

Méthodes d'obtention d'un allergène recombinant

Production of recombinant allergens

A. Casset

UMR 7199, CNRS, laboratoire de conception et application de molécules bioactives, faculté de pharmacie, université de Strasbourg,
74, route du Rhin, 67400 Illkirch, France

Résumé

Les progrès des techniques de génie génétique, et leur application en allergologie, ont fait apparaître la possibilité de programmer une cellule pour synthétiser des allergènes recombinants. L'expression de ces allergènes peut se faire dans des systèmes procaryotes (bactéries) ou eucaryotes (levures, cellules d'insecte, plantes). Le choix de la cellule hôte est principalement guidé par les modifications post-traductionnelles que doit subir la protéine pour être biologiquement active. La caractérisation biochimique et biophysique de l'allergène permet le contrôle conformationnel de la protéine produite. La réalisation de tests *in vitro* et *in vivo* conduit à la validation de l'activité immunologique de ces allergènes recombinants.
© 2011 Elsevier Masson SAS. Tous droits réservés.

Mots clés : Allergènes recombinants ; Système procaryote ; Système eucaryote

Abstract

Advances in genetic engineering techniques, applied to the field of allergy, allowed to program a cell to synthesize recombinant allergens. Recombinant allergenic proteins have been produced in a variety of different expression systems such as prokaryotic (bacterial) or eukaryotic expression systems (yeasts, insect cells, plant). One of the most important factor to consider in choosing an expression system is the processing of post-translational modifications in order to produce proteins that are correctly folded and biologically active. Biochemical and biophysical characterization of the allergen is recommended to check the conformation of the protein produced. Assessment of the immunological activity of the recombinant allergen require to perform various tests *in vitro* and *in vivo*.

© 2011 Elsevier Masson SAS. All rights reserved.

Keywords: Recombinant allergens; Prokaryotic system; Eucaryotic system

1. Introduction

L'essor du génie génétique au début des années 1970 a fait apparaître la possibilité de programmer une cellule pour synthétiser des protéines d'intérêt. Ce procédé biotechnologique permet de produire des molécules trop complexes à synthétiser chimiquement. Une protéine est qualifiée de recombinante dans la mesure où elle est produite de manière exogène dans une cellule dont l'ADN a été modifié par recombinaison génétique. Ce processus biotechnologique s'appuie sur l'isolement de l'ADN codant pour la protéine, l'emploi d'un vecteur transportant le gène d'intérêt codant pour

la protéine recherchée, l'utilisation d'une cellule hôte qui sera chargée de synthétiser la protéine via le gène inséré, l'expression, puis la séparation et l'extraction de la protéine de la cellule hôte dans laquelle elle a été produite. À la fin des années 1980, cette technique a été appliquée à la synthèse d'allergènes, Der p 1 [1] et Bet v 1 [2] furent parmi les premiers à avoir été clonés.

2. Extraits allergéniques, allergènes naturels purifiés et allergènes recombinants

Les allergènes existent sous différentes formes : naturelle et recombinante. Les allergènes naturels peuvent être contenus dans un mélange, l'extrait allergénique, ou sous forme purifiée. Les extraits allergéniques obtenus par extraction de la matière

Adresse e-mail : casset@unistra.fr.

première de la source allergénique (cultures d'acariens, grains de pollen...) peuvent être variables quant à leur contenu en molécules allergéniques à la fois sur le plan qualitatif et quantitatif. Ils contiennent également des molécules non allergéniques en proportion plus ou moins élevée. La teneur en allergènes dépend grandement des procédés d'extraction et de préparation. Il a été montré pour les extraits allergéniques d'acariens, que la teneur en allergènes majeurs peut être très variable selon qu'ils sont préparés à partir du milieu de culture des acariens ou des corps des acariens [3,4]. Les allergènes naturels purifiés, issus de la source allergénique, sont caractérisés par un polymorphisme important sur le plan moléculaire et de leur glycosylation. Les procédés d'extraction et de purification ne permettent pas d'obtenir des rendements élevés. Enfin, les allergènes recombinants correspondent à un variant précis. Ils présentent, par rapport aux extraits allergéniques, certains avantages : ils sont parfaitement caractérisés au niveau moléculaire et immunochimique, ils sont hautement purifiés et faciles à standardiser. Ils permettent une utilisation en diagnostic et en thérapeutique [5–7]. La commission de la nomenclature de l'Union internationale des sociétés d'immunologie (IUIS) a uniformisé la dénomination des allergènes en utilisant les trois premières lettres du genre de la source d'allergènes, la première ou la deuxième lettre de l'espèce, enfin un nombre arabe reflétant l'ordre dans lequel l'allergène a été découvert et/ou son importance clinique. Les allergènes naturels purifiés sont précédés de la lettre « n » tandis que les allergènes recombinants sont précédés de la lettre « r ».

