

CHAPITRE II :
INGÉNIERIE DES PROTÉINES ;
LA CONCEPTION RATIONNELLE ET ÉVOLUTION
DIRIGÉE

1. Introduction :

L'ingénierie protéique a pour but de modifier de façon plus ou moins importante la structure des protéines afin d'optimiser leurs propriétés fonctionnelles.

Deux techniques étaient principalement utilisées en vue de l'optimisation d'une protéine d'intérêt : La mutagenèse dirigée (*rational design* en anglais) et l'évolution dirigée (Fig. 1).

Les paramètres à améliorer sont variés : Par exemple, on peut obtenir une enzyme qui a une thermo-stabilité accrue ; optimiser une protéine thérapeutique; obtenir un peptide qui se lie avec une forte affinité à un ligand donné; créer *in vitro* un anticorps contre à peu près n'importe quel ligand, pour une utilisation en diagnostic; créer un ribozyme avec une nouvelle activité catalytique ; obtenir une enzyme à la fois résistante à la chaleur et à ou élargir la gamme de pH dans laquelle l'enzyme est efficace tout en augmentant son activité. Enfin les enzymes permettent aujourd'hui de remplacer certains procédés chimiques lourds et polluants par des procédés beaucoup plus respectueux de l'environnement (on parle de chimie verte).

Parmi les protéines thérapeutiques qu'il serait intéressant d'améliorer, on peut notamment citer les hormones, les cytokines, les interférons, les vaccins, et les anticorps.

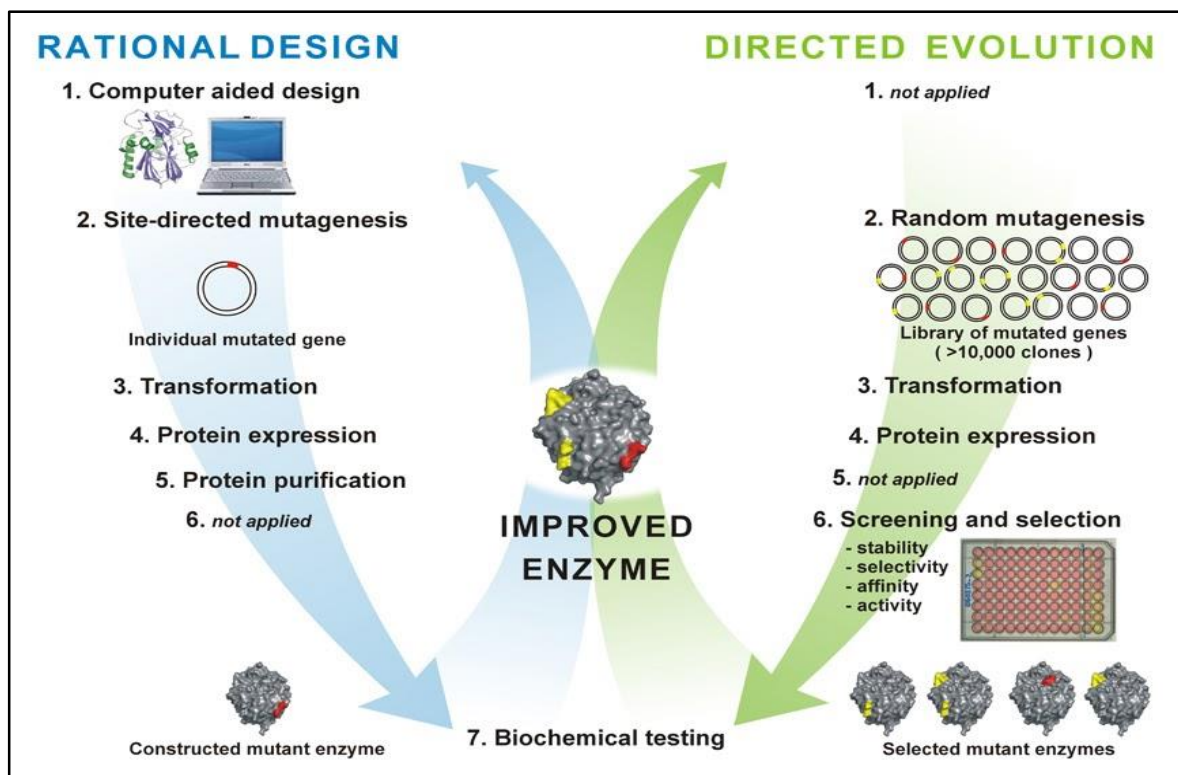


Fig. 1 : Méthodes d'ingénierie des protéines.

2. Ingénierie rationnelle ou « Protein Design »

La conception rationnelle de protéines ou « Protein Design » est une approche qui peut être relativement longue à mettre en œuvre, car elle nécessite une parfaite connaissance de la structure 3D de la protéine, et surtout du rôle précis de chaque acide aminé. Les mutants sont choisis après modélisation, réalisés, exprimés puis testés.

