations jusqu'à ce point.

La présence d'un insert à l'intérieur du vecteur a été vérifiée par PCR en utilisant les amorces T3 et T7. Dans la figure 6a nous pouvons observer l'amplification d'un fragment de 200 paires de bases à partir de l'ADN plasmidique provenant d'une colonie bleue (sensée ne pas présenter d'insert), alors que pour les autres colonies, blanches et fluorescentes, on obtient un fragment de 930 paires de bases correspondant à l'amplification du fragment GFP (0,72 + 0,2 kbases). L'apparition du même fragment pour la colonie bleue relève d'une contamination par une colonie voisine possédant elle cet insert. La figure 6b présente l'amplification T3/G1 destinée à déterminer l'orientation absolue de l'insert GFP (l'amorce G1 est interne au fragment GFP et doit se trouver en sens contraire de la T3 pour qu'il y ait amplification). On peut observer que cette amplification n'a eu lieu qu'en utilisant l'ADN plasmidique provenant des colonies fluorescentes, fait qui confirme la présence au sein de ces colonies d'un insert GFP fonctionnel en phase avec le début du peptide α .

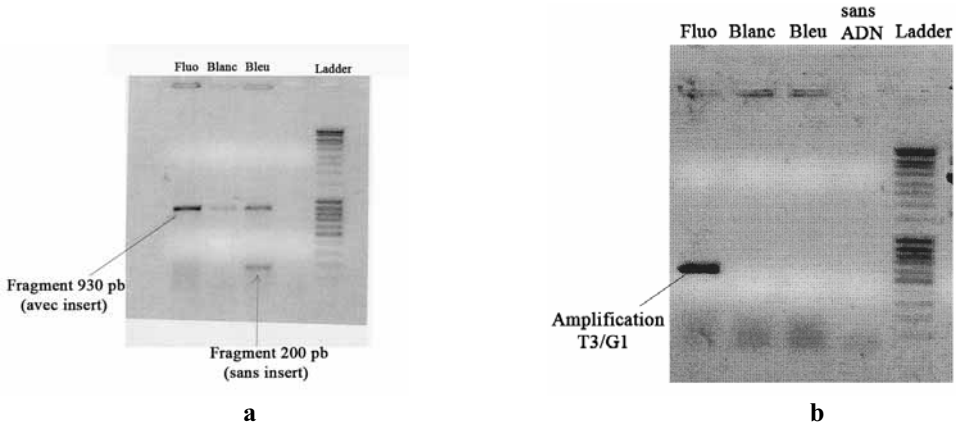


Figure 6. Amplification PCR pour vérifier la présence de l'insert (a) avec les amorces T3 et T7, et pour déterminer l'orientation absolue de l'insert (b) avec les amorces T3 et G1.

Après l'obtention de l'ADN plasmidique simple brin (réplication phagique à partir de ϕ 1+ avec phage helper) et avant de procéder aux hybridations ADN/ADN et ADN/ARN, nous avons dû réaliser la transcription *in vitro* de fragment GFP inséré par la T3 ARN polymérase. Pour ce, nous avons besoin d'ADN phagique linéaire et purifié. L'ADN plasmidique extrait des bactéries a été par conséquent digéré par l'enzyme de restriction HindIII et déposé en gel d'agarose 1%. Le fragment linearisé a été élué à partir du gel, précipité au phénol/chloroforme et a servi de matrice pour la synthèse d'ARN messager complémentaire. Tous ces résultats sont présentés dans la figure 7 a, b, c et d.

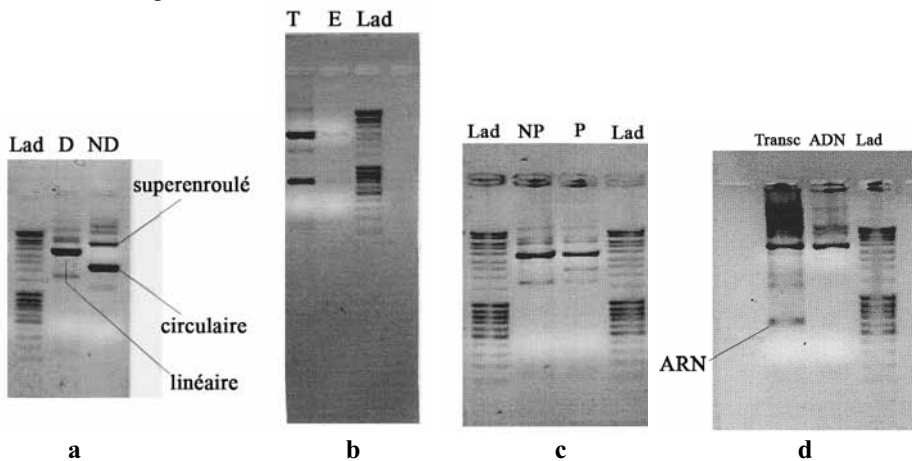


Figure 7. (a) Digestion de pSKGFP par HindIII (D=digéré, ND=non digéré) ; (b) Estimation de la quantité d'ADN après la coupure HindIII (T=témoin de concentration connue, E=échantillon) ; (c) Estimation de la quantité d'ADN restante après extraction du gel (NP=non purifié, P=purifié après élution) ; (d) Transcription *in vitro*

Une fois les simples brins étant obtenus, nous avons pu déterminer l'orientation de l'insert par hybridation de ces simples brins, aussi bien d'une manière relative (ADN/ADN) entre simples brins provenant de différentes colonies bactériennes, que de manière absolue par hybridation ADN/ARN (utilisation du produit de la transcription). Nous avons utilisé trois colonies différentes, dont deux fluorescentes (sensées contenir le fragment GFP en phase et dans le sens de la transcription) et une colonie blanche non fluorescente (contenant l'insert GFP non fonctionnel ou inséré dans l'autre sens). Les produits d'hybridation ont été déposés en gel d'agarose 0,8% dont la photo est présentée dans la figure 8.

