

**Directed evolution of a DD-peptidase
into a β -lactamase**

UCL

**Université
catholique
de Louvain**

Heykel Trabelsi

Institut des Sciences de la Vie

Université catholique de Louvain

A thesis submitted for the academic degree of PhD in Sciences

October 2014

Composition du jury

Pr. François Chaumont (Président du jury)
Université catholique de Louvain

Pr. Patrice Soumillion (Promoteur)
Université catholique de Louvain

Pr. Pascal Hols
Université catholique de Louvain

Pr. Xavier De Bolle
Université de Namur

Pr. Magali Remaud
INSA-Toulouse (France)

Pr. Bernard Joris
Université de Liège

A mes Parents

A mes frères

A ma belle soeur

A ma famille

A ma future femme

A mes futurs enfants

Rerciements

Voilà, c'est la fin de cette expérience humaine hors norme. Si on me demande de refaire la même chose, je n'hésiterais pas une seconde. Au cours de ces années de doctorat, j'ai mené un rythme de vie complètement exotique pour quelqu'un qui vient de la rive sud de la méditerranée. Une expérience qui restera gravée dans mon esprit toute ma vie. Une expérience qui m'a totalement métamorphosé. Je ne suis plus la même personne qui a quitté Sfax pour Marseille, ni celle qui a pris le TGV de Marseille pour débarquer au pays de Tintin. Cette expérience de doctorat m'a permis de voyager intellectuellement et spirituellement pour me découvrir. Ce que j'ai appris le long de ce voyage n'a pas de prix.

Je tiens donc à remercier toutes les personnes qui m'ont aidé de loin ou de près à arriver à la destination finale. Je commence par mon promoteur le Professeur Patrice Soumillion qui m'a donné la confiance et la chance d'être un membre de son équipe. Merci pour ce projet de thèse très enrichissant, merci de m'avoir donné la liberté et l'autonomie pour travailler à mon aise, merci de m'avoir continuellement motivé en me disant qu'on pourrait faire un «science» avec tes résultats. Merci pour ton encouragement et ton soutien aux moments difficiles surtout pendant la rédaction. Tes qualités de scientifique ne cessent de m'impressionner et arriver un jour à ton niveau c'est l'un de mes objectifs. Je te considère maintenant comme un grand frère.

J'aimerais également remercier le Professeur Jacques Fastrez pour ses remarques et ses discussions intéressantes pendant les «lab meetings» des premières années de mon doctorat.

Un grand merci aux membres de jury pour le temps qu'ils m'ont consacré pour lire et juger mon travail.

Merci pour ma comité d'accompagnement, les professeurs Bernard Hallet, Pierre Morsomme et Xavier Debolle, d'être au rendez vous quand j'avais besoin de vous.

Ce long voyage n'a pas pu être intéressant sans croiser des gens exceptionnels et bizarres à la fois.

Je voudrais remercier les membres de quatre mousquetaires dont je faisais partie. Merci à Jean Martin (JM) pour les moments au Guinch Bar, à Barcelone ou autour d'une table de Poker où je te laisse gagner discrètement pour embêter Val. Merci à Valéry (Val) pour ta présence, ton humour, tes moments de dépression et ton amitié. Merci à mon ami Bruno, mon vrai compagnon le long de cette bataille de thèse. Merci d'être toujours présent à chaque fois que j'ai eu besoin de ton aide ou de tes discussions et aussi pour courir les 20 km de Brussels ensemble.

Je remercie aussi Gabrielle et Geneviève les inséparables pour les bons moments qu'on a passé ensemble.

Merci à Laurence pour l'ambiance qu'elle crée au labo, pour m'avoir initié à des différents goûts musicaux et de son soutien moral.

Merci à Véronique, notre grande sœur, pour tout ce que tu as fait pour me faciliter la vie. Rien qu'une simple discussion avec toi de tout et de rien, me faisait plaisir.

Merci à Emanuele de m'avoir poussé à travailler sur les phages après son départ et à Vaidas pour les notebooks bien détaillés qui a laissé pour moi.

Je tiens de tout mon cœur à remercier Anastassia, mon alliée, pour tous les moments qu'on a passé ensemble, l'encouragement mutuel au cours de la rédaction, les discussions scientifiques et parfois politiques sur l'URSS et l'avenir de la Russie (je m'intéressais seulement à l'avenir des jolies russes) et les jours qu'on a passé en Angleterre avec Steph.

Un grand merci à Stéphanie (Steph) ma consultante «mode et Shopping» pour les agréables moments qu'on a passé d'un magasin à un autre. Aller courage pour la rédaction, tu arriveras ☺

Il ne faut surtout pas oublier Nadia, notre jolie postdoc, je voudrais te remercier pour ton soutien, tes discussions et tes conseils. Je n'oublierais jamais les moments sushi ensemble et la fête de réveillon avec Colombo (Vincent). Bon courage pour la suite.

J'ai également croisé une personne qui a vécu dans la même chambre que moi à Marseille (foyer de Luminy) mais 2 ans avant mon arrivée. Je peux rédiger un livre de 200 pages sur cette personne. Le fameux Pierre Galka, avec qui j'ai discuté plein de théories

d'astrophysique alors qu'on est pas spécialistes du domaine, avec qui j'ai inventé plein de blagues délirantes et incompréhensibles, avec qui j'ai parlé de la géopolitique de l'europe centrale pendant des heures. Pierre, merci.

Merci à Mary, l'amie et collègue libanaise, pour tous les gâteaux qu'elle a apporté au labo et pour les moments passés ensemble autour d'un café ou d'une manip qui ne marche pas.

Merci à Gilles, mon consultant Macintosh, pour les softwares et les astuces qui m'ont aidé à optimiser l'utilisation de mon Mac. Bon chance pour la fin de ta thèse.

Vielen danke für Julia, du bist großartig. Thanks for giving me the opportunity to practice my german. Thanks for your friendship and good luck with your thesis ☺

Olivier, mon mémorant puis mon collègue, merci pour ton amitié, ton sens de la reponsabilité et pour m'avoir initié au monde de jeux de société avec Philippe.

Merci à Thomas et Michael, mes mémorants, je vous souhaite une bonne continuation.

Un grand merci à toute la famille BBGM (groupes de Pascal Hols et Bernard Hallet) pour l'ambiance agréable qui règne au labo.