Dans l'ingénierie rationnelle les mutants sont créés par mutagenèse dirigée ;

2.1. Mutagenèse dirigée ou mutagenèse ponctuelle

La mutagenèse, en générale, est une technique visant à modifier artificiellement la séquence nucléotidique d'un fragment d'ADN, dans le but de modifier l'activité biologique qui en découle.

La mutagenèse dirigée (SDM) est une méthode *in vitro* de création d'une mutation dans une séquence connue. Elle est souvent effectuée par des méthodes basées sur la PCR. Typiquement, une ou deux bases sont modifiées dans la mutagenèse dirigée (une seule base ou un codon afin de modifier la séquence protéique sur un acide aminé, sans toucher à l'intégrité du gène). Ces mutations sont introduites à l'aide d'amorces oligonucléotidiques qui sont préalablement synthétisées. Ces amorces sont complémentaires à la séquence sauvage à l'exception des nucléotides qui doivent être mutés.

Ces techniques exigent par contre une bonne connaissance de base de la structure ou du mécanisme de la spécificité de substrat afin de pouvoir cibler les résidus dont la mutation peut être informative

Le principe de toute mutation introduite par PCR repose sur l'observation qu'une stricte complémentarité de l'amorce nucléotidique avec la séquence de l'ADN cible n'est pas absolument nécessaire sur toute la longueur considérée. Elle est obligatoire du côté 3' pour l'amorçage de la réaction, mais pas du côté 5'. Les mutations sont introduites dans les produits et peuvent être de trois types :

- a) Des substitutions : un ou plusieurs nucléotides sont remplacés par un même nombre de nucléotides différents.
- b) Des insertions : un ou plusieurs nucléotides sont ajoutés à la séquence cible.
- c) Des délétions : un ou plusieurs nucléotides sont enlevés de la séquence.

2.1.1. Mutagenèse par extension enzymatique d'une amorce synthétique portant la mutation.

Dans cette méthode, on fabrique un oligonucléotide (par synthèse chimique) qui contient la mutation flanquée de 10 à 15 nucléotides de séquence non mutée. Cet oligonucléotide "mutagène"

est hybridé à la séquence complémentaire sauvage présente dans l'ADN monocaténaire préparé à partir d'un clone dans un phage ou un phagemide. Il se forme donc un ADN hétéroduplexe. Cet oligonucléotide sert alors d'amorce pour la synthèse enzymatique d'ADN in vitro par une polymérase qui convertit l'ADN monocaténaire sauvage en une forme bicaténaire. L'ADN hétéroduplexe ainsi obtenu contient donc un brin sauvage et un brin muté. Cet ADN est ligaturé et transformé chez *E. coli*. Lors de la première réplication, on obtient donc un plasmide sauvage et un plasmide muté (qui se répliqueront à leur tour pour donner un mélange de plasmides sauvages et mutants). Pour discriminer ces deux types de plasmides (ou enrichir les cultures en clones mutants), un grand nombre de méthodes sont possibles : nous étudierons ici une de ces techniques : mutagenèse dirigée selon Stratagene.

Dans la mutagenèse dirigée selon Stratagene (Fig. 2), la façon de procéder nécessite une seule étape de PCR avec 2 amorces contenant des codons mutés. Dans ce cas, les séquences des amorces sont en partie complémentaires aux brins opposés du plasmide et contiennent des mésappariements au niveau des codons à muter. Le plasmide est amplifié par PCR produisant ainsi des plasmides avec des mésappariements ou « mismatch » à la position désirée. Le plasmide sauvage ayant servi de matrice est éliminé par digestion avec une enzyme de restriction (*DpnI*). Cet enzyme nécessite un site de reconnaissance méthylé pour être actif ce qui est le cas de l'ADN matrice mais pas de l'ADN amplifié. Un séquençage permet de vérifier la présence de la mutation.

