

La production de protéines recombinantes : Etat de l'art de la technologie entre la génomique et la protéomique fonctionnelle. Une approche expérimentale du contrôle de la production et de la purification de protéines

El Marjou A^{1*}, V. Le jour¹, N., Lekieffre¹, L.Cabanié¹, A. Elmselmi²

1.Laboratoire de protéine recombinantes Institut Curie /CNRS UMR144 Paris
2.Laboratoire de Biologie moléculaire, Ecole de biologie industrielle, Cergy Pontoise
*elmarjou@curie.fr

Résumé

Le développement de nouvelles technologies de production et de purification de protéines recombinantes capables de répondre aux défis imposés par les programmes de biologie structurale et de protéomique fonctionnelle est devenu une nécessité incontournable.

La sur-expression de protéines dans des systèmes hétérologues comme la bactérie et la levure s'accompagne dans la majorité des cas de précipitation des protéines dans les corps d'inclusion. Le suivi en temps réel de la production et la purification de protéines recombinantes est devenue donc un enjeu majeur des procédés industriels d'optimisation de production de protéines recombinantes. Nous avons développé un procédé de contrôle de la production de protéines recombinantes comportant des traceurs colorés (cytochrome b5 et le domaine FMN du cytochrome p450) permettant de suivre la production ainsi que la purification des protéines d'intérêt et un double tag d'affinité afin d'améliorer et de raccourcir les étapes de purification

Mots clés : Protéines recombinantes, chromatographie d'affinité, cytochrome b5, FMN.

Introduction

Plusieurs génomes procaryotes et eucaryotes sont actuellement séquencés. Le défi actuel est de déterminer la structure et la fonction de toutes les protéines codées par ces génomes (Chandonia 2006). Toutes les protéines d'un organisme vivant procaryote ou eucaryote ne s'expriment pas de la même façon dans un organisme hétérologue. Le niveau d'expression est fonction de plusieurs paramètres intrinsèques (liés à la structure et à la fonction de la protéine) et extrinsèques (liés à la physiologie de l'hôte et aux conditions de sa culture). Depuis la découverte des techniques d'ADN recombinant, et de l'expression de la première protéine humaine dans la bactérie E.Coli, plusieurs autres systèmes d'expression de protéines recombinantes ont été mis au point (Baneyx F 1999, Macauley-Patrick 2005). Ces différents systèmes se caractérisent par la capacité de production et par la qualité des protéines produites en terme de modifications post traductionnelles. Le choix d'un système par rapport à un autre dépend de la nature de la protéine à produire et de son utilisation ultérieure. Le développement de protéines de fusion a permis d'optimiser plusieurs tag qui sont utilisés à la fois dans le marquage et la localisation des protéines ainsi que leurs purification de protéines par

chromatographie d'affinité. Ces différentes étiquettes diffèrent par leur taille et la nature de substrat utilisé pour décrocher les protéines de la résine (Lichty 2005). Certains partenaires de fusion comme la protéine MBP, Nus et SUMO sont utilisés pour permettre la solubilisation des protéines (Butt TR 2005). D'autres tags qui sont des peptides de petite masse comme 6xHIS, FLAG, StrepTag n'influençant pas la structure et la fonction de la protéine sont utilisés plus globalement dans la purification. La combinaison de plusieurs des peptides d'affinités en polytag permet des purifications d'affinité pouvant atteindre jusqu'à 98% de pureté. (Lichty 2005) D'autre part l'addition de séquences contenant des sites spécifiques à des protéases virales permet d'éliminer les partenaires de fusion et d'obtenir des protéines recombinantes natives exemptes de toutes séquences étrangères (Walker PA, 1994). L'élimination des tag présente un intérêt quand il s'agit de protéines à visée thérapeutique ou pour la détermination de la structure tridimensionnelle aux rayons X. Une des limitations du procédé de production de protéines recombinantes est l'absence de système capable de suivre la synthèse de protéines en temps réel. Plusieurs travaux ont montré la possibilité d'utiliser les protéines auto fluorescentes de la famille de GFP ou coloré comme les cytochromes pour suivre de l'expression de protéines

recombinantes (DeLisa MP 1999 ; Finn 2005, Kwon 2005)

Nous avons développé une stratégie de production alliant à la fois et le suivi de l'expression de protéines et la rapidité de purification et d'élimination des tags de fusion. Cette stratégie consiste en la fusion d'un domaine composé d'un tag 10xHIS et d'un domaine SBP d'interaction avec la streptavidine fusionné avec des protéines colorées comme le cytochrome b5 bovine ou le domaine FMN du cytochrome p450 humain.