Un grand merci à mon ami Malek pour l'aventure de doctorat qu'on a mené ensemble, les moments de fou rire à chaque fois qu'on prend l'avion pour Tunis et aux moments cappucini à Rome. Je voudrais remercier aussi mon gang de tunisiens à LLN pour les moments partagés qui nous ont permis de ne pas oublier d'ou on vient. Merci à Helmi, Slim, Bilel, Sami, Anouar, Hédi, Haykel, Houcine, Rania, Maissa, Emna et Mondher.

Merci l'ISV

Merci l'UCL

Merci la Belgique

Abbreviations and Symbols

Amino-Acid	3 Letters	1 Letter
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

DNA
 RNA
 Amp
 PenG
 dATP
 dCTP
 dGTP
 dTTP
 dNTP
 GuHCl

Desoxyribonucleic acid
 Ribonucleic acid
 Ampicillin
 Penicillin G
 DesoxyAdenosine Triphosphate
 DesoxyCytosine Triphosphate
 DesoxyGuanosine Triphosphate
 DesoxyThymidine Triphosphate
 DesoxyNucleotide Triphosphate
 Guanidinium Chloride

Kan	Kanamycine
Tet	Tetracyclin
Gluc	Glucose
IPTG	Isopropyl- β -thiogalactoside
LB	Luria Bertani
PBP	Penicillin binding protein
PCR	Polymerase chain réaction
PDB	Protein Data Bank
Rmsd	Root mean square deviation
Rpm	Rotation per minute
SDS	Sodium dodecyl sulfate
D-Ala	D-Alanine
KDa	Kilo Dalton
MIC	Minimal inhibitory concentratin
Mg ²⁺	Magnesium bivalent cation
Mn ²⁺	Manganese bivalent cation
s	second
min	minute
h	hour
ml	millilitre
mM	millimolar
nm	nanomolar
bp	Base pair
IMAC	Immobilized metal ion affinity chromatography

Thesis summary

Understanding how proteins or more specifically enzymes evolve is still a challenging scientific issue. This thesis aimed at finding an artificial evolutionary trajectory between phylogenetically related proteins with the objective of uncovering some unknown aspects of the birth of new catalytic properties. It is largely accepted that serine β -lactamases, responsible of β -lactam antibiotics hydrolysis, have evolved from some ancestral DD-peptidases involved in the peptidoglycan metabolism. DD-peptidases are also called Penicillin Binding Proteins because they are covalently inhibited by β -lactams such as penicillins. In this work, a cyanobacterial DD-peptidase (PBP-A), sharing sequence similarities with class A β -lactamases, was used as a starting scaffold to engineer a β -lactamase activity. Our evolutionary scenario hypothesized a loss of the native peptidase activity to become exclusively a Penicillin-Binding Protein as a required step before evolving toward a β -lactamase function. For this purpose, highly diverse libraries of PBP-A mutants were built and drifted without any selective pressure for catalytic activity. Interestingly, libraries were enriched with clones sharing specific mutations in the active site region. The biochemical characterization of a clone with these mutations showed around 300 fold decrease in the recognition of the native substrate (peptidoglycan precursor) while maintaining the capability to react with penicillin. These two mutations were also responsible of the fitness cost readjustment of the protein. Starting from these neutralized mutants, subsequent rounds of mutagenesis and drifts under sub-lethal then lethal concentrations of β -lactam antibiotic allowed the selection of clones with an ampicillin resistance phenotype. The kinetic parameters estimation of the best mutant have shown an increase in the β -lactamase activity by at least 3 orders of magnitude.

This work have shown for the first time, how the neutralization of an enzyme after the gene acquisition promotes its neofunctionalization.

Table des matières

Introduction.....	17
I. Aperçu historique de l'évolution dirigée	19
II. Aperçu technique de l'évolution dirigée.....	23
III. Exemples de « success story » de l'évolution dirigée.....	41
VI. Le modèle général de l'évolution des protéines	46
V. Contraintes de l'évolution des protéines.....	48
VI. Evolution par duplication de gène et transfert horizontal.....	54
VII. Résistance aux antibiotiques et coût de fitness.....	57
VIII. DD-peptidases et β -lactamases.....	59
IX. Tentatives de conversion d'une PBP en β -lactamase	73
Objectives and Strategy	83
I. Objectives	85
II. Strategy	87
Results	89
Generation of PBP-A diversity and evaluation of fitness problems within the libraries	93
I. Abstract	93
II. Introduction.....	95
III. Results and Discussion	97
IV. Conclusion	108
V. Materials and methods.....	109
Neutralization of an Enzyme Promotes Its Neofunctionalization	115
I. Abstract	116
II. Introduction.....	117
III. Results and discussion.....	118
IV. Conclusion	132
Supporting information	133
V. Methods	136
Phenotypically active PBP variants : Attempts of purification and characterization.....	143
I. Abstract	143
II. Introduction.....	145
III. Results and Discussion	147
IV. Conclusion	167
V. Materials and methods.....	168
Conclusions and Perspectives	175
I. General conclusions	177

II. Perspectives.....	181
References.....	183
Appendix.....	195
Side project contribution.....	197

Introduction

I. Aperçu historique de l'évolution dirigée

Durant toute son existence, l'être humain confronté à la nature et à sa complexité illimitée, a cherché à la comprendre, la domestiquer et l'utiliser en sa faveur. Une simple lecture de l'histoire humaine nous montrera à quel point l'homme a réussi à contourner et surmonter des difficultés et des défis imposés par la nature. Il a su l'utiliser en exploitant ses ressources pour améliorer son quotidien et même son confort.

A l'aube de la civilisation, l'humanité a bien compris que la nature évolue et change constamment. Tous les organismes vivants qui peuplent la planète sont les survivants qui ont réussi à s'adapter à leur environnement. Cette conscience d'une évolution permanente pourrait être illustrée par les pratiques millénaires mises en place par l'homme pour, par exemple, améliorer et domestiquer la faune et la flore. Certes, les mécanismes sous-jacents qui mènent à l'introduction de la variabilité et des propriétés dans les populations et les espèces n'étaient pas connus mais l'ingéniosité des humains à développer des techniques de sélection est très ancienne.

Ces dernières décennies, grâce à l'explosion des technologies du vivant, l'homme est devenu capable de manipuler les molécules biologiques, dont l'ADN. Les progrès en biologie moléculaire lui ont permis de contrôler, modifier et même inventer des êtres vivants avec des caractéristiques prédéfinies.