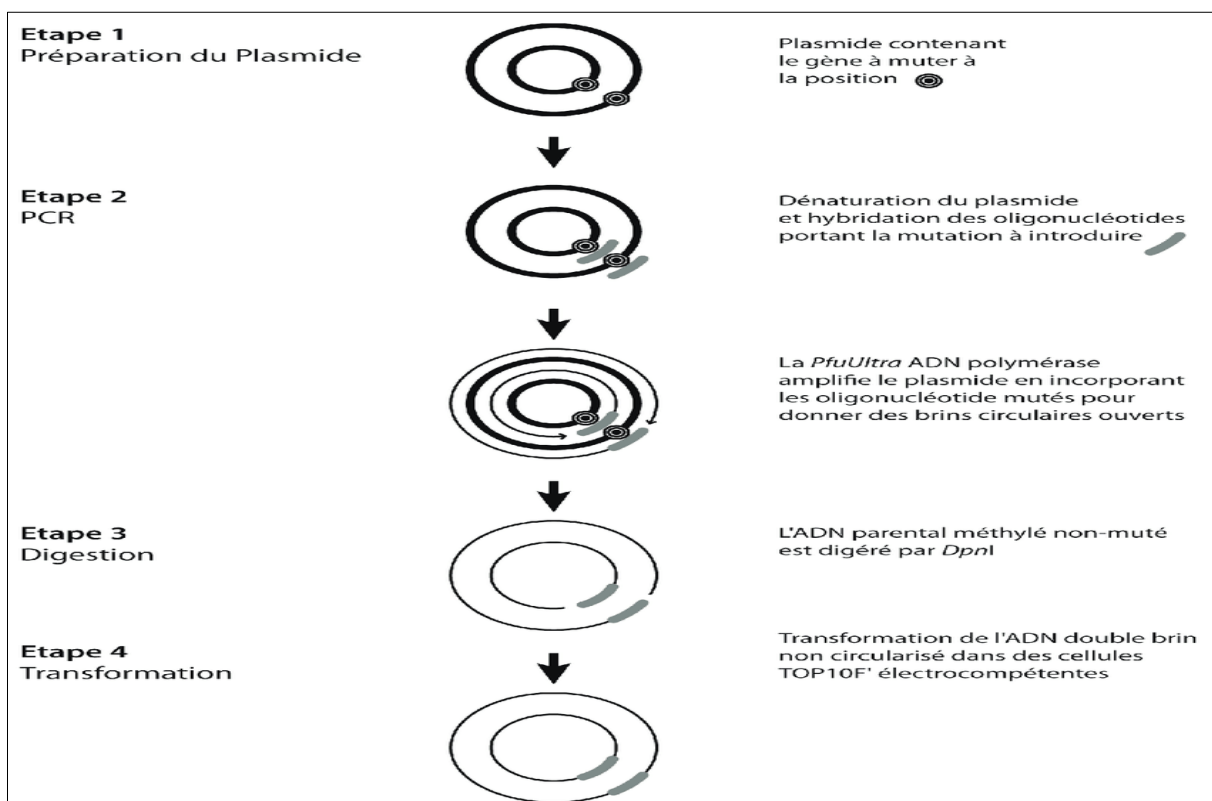


Fig.2 : Principe de la mutagenèse dirigée selon Stratagene. Figure adaptée d'après le manuel QuickChange® Site-Directed Mutagenesis Kit de Stratagene.

2.1.2. Mutagenèse par PCR (méthode des extrémités chevauchantes; overlap extension method)

La deuxième façon de procéder est la méthode des extrémités chevauchantes (overlap extension method). Elle nécessite deux amplifications du gène d'intérêt avec deux paires d'amorces. La première paire (2-3) contient les codons portant la mutation et les amorces vont s'hybrider formant un mésappariement (mismatch) alors que la deuxième paire (1-4) est entièrement complémentaire à la séquence cible. Les 2 PCR sont réalisées indépendamment et sont, par la suite, mélangées. Après dénaturation et réhybridation, une population d'ADN double-brins hétérogènes hybridés au niveau de la région mutée et comprenant des bouts cohésifs, est générée. Les bouts manquants sont, alors, complétés par simple polymérisation à l'aide de l'ADN polymérase. Le produit obtenu est amplifié par PCR avec les amorces (1-4) (Fig. 3).

Une dernière PCR utilisant les amorces **3** et **4** est alors réalisée sur ce mélange et donne, en effet, un fragment de taille souhaitée.