3. Obtention d'un allergène recombinant

3.1. Obtention de l'ADN codant pour l'allergène

La première étape de la production d'un allergène recombinant consiste au choix et à l'isolement de la séquence d'ADN codant pour cet allergène. Plusieurs approches peuvent être utilisées : l'amplification directe de la séquence codante, le criblage immunologique d'une banque d'ADN complémentaires (ADNc) exprimée dans des bactériophages ou celle obtenue par la méthode « phage display » [8,9]. Dans les cellules eucaryotes, la présence d'introns dans l'ADN génomique ne permet pas son clonage direct en vue de l'obtention d'une protéine recombinante. Il est nécessaire d'utiliser l'ARN messager et d'effectuer une conversion en ADN complémentaire par la technique de *reverse transcription – polymerase chain reaction* (RT-PCR). L'approche la plus rapide pour l'obtention de l'ADN codant pour un allergène consiste en l'amplification de la séquence codante par RT-PCR à partir des ARN extraits de cellules exprimant le gène d'intérêt. Cependant, cette méthode nécessite de connaître une partie de la séquence en aminoacides de l'allergène et plus spécifiquement la séquence N terminale, ce qui n'est pas toujours le cas. Les autres approches consistent au criblage immunologique d'une banque d'ADNc. À partir d'une source d'allergènes (culture d'acariens, grains de pollen...) les ARN messagers totaux peuvent être isolés et servir de matrice à la synthèse d'ADNc correspondants pour constituer une banque

d'ADNc, la méthode la plus classiquement utilisée est basée sur l'emploi du bactériophage lambda. Le bactériophage comporte une molécule d'ADN linéaire encapsidée qui permet d'insérer de grands fragments d'ADN étranger. La banque d'ADNc est ensuite introduite dans des cellules hôtes, des bactéries (*Escherichia coli*). Après étalement sur une boîte de pétri, les bactéries infectées par ces phages recombinants produisent de très nombreux phages avant d'être lysées. Le produit de l'expression correspondant à l'insert d'ADNc, l'allergène, est alors transféré sur une membrane qui sera incubée avec l'anticorps spécifique radioactif (IgE de sérum de sujets allergiques ou anticorps polyclonaux ou monoclonaux d'animaux). La liaison spécifique de l'anticorps avec l'allergène révèle la position des clones de l'ADNc qui expriment cet allergène [8]. Ces clones sont alors récupérés sur la boîte de pétri d'origine. Après une étape d'amplification, l'isolement et le séquençage de l'ADN permettent de vérifier la séquence codante pour l'allergène avant son intégration dans un système d'expression.

3.2. Expression de l'allergène recombinant

À partir de la séquence d'ADN codant pour l'allergène, l'expression de la protéine nécessite le clonage dans un vecteur d'expression, qui peut être un plasmide ou un virus jouant le rôle de transporteur, puis le transfert du vecteur dans une cellule hôte, qui permettra la synthèse de la protéine via le gène inséré. L'expression des allergènes peut se faire soit dans des systèmes procaryotes (bactéries), soit dans des systèmes eucaryotes (levures, cellules d'insecte, plantes). Il existe un certain nombre de différences entre ces systèmes [10]. Le choix de l'hôte est principalement guidé par les modifications post-traductionnelles que la protéine doit subir pour être biologiquement active, tandis que le choix du vecteur génétique utilisé pour exprimer le gène recombinant doit être adapté à l'hôte choisi.