Avec la compréhension et la polémique autour de la théorie de l'évolution des espèces décrite par Charles Darwin en 1859 dans son livre « On the origin of species », les scientifiques ont essayé, parfois avec succès, d'exploiter les processus évolutifs au laboratoire mais dans une échelle temporelle beaucoup plus courte que celle de la nature. C'est au début des années soixante que se sont apparus les premiers signes d'une nouvelle ère de biologie moléculaire, l'ère de l'évolution moléculaire.

Cette discipline utilise un ensemble de techniques et d'outils permettant de reproduire un processus évolutif au laboratoire en imitant l'évolution naturelle. Elle vise à créer des nouvelles entités biologiques (organismes, cellules, molécules) dotées de propriétés nouvelles ou améliorées (1). De nos jours, le terme d'évolution dirigée est préféré à celui d'évolution moléculaire car il souligne l'objectif des chercheurs consistant à faire évoluer une entité biologique choisie, généralement une protéine, dans une direction choisie également. Au niveau expérimental, le processus d'évolution dirigée est divisé en deux étapes : la première consiste à introduire des mutations aléatoires dans un gène cible pour créer une collection de mutants la plus diverse possible et la deuxième étape consiste à y rechercher les variants qui présentent la propriété souhaitée par une méthode de sélection ou de criblage appropriée. Si nécessaire, les variants identifiés peuvent ensuite être le point de départ d'un nouveau cycle d'évolution (2) (Figure 1).

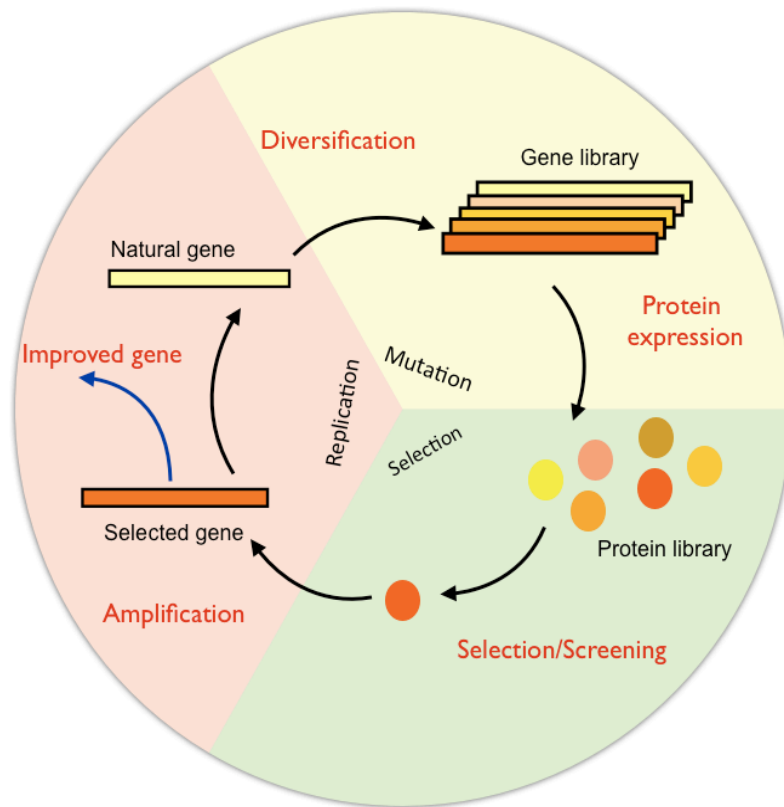


Figure 1: Schéma d'évolution dirigée d'une protéine. Le gène codant pour la protéine de départ subit une mutagenèse aléatoire pour créer une banque de gènes, une fois exprimées les protéines d'intérêt sont isolées à l'aide d'une méthode de sélection ou de criblage pour obtenir le ou les meilleurs variants. Le ou les gènes sélectionnés peuvent servir comme point de départ pour un second tour d'évolution.

L'évolution dirigée, utilisée comme approche de prédilection dans tous les laboratoires qui s'intéressent à l'ingénierie des protéines, a vu le jour grâce à la succession de plusieurs découvertes surtout d'ordre technique. Au début des années soixante, la mise en évidence de composés chimiques mutagènes, capables de réagir avec l'ADN et muter quelques bases a permis aux scientifiques de les exploiter. En effet, avec ces réactifs, introduire intentionnellement des mutations dans un organisme avec des fréquences

élevées est devenu faisable. En 1964, ceci a notamment permis à Lerner et ses collaborateurs de conférer, à la bactérie *Aerobacter aerogenes*, la capacité d'utiliser le xylitol après une mutagénèse chimique aléatoire, ceci dans une tentative de comprendre l'émergence de nouvelles fonctions (3). Un peu plus tard, les premières expériences d'évolution dirigée extracellulaires ont été réalisées par le groupe de Spiegelman : afin de montrer les potentialités évolutives d'un système d'acides nucléiques auto-réplicatifs, des répliques ARN-dépendantes ont été reconstituées *in vitro* pour faire évoluer des molécules d'ARN par réplication durant plusieurs générations et avec la réduction du temps de biosynthèse de l'ARN comme pression de sélection. Après 74 générations, la molécule auto-répliquative s'est débarrassé de 83% de sa séquence pour devenir la plus petite entité auto-répliquative connue (4-6). Plus tard, en 1973, Campbell et ses collaborateurs ont observé l'émergence d'une nouvelle activité β -galactosidase dans une souche d'*E. coli* dépourvue du gène LacZ (7). Ce travail a mis en évidence au laboratoire le pouvoir de la sélection naturelle, capable de façonner le destin d'une population de bactéries. En effet, la culture d'*E. coli* privée d'une activité β -galactosidase dans un milieu contenant du lactose a mené à la sélection d'une autre protéine codée par un autre gène pour survivre. Cinq mutations ont été détectées au niveau d'un gène encodant un enzyme immuno-génétiquement et cinétiquement différent de la β -galactosidase encodée dans l'opéron lactose. Cette protéine a donc évolué vers une nouvelle β -galactosidase fonctionnelle.

II. Aperçu technique de l'évolution dirigée

Comme déjà mentionné, toute expérience d'évolution dirigée est composée de deux étapes indispensables. La première est la création de la diversité génétique et la deuxième est la recherche des mutants présentant la propriété souhaitée. Dans cette partie, nous allons décrire plus en détails les différentes méthodes et stratégies possibles pour réaliser cela.