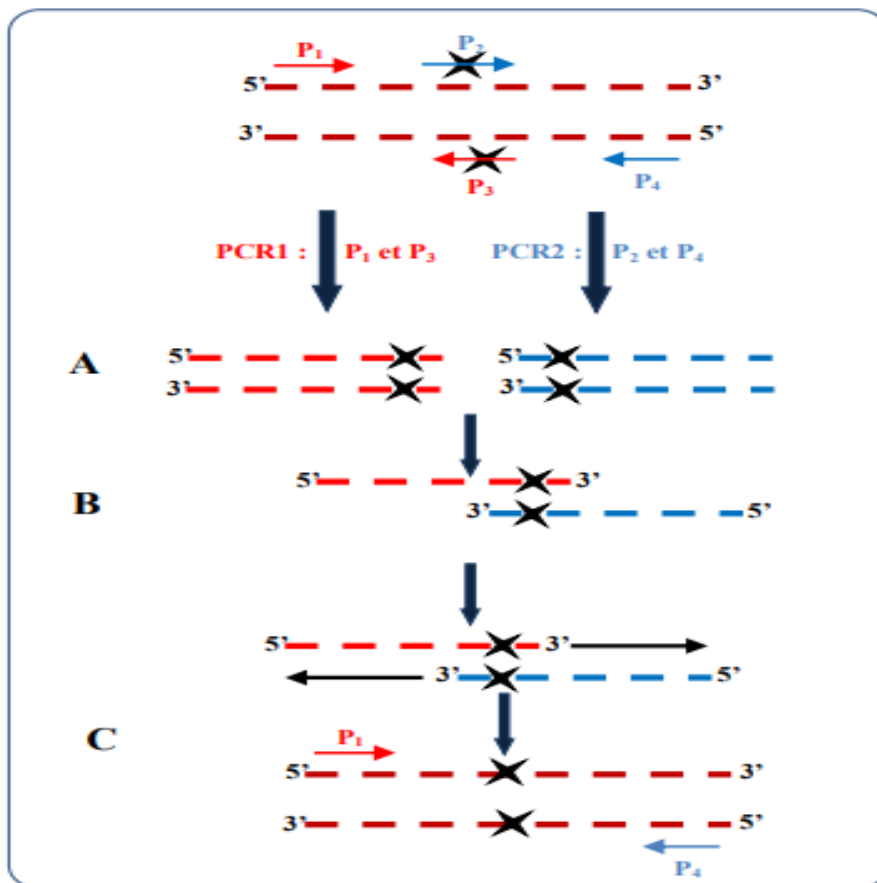


Fig. 3 : Schéma général de la méthode des extrémités chevauchantes : A : Les 2 PCR réalisées séparément avec les 2 paires d'amorces (1-3) d'une part et (2-4) d'autre part. B : Dénaturation et réhybridation des deux produits de PCR produisent une population de molécules d'ADN hétérogènes qui

sont par la suite complétées par l'ADN polymérase. C : La dernière étape d'amplification par les amorces (1-4) produit des séquences de départ avec les mutations codées par les amorces 2 et 3.

En résumé, la technique rationnelle de mutagenèse dirigée présente l'avantage d'être relativement rapide mais ne permet pas de définir une image globale de la relation structure fonction, ne serait-ce même que dans une famille de gènes. Les mutations observées sont majoritairement spécifiques à un seul substrat ou un seul enzyme. C'est pour cela que les techniques d'évolution dirigée ont été développées, afin d'augmenter les informations recueillies en augmentant le nombre de mutants générés.