Matériel et Méthodes

Vecteurs d'expression

La cassette d'expression composée de 10 histidines est réalisée par hybridation de deux oligonucleotides HIS1 et HIS2. Cette cassette est insérée dans le vecteur d'expression pET15b (Novagen). Le domaine SBP est amplifié par PCR à partir du vecteur p2yX en utilisant les amorces SBP1 et SBP2. Le produit de PCR est purifié à partir du gel inséré dans le vecteur pET15-10xHIS

Le cytochrome b5 et le domaine FMN p450 sont amplifiés par RT-PCR à partir des ARN totaux extraits du foie de Rat et Humain. Le cytochrome b5 est amplifié par les amorces B5S et B5AS, le cytochrome p450 par les amorces FMNS et FMNAS. Les réactions d'amplification PCR sont réalisées en utilisant la polymérase haute fidélité Pfu-Ultra (Stratagene) ;

Les fragments amplifiés sont insérés au niveau du site du vecteur pET15b-HSBP. Toutes les constructions obtenues sont séquencées par le protocole de séquençage ABI PRISM 377 DNA Sequencer.

Tableau 1 : Liste des oligonucleotides utilisés dans le travail.

HIS1	CATGGCCTCGTACTACCATCACC ATCACCATCACCATCACCATCAC TGCGGCCGC
HIS2	GCGGCCGACAGTGATGGTGATGG TGATGGTGATGGTGATGGTAGTA CGAGGC
SBP1	GCGGCCGCGGCTCAGGGATGGA CGAGAA
SBP2	TTAATTAAGGTACCGGATCCTGG TTCAC
B5S	ATCGCCGAGCAGTCAGACAAG
B5AS	AAGGGTTTCCGAAGGCTT
FMNS	ATCAGCAGCTTTGTGGAAAAGAT G
FMNAS	GCTGGAACCTCGCCAGT

Production et purification de protéines recombinantes

Pour tester la validité de notre système, des bactéries BL21DE3 pLysS sont transformées par les constructions pET-HSBP-b5 et pET-HSBP-FMN. Les bactéries recombinantes sont validées par PCR. La production de protéines est réalisée dans 100 ml de milieu LB contenant 100µg/ml d'Ampicilline et 34µg/ml de chloramphénicol ensemencés avec des clones pET-b5 et pET-FMN,

Les bactéries recombinantes sont cultivées jusqu'à atteindre la phase exponentielle de croissance à DO₆₀₀ de 0,5 puis la production de protéines est induite par 1 mM de l'IPTG, pendant 3 heures à 37 degrés sous 250 rpm d'agitation. Les bactéries induites sont collectées par centrifugation à 6000 rpm pendant 10 min et le culot de bactéries est lavé avec du PBS puis suspendu dans le tampon de lyse contenant du PBS, 0,5 mg/ml de lysozyme, 0,5 % Triton X-100, 10 mM MgCl₂, 100 µg/ml DNase I, plus les inhibiteurs de protéases (complet de chez roche)

La suspension bactérienne est incubée dans le tampon de lyse avec agitation à 4°C. Après 1 heure d'incubation 1M de NaCl est ajouté aux lysats de la suspension bactérienne puis ré-incubé pendant 1 heure. L'extrait est centrifugé à 25000 rpm pendant une heure et le sur-nageant contenant la protéine est récupérée puis chargé sur la colonne d'affinité His-Select (Sigma). La colonne est lavée avec du PBS contenant 350 mM NaCl et 15 mM Imidazole. Les protéines sont éluées avec le tampon PBS contenant 250 mM imidazole. Les fractions contenant les protéines recombinantes cytochrome b5 et FMN sont visibles à l'œil nu. Ces fractions sont regroupées puis la concentration de protéine est déterminée par le protocole de Bradford. Le pool de protéines est alors chargé sur une colonne de streptavidine-agarose. Après le lavage au PBS, les protéines sont décrochées par un tampon PBS contenant 10mM des-thiobiotin. Les fractions contenant les protéines sont contrôlées par le chromatogramme aux longueurs d'onde d'absorption 280nm et 412 pour le cytochrome b5 et 450 pour le domaine FMN