II.1. Création de la diversité

II.1.1. Agents chimiques mutagènes

La méthode la plus ancienne et la plus simple pour introduire des mutations de manière aléatoire dans l'ADN est de traiter un organisme avec des composés mutagènes (8). Ces composés vont induire des erreurs d'incorporation au niveau de la réplication. Le grand handicap de cette méthode est que les mutations ne ciblent pas un gène bien défini mais tout le génome, ce qui rend souvent très difficile l'analyse des résultats de sélection. Les chances d'obtenir plus qu'une mutation dans un même gène, ce qui pourrait être requis pour voir l'émergence d'une nouvelle propriété, sont également extrêmement faibles avec cette méthode. C'est pourquoi, de nos jours, les agents chimiques sont abandonnés en faveur de nouvelles approches plus performantes et mieux ciblées.

II.1.2. Utilisation de souches mutagénisantes

La suppression chez *E. coli* des systèmes de réparation des erreurs génétiques (mutations) a été utilisée pour augmenter le nombre de mutations spontanées. A titre d'exemple, la souche XL1-Red présente des altérations des gènes *mutT* et *mutS* impliqués dans la réparation de mésappariements au sein de l'ADN (9), et aussi au niveau de *mutD* encodant une sous-unité de la polymérase III responsable de la fidélité de l'holoenzyme (activité d'édition exonucléase 3'-5') (10). Le taux de mutations de cette souche est 5000 fois

plus élevé que celui de la souche sauvage. Les chercheurs ont utilisé ce type de souche pour générer des nouveaux phénotypes. L'inconvénient est que les mutations touchent tout le génome et non pas un gène précis. De ce fait, la souche est instable génétiquement et la détermination de l'origine des nouveaux phénotypes est très compliquée.

Cette technique est donc moins utilisée aujourd'hui mais suscite un regain d'intérêt suite aux récents travaux du groupe de Liu qui présente un système d'évolution dirigée et continue par mutagenèse *in vivo* d'un phage porteur du gène d'intérêt (11). Dans leur système appelé PACE (Phage-Assisted Continuous Evolution), un lagon (chambre de culture) est utilisé en batch continu pour faire évoluer une protéine encodée dans le génome d'un phage filamenteux de type M13 (Figure 2). Un taux de mutagenèse élevé est assuré par un plasmide de mutagenèse qui code pour une polymérase sans activité d'édition (dnaQ926) (12) sous le contrôle d'un promoteur inductible à l'arabinose (araBAD) et deux gènes *umuC* et *umuD* (polymérase V) impliqués dans la réponse SOS qui permettent la fixation des mutations. L'arabinose est introduit uniquement dans le lagon qui est continuellement alimenté en bactéries génétiquement stables (sans arabinose). Les bactéries ne deviennent instables génétiquement qu'une fois dans le lagon où la mutagenèse *in vivo* est déclenchée. Comme le génome, le phage est alors soumis à cette mutagenèse mais ce dernier pourra être transféré vers les nouvelles bactéries arrivantes par infection. En effet, le phage est continuellement empaqueté en particules infectieuses, grâce à la présence d'un plasmide assistant encodant la protéine pIII essentielle à l'infectivité du phage (« helper »). Les bactéries arrivantes sont donc continuellement infectées. Un flux continu renouvelle sans cesse les bactéries, ce qui permet de cibler les mutations sur le phage uniquement.

Leur travail a permis avec succès l'évolution de l'ARN polymérase du phage T7 en mutants capables de reconnaître d'autres régions promotrices. La pression de sélection a été simplement installée en plaçant le gène de la

protéine pIII sous le contrôle du nouveau promoteur cible de l'ARN polymérase de T7. Ainsi, seuls les enzymes mutants capables de transcrire le gène de la protéine pIII permettront la production de particules virales infectieuses elles-mêmes capables de se maintenir dans le lagon. Le caractère continu de l'expérience permet de fortement accélérer le processus évolutif par rapport aux approches classiques (voir plus loin) qui séparent l'étape de mutagenèse de l'étape de sélection ou de criblage. Grâce au lagon, 200 tours de mutagenèse ont été effectués en 8 jours.

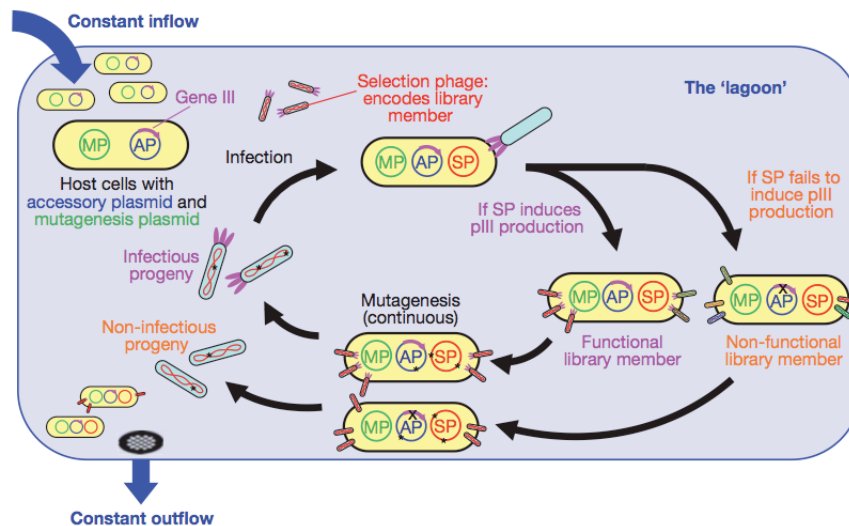


Figure 2: Vue globale du système PACE. Les cellules hôtes arrivent continuellement au lagon où elles seront infectées par les phages de sélection (SP) qui codent pour la banque de mutants. Les membres fonctionnels de la banque induisent la production de pIII du plasmide accessoire (AP) pour libérer une progéniture de phages capables d'infecter les nouvelles cellules hôtes. La mutagenèse est assurée par l'induction du plasmide de mutagenèse (MP). Les cellules hôtes quittent le lagon à une vitesse moyenne plus rapide que celle de la réplication ce qui limite l'accumulation des mutations seulement aux phages en réplication. Figure reprise de la référence (11).