3.2.1. Expression en système procaryote

La bactérie est généralement l'hôte de choix pour produire une protéine recombinante en raison de ses propriétés de croissance rapide, de son faible coût de production et de sa génétique très étudiée. Parmi les bactéries, *E. coli*, qui a été très étudiée depuis les années 1960, est sans doute le système d'expression le plus populaire [11]. Sa génétique est très bien connue, elle est facile à cultiver par fermentation, ce qui peut être réalisé à grande échelle en bioréacteur. Pour exprimer un gène dans *E. coli*, celui-ci doit être inséré dans un vecteur de clonage, un plasmide. Ce plasmide d'expression doit contenir certains éléments en particulier une origine de réplication pour permettre au gène d'être répliqué dans la bactérie, un promoteur pour initier la transcription du transgène et un marqueur de sélection pour aider la sélection des bactéries ayant incorporé le vecteur. L'un des avantages de l'utilisation d'*E. coli* comme cellule hôte est qu'il existe de nombreux outils de clonage commerciaux permettant de jouer sur le type de protéines exprimées, que ce soit sous forme de protéines de fusion possédant une séquence « tag », séquence d'acides aminés fusionnée permettant la purification ultérieure, ou sous

forme de protéines non fusionnées. De plus, la protéine recombinante peut représenter jusque 25 % des protéines totales produites dans *E. coli*, ce qui correspond à un bon rendement de plusieurs grammes par litre d'allergène recombinant produit. Toutefois, la surproduction d'une protéine étrangère et le fait que *E. coli* possède un faible potentiel de sécrétion peuvent entraîner leur agrégation dans un état le plus souvent biologiquement inactif, appelé corps d'inclusion [12]. Un autre facteur contribuant à la formation d'agrégats ou de ces corps d'inclusion est le fait que la cellule procaryote est incapable d'effectuer les modifications post-traductionnelles que la protéine requiert pour atteindre sa conformation native. La formation de ces corps d'inclusion nécessitera une stratégie de resolubilisation/repliement *in vitro* pour obtenir l'allergène recombinant sous forme soluble. Ainsi, l'utilisation d'*E. coli* présente quelques limites : la faible sécrétion de l'allergène produit qui peut compliquer sa purification ultérieure, l'absence de modifications post-traductionnelles telles que la glycosylation ou la formation de ponts disulfures du fait de l'environnement réducteur du cytoplasme conduit le plus souvent à des problèmes de repliement ou modifie l'activité de l'allergène. Ces derniers points peuvent s'avérer critiques dans le cas de certains allergènes recombinants pour leur utilisation en diagnostic lorsque la réponse IgE repose sur la conformation native de la protéine ou pour étudier la structure de ces allergènes si l'activité enzymatique native est absente. D'autres bactéries peuvent être utilisées qui ont l'avantage de présenter une capacité de sécrétion supérieure à *E. coli* mais dont la génétique est moins bien connue et le rendement de production de protéines plus faible. Le système d'expression *E. coli* est donc adapté pour la production de protéines non glycosylées ou pour lesquelles la glycosylation n'est pas nécessaire. Ce système permet la production d'une grande quantité de protéine soluble correctement repliée.

3.2.2. Expression en système eucaryote

Le principal avantage des systèmes eucaryote est leur capacité à effectuer des modifications post-traductionnelles, parmi lesquelles l'adressage par des peptides signaux, le repliement, la formation de ponts disulfures ou la glycosylation. Parmi les systèmes eucaryotes disponibles, les plus utilisés pour l'expression d'allergènes recombinants sont les levures, les cellules d'insecte et de plante.