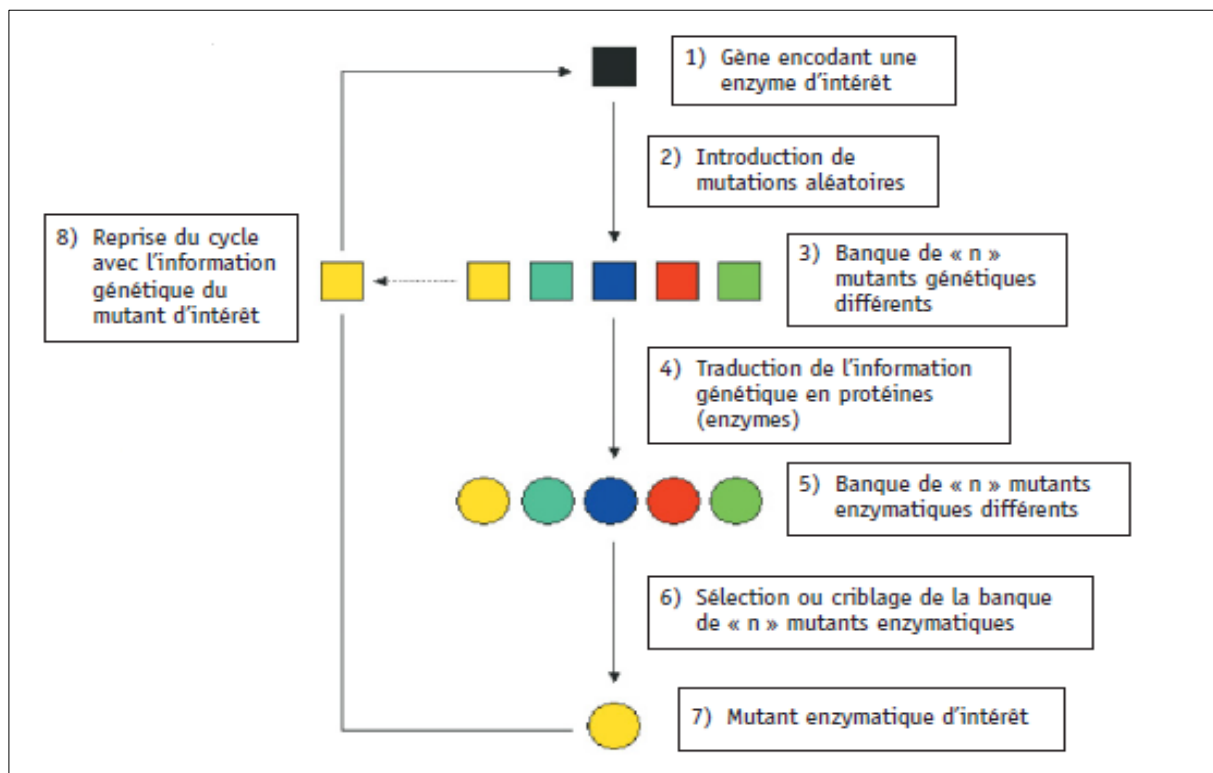
2. Evolution dirigée

2.1. Définition

L'évolution dirigée est un ensemble de technologies permettant d'améliorer une protéine ou un acide nucléique, en reproduisant artificiellement le processus naturel de l'évolution mais en cherchant à l'orienter dans une direction choisie. Les approches d'évolution dirigée nécessitent deux étapes essentielles :

- une étape de création de la diversité génétique
- une étape de criblage ou de sélection des variants présentant des propriétés recherchées

La première étape de création de diversité consiste à créer à partir d'un gène initial une population de gènes codants pour des formes modifiées de la protéine d'intérêt. Ces modifications peuvent être totalement aléatoires par leur position et leur nature ou bien aléatoires mais localisées à des positions particulières. Elles peuvent aussi résulter d'un processus de recombinaison de séquences



préalablement modifiées. On appelle banque, ou bibliothèque, la population ainsi obtenue. Les grandes étapes de ces technologies sont exposées dans la figure 4.

Fig. 4 : Schéma de l'évolution dirigée.

2.2. L'évolution moléculaire dirigée

Grâce aux outils de la biologie moléculaire, il est possible de mimer en laboratoire les processus à l'origine de l'évolution naturelle, ce qui a permis de modifier les propriétés des molécules biologiques telles que les protéines, ce qui aurait été sans doute impossible à obtenir autrement. Les deux composantes de l'évolution naturelle sont reproduites artificiellement : la génération de la diversité est obtenue en agissant sur l'ADN (génotype) et la population est conduite artificiellement à évoluer par la sélection des mutants ayant intégré le caractère d'intérêt (phénotype). L'objectif essentiel de ce processus est de parvenir à générer des mutants ayant la caractéristique recherchée et puis de parvenir à isoler l'information génétique correspondante et à l'amplifier.

La toute première protéine qui a été sujette à l'évolution dirigée est une forme de β -galactosidase.

2.3. Les techniques d'évolution moléculaire dirigée des protéines

La réussite du processus d'évolution réside dans la bonne combinaison entre **les méthodes de génération de diversité** et celles qui permettent de **sélectionner ou de cribler les mutants** ayant incorporés la propriété voulue.

Pour mimer l'évolution *in vitro*, le processus naturel doit être accéléré pour que la diversité soit créée puis criblée/sélectionnée dans un temps compatible avec les exigences expérimentales. Les méthodes de génération de diversité les plus communément utilisées sont la mutagenèse aléatoire et la recombinaison. A titre d'exemple on peut citer : le traitement de l'ADN ou de la bactérie par des agents mutagènes chimiques, la PCR mutagène (error-prone PCR, epPCR), la RCA mutagène (rolling circle error-prone PCR, epRCA) et la mutagenèse à saturation.

Le choix d'une stratégie ou d'une autre pour atteindre un but choisi n'est pas soumis à des règles précises, mais chaque méthode comporte des avantages particuliers et des limites.

2.3.1. Mutagenèse aléatoire

La mutagenèse aléatoire vise à introduire dans un fragment d'ADN des substitutions de nature et de position non contrôlées.

a) Mutagenèse chimique :

Historiquement, la mutagenèse aléatoire était réalisée au moyen de procédés chimiques altérant la structure de l'ADN. Cette méthode implique l'utilisation d'agents chimiques tels que :

- **L'Ethylméthane sulfonate (EMS)** qui agit sur les bases guanines ce qui induit des erreurs lors de la réplication de l'ADN.
- **L'acide nitrique (HNO₂)** qui agit en désaminant les adénines et les cytosines causant une mutation ponctuelle par transversion (A/T en G/C).
- **Le bisulfite de sodium** qui agit spécifiquement sur la cytosine d'une région d'ADN simple brin. Il a été utilisé par l'équipe d'Ermakova-Gerdes pour muter aléatoirement et *in vitro* le gène psbDI.

b) « Error Prone » PCR : epPCR :

Cette technique est basée sur une amplification par PCR de la séquence à modifier mais dans des conditions favorisant l'augmentation de la fréquence globale des erreurs de la DNA polymérase (Fig.5). De très nombreux documents ont décrit des techniques de mutagenèse aléatoire basées soit sur l'utilisation d'une polymérase modifiée ayant structurellement une faible fidélité de réplication, soit sur l'utilisation d'une polymérase non modifiée, mais dans des conditions particulières d'amplification menant à un fort taux de mutation. Le taux d'erreur peut être augmenté par l'ajout de Mn²⁺ (diminution de l'efficacité de l'hybridation des bases) ou par l'augmentation de la concentration en Mg²⁺ dans le but de stabiliser les paires de bases non complémentaires.