II.1.3. Usage d'oligonucléotides dégénérés

Cette technique est utilisée généralement quand une courte région de la protéine est ciblée (boucle, site de reconnaissance du substrat, site actif). Elle

consiste à synthétiser chimiquement des oligonucléotides partiellement dégénérés (synthèse combinatoire) pour ensuite les introduire de manière site-spécifique dans un gène. La méthode la plus simple est d'utiliser une dégénérescence complète des codons de type NNN (N correspondant à un mélange équimolaire des bases A, T, C et G) où chacun des 64 codons est représenté de manière équivalente (13). Il est cependant plus judicieux de créer des banques biaisées. En effet, la dégénérescence NNN couvre l'entièreté du code génétique mais génère une représentation favorisant certains acides aminés et comprenant 5% de codons stop (3/64). Pour mieux équilibrer cette représentation, les chercheurs utilisent des codons de type NNK ou NNS (K = G ou T et S = C ou G). Cette dégénérescence représente 32 codons encodant les 20 acides aminés mais avec un meilleur équilibre de représentation et seulement 3% de codons stop (1/32).

D'autres dégénérescences, telles que NDT, KMT ou même TMT (D = A ou G ou T, M = A ou C), sont aussi utilisées. Par exemple, pour réduire le nombre de clones à cribler tout en gardant une grande chance de trouver des clones d'intérêt (Tableau 1), il peut être avantageux d'utiliser la dégénérescence NDT qui encode 12 acides aminés représentant la variété physicochimique des chaînes latérales (14).

Tableau 1: Comparaison des nombres de mutants à cribler pour couvrir 95% des résidus possibles pour 1 à 5 codons NNK ou NDT. Si seulement 4 positions sont rendues aléatoires, pour une banque de type NNK permettant les 20 AA, il faut cribler plus de 3 millions de clones alors que pour les banques de type NDT (12 AA) il suffit de cribler environ 50 fois moins. Tableau adapté de Reetz et ses collaborateurs (14).

Surveying 20 AA		Surveying 12 AA	
<i>No. of AA to be randomized</i>	<i>No. of mutants to be screened at 95% library coverage</i>	<i>No. of AA to be randomized</i>	<i>No. of mutants to be screened at 95% library coverage</i>
1	94	1	34
2	3,066	2	430
3	98,163	3	5,175
4	3,141,251	4	62,118
5	100,520,093	5	745,433

II.1.4. PCR mutagénisante ou error prone PCR (epPCR)

C'est la technique la plus utilisée en ingénierie des protéines. Elle permet de cibler des mutations aléatoires sur n'importe quel fragment d'ADN amplifiable par PCR, typiquement un gène complet. Le taux de mutation est plus ou moins contrôlable et sera choisi en fonction de la stratégie ou des objectifs du chercheur.

En epPCR, la fidélité de la synthèse de l'ADN par la polymérase est intentionnellement réduite. Généralement on utilise la Taq polymérase qui est dépourvue d'activité exonucléase et qui a un taux d'erreur 10 fois supérieur aux enzymes à haute fidélité telle que la Pfu polymérase. Les conditions expérimentales peuvent également être modifiées pour augmenter encore plus la fréquence des erreurs. Ainsi, ajouter du manganèse (Mn^{2+})

permet de réduire la spécificité d'appariement des paires de bases et d'augmenter le taux d'erreur (environ 25 fois avec 0,7 mM de MnCl_2) (15). D'autre part, l'utilisation d'un mélange non stoechiométrique des quatre dNTPs provoque également des mésincorporations. Le plus souvent, des concentrations déséquilibrées de dNTPs sont utilisées en combinaison avec MnCl_2 (16, 17). Des analogues de nucléotides comme le 8-oxo-dGTP ont été aussi utilisés pour perturber l'appariement des paires de bases et induire des mutations (18). Finalement, le nombre de mutations augmente également avec la concentration de MgCl_2 et le nombre de réplifications qui peut être contrôlé par la concentration d'ADN matrice et le nombre de cycles d'amplification.

II.1.5. DNA Shuffling

Cette approche introduite par Stemmer en 1994 est devenue un outil de mutagenèse largement utilisé par les chercheurs, leur permettant d'explorer des régions de l'espace de séquences non atteintes auparavant (16,17).

La technique est également connue sous le nom de PCR sexuelle et est illustrée à la Figure 3. Elle tire avantage de la recombinaison de fragments d'ADN homologues. La première étape consiste à préparer l'ADN de départ, soit en réalisant une epPCR pour créer une diversité initiale, soit en groupant des gènes homologues (partageant plus de 70 % d'identité de séquences) appartenant à différents organismes. La seconde étape est une fragmentation de l'ADN en petits morceaux de 10 à 50 pb par traitement à la DNase I. A ce stade, il est également possible d'ajouter des oligonucléotides dégénérés pour favoriser l'incorporation de mutations dans des régions spécifiques. Les fragments et éventuels oligonucléotides synthétiques sont ensuite rassemblés et recombinaisonnés en réalisant une PCR sans amorces. A la fin, une PCR classique avec des amorces terminales permet d'amplifier les gènes réassemblés. Le mélange final est cloné pour constituer une banque de

mutants. La méthode permet notamment le transfert des mutations bénéfiques entre séquences issues de gènes naturellement homologues ou provenant d'une expérience d'évolution dirigée préalable.

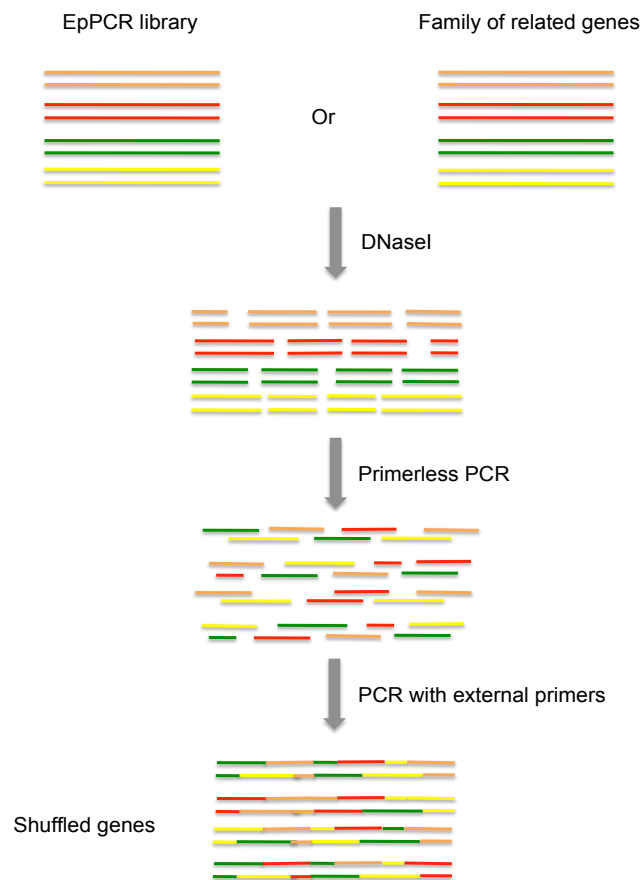


Figure 3: Méthode de DNA Shuffling pour l'évolution dirigée. Après une fragmentation aléatoire, les fragments d'ADN sont rassemblés par PCR sans amorces, puis une dernière PCR avec deux amorces externes permettent de reconstituer le gène.