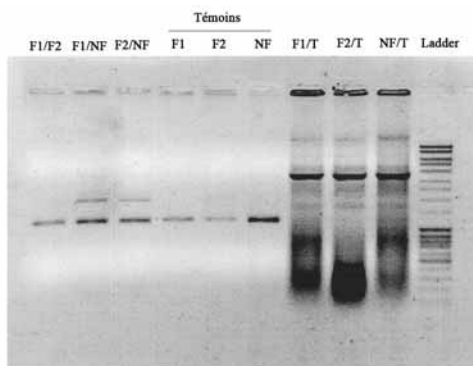


Figure 8. Hybridations des simples brins. F1 et F2 correspondent aux deux colonies fluorescentes 1 et 2. NF correspond à la colonie blanche non fluorescente. T représente le transcrit ARN.

Le fait que F1 et F2 ne s'hybrident pas montre que ces bactéries contiennent la même construction pSKGFP dans laquelle le fragment GFP est orienté dans le même sens, probablement celui de la transcription car la fluorescence des colonies est un signe de la production in vivo d'une partie fonctionnelle de la protéine GFP. Au contraire, NF s'hybride aussi bien avec F1 qu'avec F2, ce qui signifie que les bactéries blanches non fluorescentes prélevées contiennent bien l'insert GFP mais dans le sens opposé. Les résultats d'hybridation à l'ARN confirment l'hypothèse pour F1 et F2, malheureusement pour NF une hybridation non spécifique a du avoir lieu, certainement la formation d'un hétéroduplex à partir de deux brins qui ne sont pas complémentaires sur toute la séquence mais qui présentent des régions proches. Il faut préciser que la T3 ARN polymérase ne possède pas une activité précise de terminaison de la transcription, elle continue le long de la matrice jusqu'à certaines séquences terminatrices que notre construction ne contient pas. C'est pour cette raison que nous avons dû linéariser l'ADN avant la transcription, imposant ainsi à l'enzyme l'arrêt à un site en aval de la région d'intérêt (la coupure HindIII). On peut considérer que l'appariement non spécifique NF/T se situe non pas au niveau du fragment GFP, mais un peu en aval de ce dernier. Utiliser une autre enzyme de restriction que HindIII pourrait résoudre ce manque de spécificité.

Finalement, l'aboutissement du travail a consisté en l'électrophorèse en gel PAGE dénaturant de l'extrait protéique bactérien provenant de colonies fluorescentes. Nous avons essayé de mettre en évidence la protéine GFP par rapport aux marqueurs de migration de poids moléculaires connus. Nous avons obtenu (Figure 9) une bande correspondant à 31kDa, équivalent de la protéine de fusion entre GFP (28 kDa) et le peptide α de la β -Galactosidase.

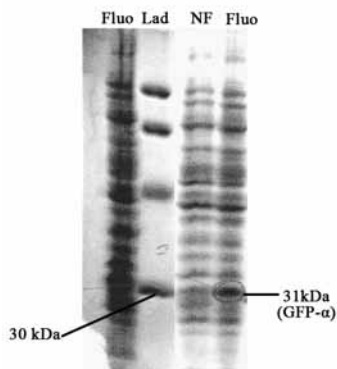


Figure 9. Expression in vivo de la GFP dans les colonies fluorescentes (Fluo) et très peu dans les colonies blanches non fluorescentes (NF). Lad=Marqueur de migration.

CONCLUSIONS

La suite d'expériences présentée nous a permis non seulement le production réussie in vivo de la Green Fluorescent Protein (GFP) sous forme d'une protéine de fusion, mais également la comparaison de différentes méthodes de travail en vue de l'optimisation de telles approches moléculaires. Nous avons ainsi comparé plusieurs techniques de différentes efficacités en vue de la mise au point d'un enchaînement moléculaire muni d'un bon rendement de production et de vérification. Cette approche complexe, quoique minutieuse, ouvre la voie à différents essais de génie génétique destinés à faire produire par des microorganismes différentes protéines végétales, animales, voire humaines, à l'échelle du laboratoire ou même industrielle.

RÉFÉRENCES

1. Yang F., Moss L.G. and Phillips G.N.Jr., 1996, The molecular structure of green fluorescent protein. *Nat Biotechnol*, **14(10)**, 1246-1251.
2. Prasher D., Eckenrode V., Ward W., Prendergast F. and Cormier M., 1992, Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene*, **111**, 229-233.
3. Chalfie M., Tu Y., Euskirchen G., Ward W. and Prasher D., 1994, Green fluorescent protein as a marker for gene expression. *Science*, **263**, 802-805.
4. Cubitt A., Heim R., Adams S., Boyd A., Gross L. and Tsien R., 1995, Understanding, improving and using green fluorescent proteins. *TIBS*, **20**, 448-455.
5. Altling-Mees M.A., Sorge J.A. and Short J.M., 1992, pBluescriptII: multifunctional cloning and mapping vectors. *Meth Enzymol*, **216**, 483-495.

1) Université «Alexandru Ioan Cuza», Iași, Roumanie

* correspondance à adresser à : luciannegura@yahoo.fr