Analyse sur Gel SDS-PAGE

La qualité des protéines purifiées est analysée sur gel 4-12% SDS-PAGE Nupage (Invitrogen). Les protéines séparées sont révélées par coloration au bleu de commassie

Résultats et discussion

Structure des cassettes

Nous avons développé deux cassettes d'expression composées d'une séquence de 10XHIS, d'une séquence d'interaction avec la streptavidine (SBP) et de deux protéines colorées de la famille des cytochromes, le cytochrome b5 de rat (rouge) et le domaine FMN du cytochrome p450 Humain (jaunes).

Ces constructions comportent également une séquence de clivage par la protéase TEV qui permet d'éliminer cette cassette de la protéine d'intérêt une fois purifiée.

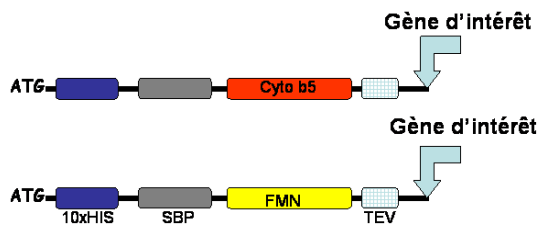


Figure 1-composition des deux cassettes d'expression colorées. Ces séquences sont insérés dans la partie terminale des gènes d'intérêt et se terminent par une séquence TEV de 6acides aminés

Procédé de production

Pour valider ces constructions, ces cassettes sont insérés dans le vecteur d'expression pET15b et sont transformées dans les bactéries BL21DE3.

Après induction et centrifugation des bactéries recombinantes il est facile d'observer l'expression des protéines de fusion avec le cytochrome b5 (figure 2a). A cette étape il est très difficile de savoir si la protéine est soluble, ou sous forme de corps d'inclusion. Après extraction et séparation des fractions solubles et insolubles, la protéine soluble colorée est présente dans le surnageant (figure 2b). Le chargement de l'extrait de protéines permet de suivre son accrochage sur la résine d'affinité NiNTA (figure 2c). L'élué de la colonne montre que la protéine attendue est bien localisée dans les fractions colorées (figure 2d)

Ces résultats montrent que le cytochrome b5 est d'une part un puissant traceur de l'expression de protéines recombinantes et d'autre part il a été montré qu'il permet également la solubilisation de certaines protéines normalement insolubles dans la bactérie. (Mitra 2005). Sa masse moléculaire réduite moins de 15kDa et sa coloration rend cette protéine plus avantageuse que d'autres partenaires de fusions comme les protéines GST et GFP