II.2. Sélections

Chez les organismes vivants, l'association entre le génotype et le phénotype assurée par la compartimentation cellulaire est essentielle pour l'évolution. En évolution dirigée, les chercheurs utilisent cette propriété naturelle mais ont également développé des stratégies innovantes de criblage ou de sélection basées sur des couplages artificiels entre gènes et protéines encodées. Dans ce chapitre, nous présenterons les différentes approches de sélection ainsi que les technologies qui permettent de les mettre en œuvre.

II.2.1. Sélection *in vivo*

Il est dans certains cas possible de faire émerger *in vivo*, au départ d'une banque de mutants, les clones présentant une propriété recherchée. Pour ce faire, il faut que la propriété confère un avantage biologique significatif à l'hôte du gène. Généralement, l'avantage sélectif peut être apporté par inactivation d'un composé toxique ou par production d'un composé utile pour la croissance. Une grande variété d'activités catalytiques peut ainsi être directement associée à un processus sélectif. Par exemple, en 1994, Stemmer a pu générer un mutant de Tem1 β -lactamase capable de conférer une résistance à la cefotaxime 32000 fois supérieure à la bactérie hôte. Ces clones ont été obtenus après 3 tours de « DNA shuffling » suivis de 2 tours de « back crossing » (« shuffling » avec le gène de départ permettant d'éliminer les mutations délétères), et en augmentant progressivement la concentration de l'antibiotique au cours des tours de sélections (20).

D'autres équipes ont également développé des stratégies de sélection très ingénieuses pour coupler la propriété recherchée à un processus sélectif. Ainsi, le groupe de Peter Schultz, a utilisé une approche très simple de sélection basée sur la suppression de codons stop pour étendre le code génétique et permettre l'incorporation d'acide aminés non naturels (UAAs) dans les protéines. Pour cela, ils ont fait évoluer un couple d'aaRS/tRNA (aaRS pour aminoacyl-tRNA synthétase) pour que le tRNA soit chargé par

un acide aminé non naturel et qu'il supprime un codon stop introduit dans le gène d'intérêt à la position voulue. Une sélection en deux étapes a été mise au point à cet effet. Après la transformation d'une banque de tRNA supresseurs dans *E. coli* contenant déjà un gène toxique avec un codon stop dans son cadre ouvert de lecture, seuls les clones exprimant des tRNA mutants qui ne sont pas substrats des aaRS endogènes peuvent survivre. Les tRNA mutants sélectionnés ont été par la suite produits dans des cellules d'*E. coli* exprimant une aaRS hétérologue et contenant un gène encodant une β -lactamase mais avec un codon stop à une position qui tolère n'importe quel résidu d'acide aminé. Dans cette sélection positive, la présence d'ampicilline dans le milieu tue tous les clones contenant des tRNA non chargés et permet la survie des clones ayant des tRNA aminoacylés exclusivement par l'aaRS hétérologue (21, 22). Le nouveau couple tRNA/aaRS peut alors être qualifié d'orthogonal. Une sélection similaire en deux étapes a ensuite été conçue pour sélectionner des aaRS mutantes capables d'aminoacyler le tRNA orthogonal avec un acide aminé non naturel (Figure 4). Une banque de aaRS présentant des mutations aléatoires de résidus positionnés au niveau du site de fixation du substrat a été produite dans des cellules d'*E. coli* exprimant le tRNA orthogonal et un gène encodant la chloramphénicol acetyltransferase (CAT) ayant un codon stop. Dans cette sélection positive, la croissance cellulaire a été réalisée en présence de chloramphénicol et de l'acide aminé non naturel d'intérêt. A ce stade, seulement les clones ayant des mutants d'aaRS capables d'aminoacyler le tRNA orthogonal avec l'acide aminé non naturel ou les acides aminés endogènes ont pu survivre. Ces mutants ont été ensuite exprimés dans des bactéries qui produisent le tRNA et qui contiennent un gène toxique avec un codon stop. Dans cette sélection négative, la croissance a été réalisée en absence de l'acide aminé non naturel de manière à éliminer les mutants d'aaRS qui chargent le tRNA orthogonal avec des acides aminés endogènes (production du gène toxique). Les

survivants produisent donc des aaRS spécifiques pour un acide aminé non naturel bien défini (23).

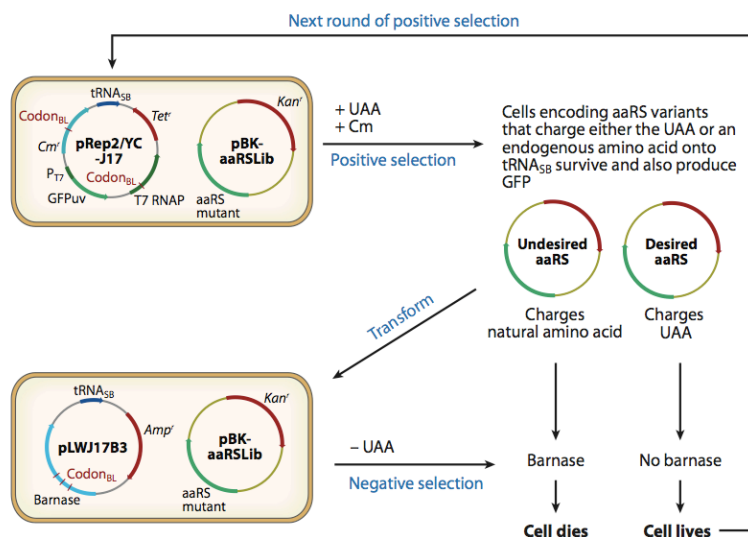


Figure 4: Schéma de sélection en deux étapes pour l'évolution dirigée d'un couple orthogonal aaRS/tRNA pour l'incorporation d'acides aminés non naturels chez *E. coli*. Abréviations : aaRS= aminoacyl-tRNA synthétase; *codon_{BL}*= codon blanc (= codon stop) ; GFP_{UV}= protéine fluorescente verte; P_{T7}= promoteur T7; T7 RNAP= T7 RNA polymérase; *tRNA_{SB}*= tRNA orthogonal suppresseur de codon stop; UAA = acide aminé non naturel. Figure tirée de la référence (23).