Dans un cas comme dans l'autre, l'enzyme introduit des mutations au fil des cycles ; à la fin de la réaction, de nombreuses copies de la molécule initiale sont obtenues, chacune de ces molécules comportant une ou plusieurs mutations différentes.

Après la PCR, la population des séquences modifiées doit être insérée dans un plasmide de façon à obtenir un ensemble de clones constituant la bibliothèque. Dans l'epPCR, les mutations ponctuelles sont les plus fréquentes mais des délétions et donc des décalages du cadre de lecture peuvent aussi avoir lieu.

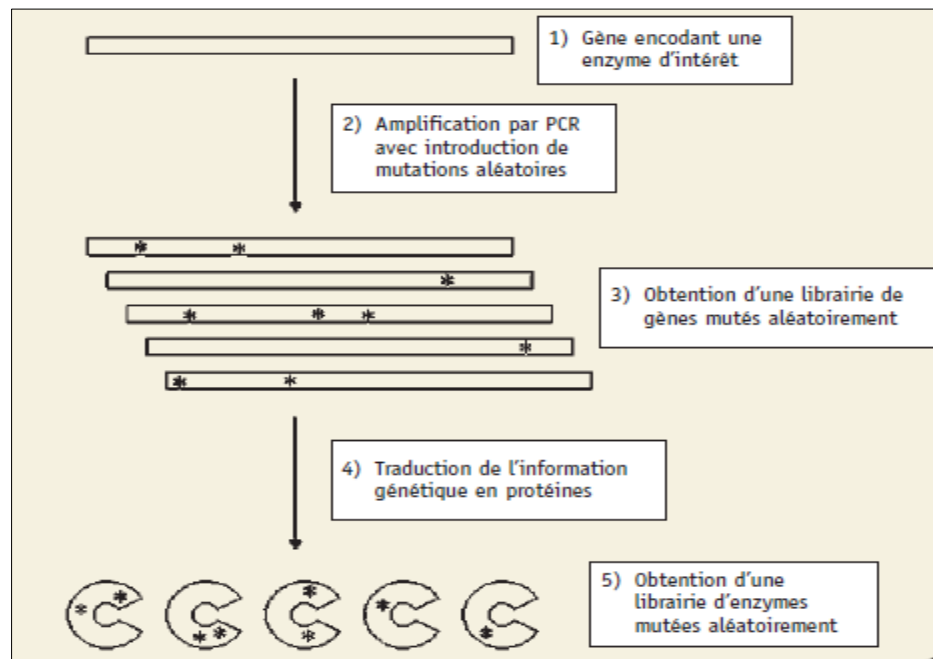


Fig.5 : Mutagenèse aléatoire par PCR pro-erreurs

Exemple :

Un exemple, impliquant cette technique, est l'ingénierie de la sous-unité HycE de l'hydrogénase-3 d'*E. coli* pour laquelle, 7 variants améliorés ont été obtenus. La production d'hydrogène obtenue par la meilleure séquence est multipliée 17 fois relativement à celle obtenue avec la protéine initiale.

Une autre procédure, MutaGen, est basée sur l'utilisation des ADN polymérases Pol β et Pol η (η) à très faible fidélité. Cette technique a été appliquée pour muter un gène codant l'amylosucrase. Les auteurs ont déterminé que les mutants obtenus par l'ADN polymérase Pol β sont 4 à 7 fois moins mutés que ceux obtenus par la polymérase Pol η , et que les spectres de mutations obtenus par l'une et l'autre sont complémentaires. La polymérase Pol η a engendré des modifications comprenant des substitutions de bases et des délétions de codons, ce qui est rarement obtenu par d'autres méthodes.

Une stratégie aussi intéressante, intitulée « frame shuffling » nécessite l'utilisation de l'ADN polymérase de famille Y qui introduit des taux importants de mutations « frame shift » pour changer les cadres de lectures des gènes parentaux. Les mutants produits montrent des échanges de segments de séquences et des propriétés physicochimiques différentes de celles des mutants obtenus par des méthodes de mutations conventionnelles.