Les levures, organismes unicellulaires, possèdent l'avantage de se cultiver simplement et de présenter une croissance rapide. *Pichia pastoris* et *Saccharomyces cerevisiae* sont les deux espèces utilisées, avec une préférence pour *P. pastoris* qui permet d'obtenir une production de protéines dix à 100 fois plus importante [13]. De plus, *S. cerevisiae* peut conduire à une hyperglycosylation des protéines exprimées, alors que ce phénomène semble moindre avec *P. pastoris*. Moins de résidus mannose sont ajoutés dans chaque N-glycosylation (huit à 14 par chaîne alors que *S. cerevisiae* en comporte 50 à 150). *P. pastoris* est capable d'effectuer une O-glycosylation ou une N-glycosylation. Cependant, l'importante glycosylation de l'allergène recombinant le rend très différent de l'allergène naturel et peut engendrer des difficultés à son utilisation en

diagnostic ou en thérapeutique. *P. pastoris* possède néanmoins un bon rendement d'expression (environ 5 à 200 mg par litre) et un potentiel de sécrétion naturelle élevée. Il existe de nombreux vecteurs disponibles pour ces deux espèces.

L'expression d'allergènes recombinants dans des cellules d'insectes comme cellule hôte utilise des vecteurs développés à partir de *Baculovirus*, qui est un virus lytique spécifique des larves de lépidoptères. Les virus utilisés dérivent généralement de *Autographa californica nuclear polyhedrosis virus* (AcNPV), possédant un ADN bicaténaire capable d'intégrer un transgène de 10 kb. L'infection de la cellule hôte, la lignée cellulaire Sf9 de *Spodoptera frugiperda*, permet la production de l'allergène recombinant [8]. Les protéines produites subissent des modifications post-traductionnelles presque complètes et peuvent être excrétées. Le rendement d'expression est élevé (environ 1 g/L de culture), néanmoins, le coût de la production est élevé.

Enfin, les cellules végétales, plantes entières ou lignées cellulaires, peuvent être utilisées comme système d'expression d'allergènes recombinants. Les vecteurs utilisés sont des virus, comme le virus de la mosaïque du tabac (TMV) qui se réplique dans la cellule hôte de plante (*Nicotiana benthamiana*) et dans laquelle est synthétisée la protéine. Les plantes peuvent réaliser la plupart des glycosylations et permettent d'obtenir un repliement satisfaisant. Le rendement de production peut être élevé (0,5 à 2 mg/g de feuille de tabac). La production dans la plante entière peut nécessiter des étapes d'extraction et de purification plus longues pour extraire la protéine des tissus.

3.3. Extraction et purification de l'allergène recombinant

Lorsque la protéine est sécrétée par la cellule productrice, la purification se fait à partir du milieu de culture tandis que lorsque la protéine est intracellulaire, il est nécessaire de détruire les structures cellulaires pour l'en extraire. La purification peut se faire par les techniques de filtration sur gel, de chromatographie d'échanges d'ions ou d'affinité. L'expression de protéines de fusion possédant une séquence « tag », séquence d'acides aminés ajoutés en fusion à la protéine d'intérêt, permet d'améliorer la production, la détection, la stabilité ou la purification. Par exemple, l'ajout d'une séquence homopolymérique d'histidine (His-tag) crée un site de haute affinité pour les cations divalents tel que le nickel ou le cobalt, qui permet à la protéine de fusion de s'accrocher sur une colonne d'affinité et de faciliter la purification de l'allergène. Cette séquence, de petite taille, pourra persister sur la protéine. L'utilisation de séquences plus longues peut nécessiter d'insérer entre les deux peptides une séquence qui va coder pour un site de coupure. Ce site de coupure permettra d'enlever la fusion entre les deux protéines après leur expression *in vitro*.

4. Caractérisation de l'allergène recombinant

La caractérisation du produit final est nécessaire tant sur le plan physicochimique que sur le plan de l'activité immunologique pour valider la qualité de l'allergène recombinant [14]. La caractérisation biochimique et biophysique de l'allergène permet le contrôle conformationnel de la protéine