II.2.2. Sélection in vitro

Bien que les méthodes de sélections *in vivo* ont eu beaucoup de succès, elles sont devenues de plus en plus handicapantes face aux besoins croissants de la communauté scientifique à identifier de nouvelles molécules avec des caractéristiques hautement désirées par les industries biotechnologiques et pharmaceutiques (grandes affinités par exemple). En effet, les taux de transformation des cellules hôtes (levures ou bactéries) qui ne dépassent pas la diversité requise pour une banque de mutants, l'impossibilité de sélectionner des protéines toxiques, ni l'utilisation d'acides aminés non naturels, et les biais de la diversité parfois dus à la présence d'un mutant favorisé par son taux de croissance plus élevé par rapport aux autres clones

mais sans avantage sélectif, a poussé les scientifiques à innover et inventer des nouvelles méthodes dites *in vitro* où le passage par une phase *in vivo* est devenu limité voire complètement éliminé.

II.2.2.1. Phage display

Cette technique est essentiellement utilisée pour sélectionner des protéines dotées de propriétés de liaison spécifique à une cible choisie. Elle est basée sur une association physique entre la protéine d'intérêt et son gène via une particule virale. Par ingénierie génétique, le gène d'intérêt est fusionné à celui encodant une protéine du manteau d'un bactériophage. Généralement, ce sont des phages filamenteux de type Fd, f1 et M13 qui sont utilisés et la protéine de manteau est la pIII présente (3-5 copies/phage) à une extrémité de la particule filamenteuse (24). Le gène fusionné sera encapsulé dans les particules phagiques qui exprimeront la protéine à leur surface. Au départ d'une collection de phages, il est alors possible de réaliser des sélections basées sur une capture d'affinité (« panning ») à l'aide d'une cible immobilisée (Figure 5). Après incubation, les phages non liés sont éliminés par lavages. Ensuite, une étape d'élution permet la récupération des phages capturés exprimant les protéines d'intérêt à leur surface. Les phages sont finalement amplifiés pour pouvoir être caractérisés ou pour recommencer un nouveau cycle de sélection. Cette technique a rencontré beaucoup de succès, et est notamment largement utilisée pour la sélection d'anticorps monoclonaux (25).

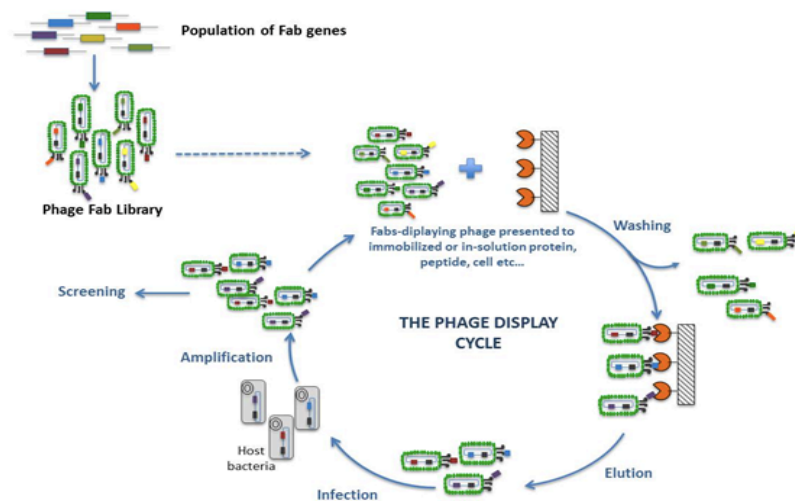


Figure 5: Cycle de sélection par phage display. L'ADN codant pour une collection de protéines (dans ce cas-ci une collection de fragments d'anticorps Fab) est cloné en fusion avec un des gènes de protéines de la capside (pIII par exemple). Les phages exprimant des protéines capables de se fixer à une cible immobilisée peuvent être sélectivement capturés. Après plusieurs lavages, les phages non désirables sont écartés et les plus intéressants sont élués puis amplifiés pour être analysés ou entrer dans un nouveau cycle. Figure prise du site : www.fjb.pt/phadis.html.

II.2.2.2. Ribosome display

Contrairement au phage display où la production des phages est réalisée *in vivo* par infection de bactéries et seule l'étape de capture est effectuée *in vitro*, la technique de ribosome display est une technique entièrement *in vitro* (Figure 6). La technologie est basée sur la liaison de la protéine à son ARN messenger dans un complexe ARNm-ARNt-protéine-ribosome formé après une réaction de traduction *in vitro*. Le complexe est maintenu stable grâce à une structure en épingle à cheveux placée en fin d'ARNm, en lieu et place du codon stop. La protéine d'intérêt est attachée au dernier ARNt utilisé lors de la traduction et via un peptide espaceur qui traverse le tunnel de sortie du ribosome et permet le repliement de la protéine. Le reste des opérations est similaire au phage display. L'étape d'amplification du matériel génétique sélectionné se fait par transcription inverse et PCR. Comme il est possible de partir directement d'une collection de fragments d'ADN générés par PCR, il

n'y a pas de limitation due à la transformation de microorganismes et la méthode permet de travailler avec des collections de très grande diversité (jusqu'à 10^{12} clones différents). D'autre part, le ribosome display permet d'échapper à certains des problèmes associés à l'expression des protéines toxiques difficilement évolvables *in vivo*.