2.3.2. Mutagenèse à saturation (Site Saturation mutagenesis : SSM)

La mutagenèse à saturation est basée sur l'utilisation d'oligonucléotides (amorce) dégénérés : Lors de la synthèse des oligonucléotides, il est facile de créer une dégénérescence à n'importe quel endroit de la séquence de l'oligonucléotide (amorce) : au niveau de la position souhaitée, on utilise alors non pas une base unique, mais un mélange équimolaire de plusieurs bases. Dans une séquence, on note conventionnellement N le mélange équimolaire des quatre bases. Par exemple, ATN correspond à 25% de ATA, 25% de ATT, 25% de ATG et 25% de ATC. Un oligonucléotide contenant deux positions dégénérées est en fait constitué d'un mélange équimolaire de 16 oligonucléotides. Un oligonucléotide avec trois positions dégénérées N est constitué d'un mélange équimolaire de 64 oligonucléotides.

La SSM peut être réalisée de deux manières différentes : soit en une seule étape de PCR soit en 2 étapes de PCR (les deux méthodes expliquées dans la partie mutagenèse dirigée mais en utilisant des amorces dégénérées).

2.3.3. La mutagenèse par mélange d'ADN (La recombinaison ou DNA shuffling)

La mutagenèse par mélange d'ADN consiste à recombinaison des séquences partiellement homologues, isolées à partir de différents organismes. Par exemple, si l'on travaille sur une enzyme, la première étape d'une approche par mélange d'ADN reviendra à isoler un grand nombre de gènes homologues à cette enzyme (à partir de collections de souches par exemple). Différentes approches existent alors pour mélanger les domaines de ces gènes homologues et générer une banque de molécules d'ADN « chimériques » (Fig.6), c'est-à-dire constituées de plusieurs domaines ayant des provenances différentes. Il est attendu que ces molécules contiennent ainsi des caractéristiques nouvelles, et notamment des propriétés cumulées de deux ou plusieurs gènes parentaux.

Ainsi, par exemple, si l'on dispose au départ de deux gènes homologues d'une même enzyme, l'un connu pour être très actif, et l'autre pour être thermostable (ce dernier ayant par exemple été isolé à partir d'un organisme thermophile), on peut espérer que certaines des molécules obtenues par mélange de ces deux gènes, - contenant alors certains domaines du premier et certains autres du second -, auront des propriétés cumulées de forte activité et de thermostabilité (une telle additivité n'est pas évidente et pas toujours assurée mais est en pratique assez fréquemment observée). Dans certains cas, non seulement des propriétés cumulées, mais également des propriétés nouvelles (par exemple une activité supérieure à celle de tous les gènes naturels de départ), ont été obtenues par mélange de gènes.

Ces approches de mélange d'ADN se sont révélées être particulièrement efficaces dans le domaine de l'amélioration des enzymes. Par contre, cette approche n'est pas adaptée au domaine des

protéines thérapeutiques, d'une part parce que les polymorphismes humains sont relativement limités, et d'autre part parce qu'il n'est guère concevable, pour des raisons d'immunogénicité notamment, de penser que le mélange d'ADN à partir de protéines de différentes espèces puisse apporter un bénéfice notable en thérapeutique humaine.