produite. Un ensemble de techniques biochimiques telles que l'électrophorèse ou l'analyse de la séquence en aminoacides, et biophysiques telles que la spectrométrie de masse ou le dichroïsme circulaire, peuvent être conduites afin de tester l'intégrité, l'homogénéité et la qualité des structures protéiques. Les tests d'activités biologiques, dépendant de l'allergène produit, sont très variés comme par exemple la mesure de l'activité enzymatique ou la capacité de liaison d'un ligand à son site de fixation. Enfin, dans le cas de l'expression dans un système procaryote, la concentration en endotoxines peut être déterminée. La validation de l'activité immunologique des allergènes recombinants repose sur l'utilisation de tests *in vitro* et *in vivo*. L'évaluation de la capacité de liaison aux IgE, de la dégranulation des cellules effectrices se fera par des tests *in vitro*. L'évaluation de l'allergénicité se fera chez l'homme par la réalisation de tests cutanés et de tests de provocation.

5. Conclusion

La production d'allergènes recombinants nécessite de porter une attention particulière au choix du système d'expression en prenant en considération les avantages et les inconvénients de chacun et en se basant sur les connaissances du degré de glycosylation, la présence de ponts disulfures, la séquence N-terminale en aminoacides ou les propriétés de stabilité de la protéine. Les allergènes produits devront être validés par rapport à l'allergène naturel purifié, lorsque celui-ci a été isolé. Ainsi, les allergènes recombinants, en vue de leur application en diagnostic et en thérapeutique, permettent la production de préparations allergéniques de qualités pharmaceutiques constantes.

Conflit d'intérêt

Aucun.

Références

- [1] Thomas WR, Stewart GA, Simpson RJ, Chua KY, Plozza TM, Dilworth RJ, et al. Cloning and expression of DNA coding for the major house dust mite allergen Der p 1 in *Escherichia coli*. *Int Arch Allergy Appl Immunol* 1988;85:127–9.
- [2] Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D, Rumpold H, et al. The gene coding for the major birch pollen allergen Bet v 1, is highly homologous to a pea disease resistance response gene. *EMBO J* 1989;8:1935–8.
- [3] Tovey ER, Baldo BA. Comparison by electroblotting of IgE-binding components in extracts of house dust mite bodies and spent mite culture. *J Allergy Clin Immunol* 1987;79:93–102.
- [4] Eraso E, Martinez J, Garcia-Ortega P, Martinez A, Palacios R, Cisterna R, et al. Influence of mite growth culture phases on the biological standardisation of allergenic extracts. *J Investig Allergol Clin Immunol* 1998;8: 201–6.
- [5] Valenta R. Recombinant allergen-based concepts for diagnosis and therapy of type I allergy. *Allergy* 2002;57:66–7.
- [6] Valenta R, Niederberger V. Recombinant allergens for immunotherapy. *J Allergy Clin Immunol* 2007;119:826–30.
- [7] Thomas WR. Advent of recombinant allergens and allergen cloning. *J Allergy Clin Immunol*. doi:10.1016/j.jaci.2010.12.1084 (in press).
- [8] Schmidt M, Hoffman DR. Expression system for production of recombinant allergen. *Int Arch Allergy Immunol* 2002;128:264–70.
- [9] Cramer R, Janussi R, Menz G, Blaser K. Display of expression products in cDNA libraries on phage surfaces: versatile screening system for selective isolation of genes by specific gene-product/ligand interaction. *Eur J Biochem* 1994;226:53–8.
- [10] Demain AL, Vaishnav P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* 2009;27:297–306.
- [11] Wallner M, Gruber P, Radauer C, Maderegger B, Susani M, Hoffmann-Sommergruber K, et al. Lab scale and medium scale production of recombinant allergens in *Escherichia coli*. *Methods* 2004;32:219–26.
- [12] Betton JM, Chaffotte A. Replément et production de protéines recombinantes. *Med Sci* 2005;21:613–7.
- [13] Pokoj S, Lauer I, Fötisch K, Himly M, Mari A, Enrique E, et al. *Pichia pastoris* is superior to *E. coli* for the production of recombinant allergenic non-specific lipid transfer proteins. *Protein Expr Purif* 2010;69:68–75.
- [14] Cromwell O, Suck R, Kahlert H, Nandy A, Weber B, Fiebig H. Transition of recombinant allergens from bench to clinical application. *Methods* 2004;32:300–12.