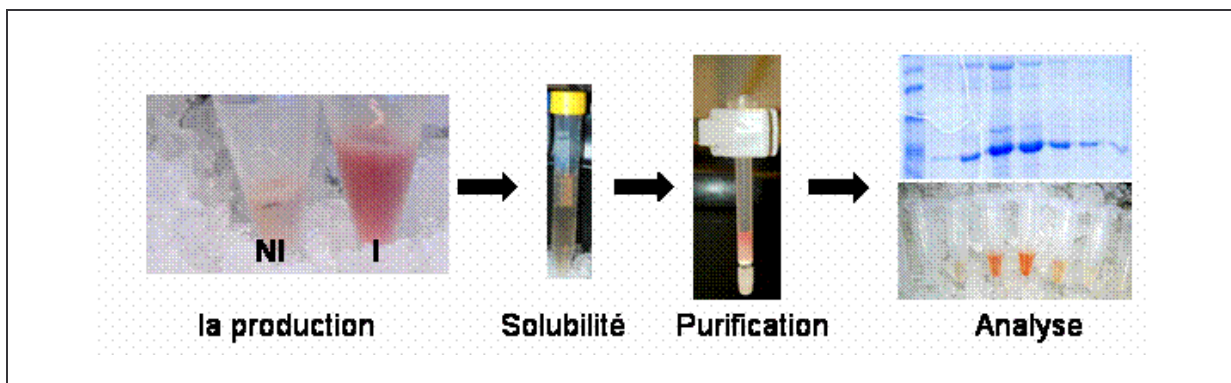


Figure 2-Contrôle des différentes étapes du procédé de production de protéines recombinantes utilisant les traceurs colorés. Expression dans les cellules induites (a), présence de la protéines dans la fraction soluble (b), liaison sur la colonne d'affinité (c) correspondance entre le contenu des fractions et la présence de la protéine d'intérêt (d)

La fusion de plusieurs domaines entraîne parfois des encombrements stériques et l'inaccessibilité de certaines parties de la protéine à la liaison sur les colonnes d'affinité. Pour tester l'accessibilité du domaine SBP, les protéines purifiées de la colonne His-Select sont chargées sur la streptavidine agarose. La purification est suivie directement par deux longueurs d'onde 280nm pour la présence des protéines et 412nm ou 450 nm pour la présence du cytochrome b5 et p450. Le résultat de la figure 3 montre que le cytochrome b5 est plus accessible comme en témoigne le rapport 412/280 proche de 1 alors que l'absorbance du domaine FMN est influencé par la fusion avec le tag 10xHIS-SBP le rapport de 450/280 est proche de 0. Il est donc impératif de modifier la séquence charnière de fusion pour améliorer l'accessibilité de ce domaine.

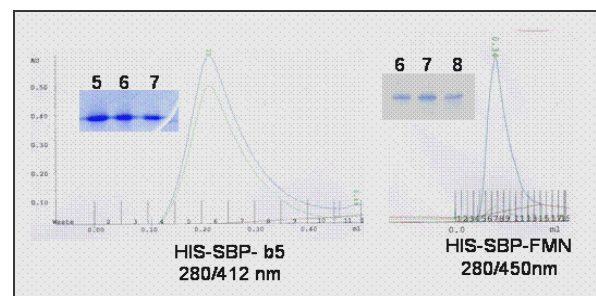


Figure 3-purification des protéines HIS-SBP-cytochrome b5 et HIS-SBP-FMN sur colonne Streptavidine agarose/

La purification est contrôlée par l'absorbance de protéines à 280nm et celle du cytochrome b5 à 412 nm et celle du domaine FMN à 450nm. Les

fractions correspondantes aux maximums du pic d'absorbance sont analysées sur GEL SDS-PAGE

L'analyse de protéines correspondant au pic d'absorbance confirme bien la présence des deux protéines en question

Bibliographie

Chandonia JM, Brenner SE. 2006 The impact of structural genomics: expectations and outcomes. *Science*, 311:347-51.

Kwon SY. et al 2005; Highly efficient protein expression and purification using bacterial hemoglobin fusion vector. *Plasmid*, 53:274-82.

Finn RD, Kapelioukh I, Paine MJ. 2005. Rainbow tags: a visual tag system for recombinant protein expression and purification. *Biotechniques*, 38:387-8,

Butt TR, Edavettal SC, Hall JP, Mattern MR 2005. SUMO fusion technology for difficult-to-express proteins. *Protein Expr Purif.*;43:1-9.

Baneyx F. 1999. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol.*;10(5):411-21.

Macauley-Patrick S, 2005 Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*, 22:249-70.

Lichty et al 2005 ; Comparison of affinity tags for protein purification. *Protein Expr Purif.*;41:98-105.

DeLisa MP, Li J, Rao G, Weigand WA, Bentley WE. 1999. Monitoring GFP-operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor. *Biotechnol Bioeng*, 5;65:54-

Walker et al 1994; Efficient and rapid affinity purification of proteins using recombinant fusion proteases. *Biotechnology (N Y).*;12:601-5.

Mitra A. et al 2005. High level expression of peptides and proteins using cytochrome b5 as a fusion host. *Protein Expr Purif.*;41:84-97.