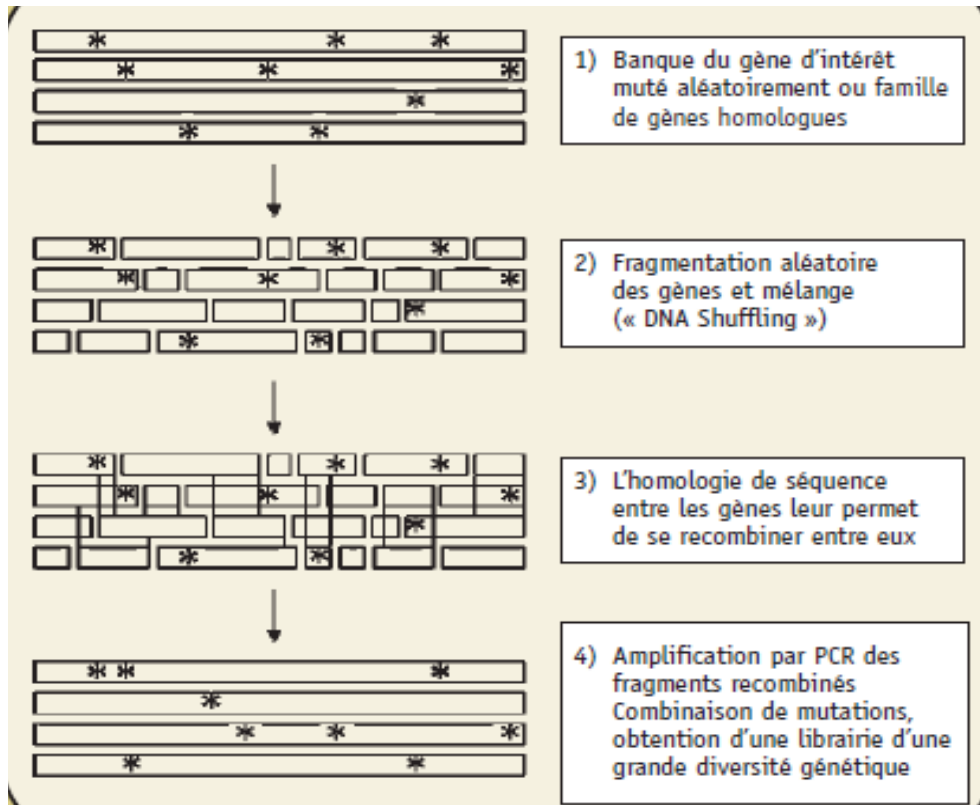


Fig. 6 : DNA shuffling

Une approche a été d'utiliser une matrice exogène similaire mais pas identique aux gènes mélangés pour hybrider les morceaux issus de la fragmentation (méthode RACHITT, Fig.7)

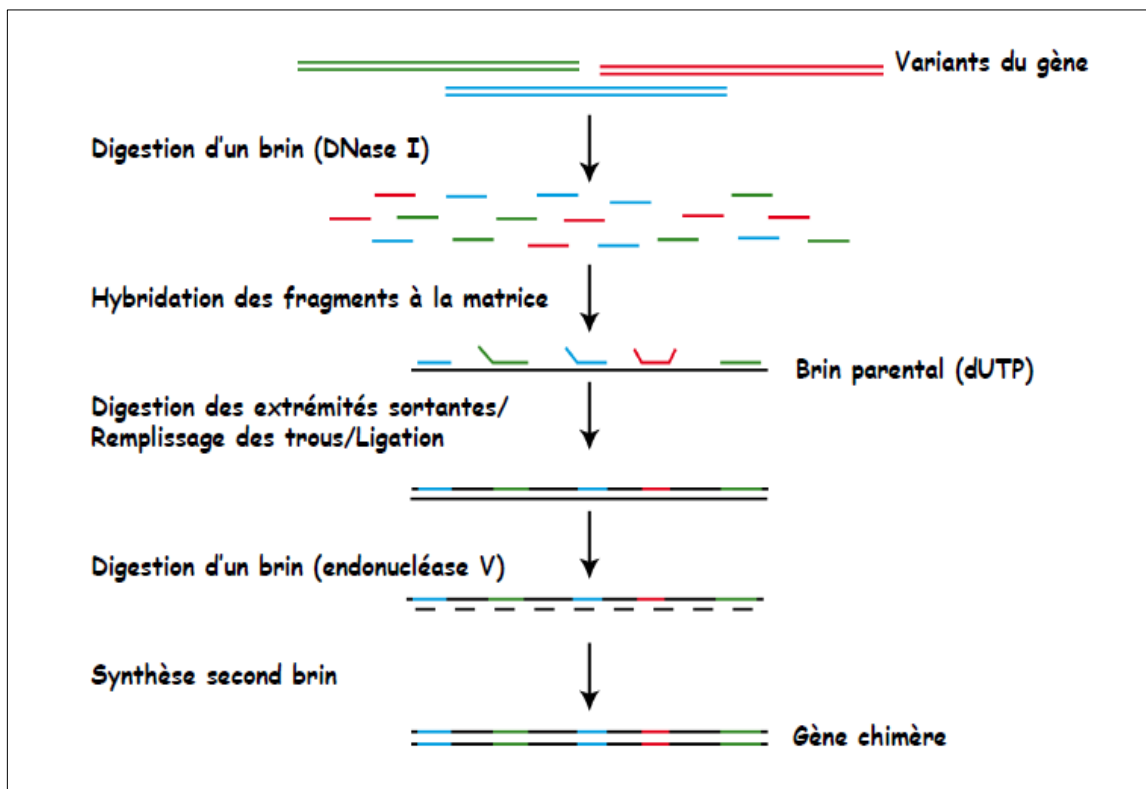


Fig. 7 : méthode RACHITT

La deuxième étape du processus d'évolution dirigée est l'étape de sélection des variants ayant incorporé la propriété d'intérêt. Vous allez voir les méthodes de sélection dans le prochain chapitre.

Références :

EP1544296A1, European Patent Office. *Procédé de mutagenèse dirigée massive.*
<https://patentimages.storage.googleapis.com/ff/49/66/67cba44c01b4b7/EP1544296A1.pdf>

Ryckelynck, M. (2009). *Evolution dirigée de macromolécules biologiques.*
<http://www.lbgi.fr/~ripp/PublicDirectory/Vrac/MBMRyckelynck2009.pdf>

Doucet, N. et Pelletier, J. N. *L'évolution dirigée : la recherche d'enzymes toujours plus performantes*
http://www.profs.inrs.ca/ndoucet/pdf/2004a_CQVB.pdf

Guellouz, A. *Création par évolution dirigée de protéines artificielles en alternatives aux anticorps.* Thèse de doctorat. Université Paris Sud 11. <https://tel.archives-ouvertes.fr/tel-00767675/document>