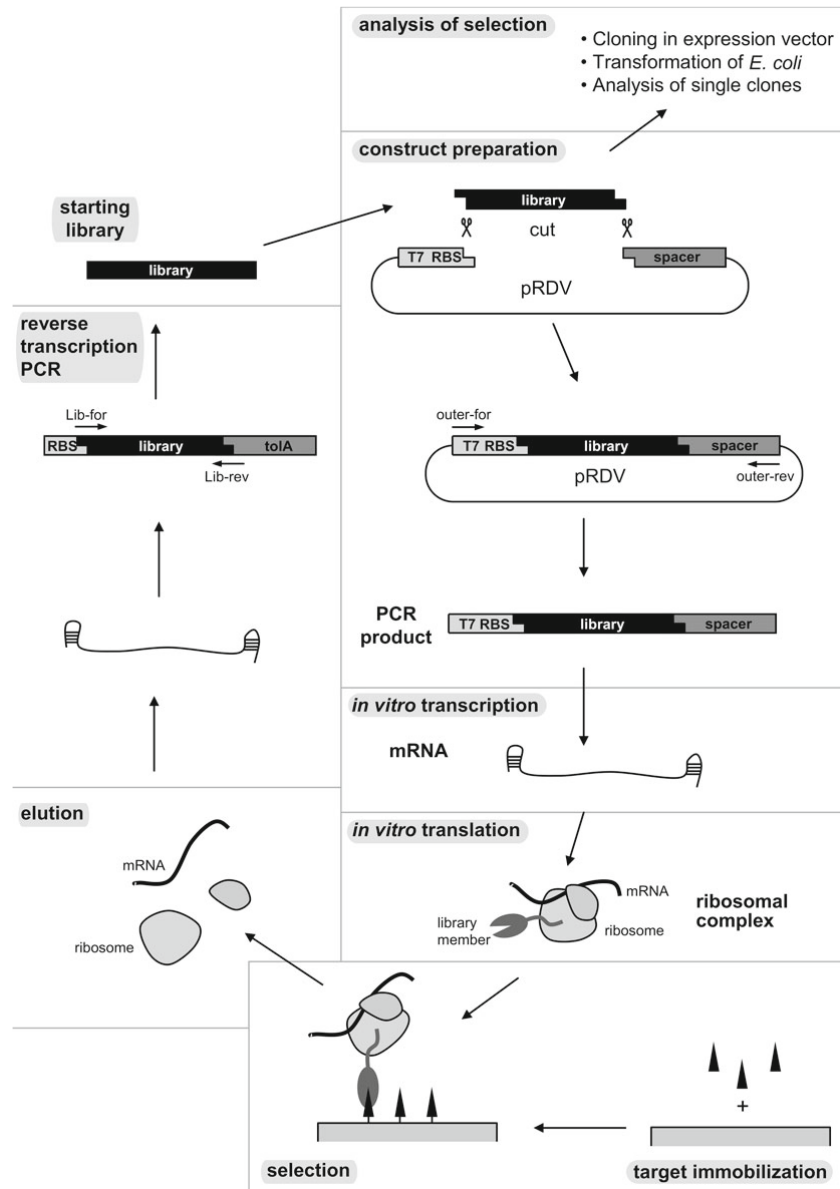


Figure 6: Cycle de sélection par ribosome display. La banque d'ADN sous forme d'un produit PCR est liguée au vecteur de ribosome display (pRDV) pour être en phase avec un espaceur (spacer) qui connectera l'extrémité C-terminale de la protéine au tRNA via le tunnel de sortie du ribosome. En aval, se trouve un promoteur T7 pour la transcription *in vitro* et la région RBS pour initier la traduction (ribosome binding site).

II.2.2.3. mRNA display

Cette stratégie est similaire au ribosome display. Une fusion covalente directe entre l'ARNm et le polypeptide qui l'encode est cependant réalisée grâce à une fonction puromycine (antibiotique analogue à un ARN de transfert chargé) située à l'extrémité 3' de l'ARNm. La traduction se fait *in vitro* et, en bloquant le ribosome à l'aide d'une structure en épingle à cheveux, le groupement puromycine peut alors réagir de manière « intramoléculaire » (intra-complexe) sur l'extrémité C-terminale de la protéine. Le ribosome peut alors se dissocier et l'ARNm se retrouve lié d'une manière covalente à la protéine qu'il encode (26). Cette technique a l'avantage de s'affranchir du ribosome et de l'ARNt mais nécessite une préparation plus complexe de l'ARN messager.

II.3. Criblage

Le criblage est l'analyse systématique de chacun des mutants d'une banque pour identifier et récupérer ceux qui ont les caractéristiques souhaitées. Les tests de criblage sont généralement basés sur une détection par spectroscopie de fluorescence ou d'absorbance et plus rarement à l'aide de tests chromatographiques ou biologiques (effet sur la croissance d'un organisme). Plus le test de criblage sera rapide, plus grandes seront les collections qui pourront être criblées. Ainsi par exemple, l'utilisation d'un substrat fluorogénique ou chromogénique adapté à une activité enzymatique recherchée permet d'analyser un très grand nombre de clones dans un temps très court. Cependant, les criblages classiques ne permettent d'analyser que quelques centaines à quelques dizaines de milliers de clones par jour et n'ont donc pas la puissance des méthodes de sélection. Les progrès technologiques en automatisation, robotisation et miniaturisation ont néanmoins mené au développement des méthodes dites à haut débit (High Throughput Screening ou HTS) permettant de cribler entre un million et un milliard de clones par jour (27–29). En particulier, les trieurs activés par fluorescence (FACS pour

« fluorescence activated cell sorter ») peuvent analyser plusieurs millions de cellules par minute tout en récupérant les clones d'intérêt. Dans certains cas, des stratégies spécifiques, assurant le maintien du couplage entre la protéine d'intérêt et son gène, doivent cependant être envisagées pour être compatibles avec le criblage par trieur. Nous allons en détailler deux ci-après.

II.3.1. Expression en surface de cellule

Il est possible de présenter une banque de protéines ou de peptides à la surface externe d'une cellule telle qu'*E. coli* ou *S. cerevisiae*. Ceci est réalisable en fusionnant le gène d'intérêt au gène d'une protéine ou d'un domaine protéique naturellement exposé vers l'extérieur. Ainsi, Francisco et ses collaborateurs ont pu exprimer une β -lactamase à la surface d'*E. coli* et ont montré qu'une protéine périplasmique pouvait être transportée à travers la membrane externe et demeurer stablement accrochée. Ceci est rendu possible en fusionnant la séquence signal et les 9 premiers acides aminés de la lipoprotéine Lpp d'*E. coli*, responsables de l'adressage hors de la cellule, à la protéine membranaire OmpA (du 49^{ème} au 159^{ème} acides aminés) elle-même fusionnée à la séquence totale de la β -lactamase (30).

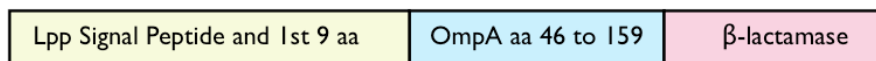


Figure 7: schéma de la fusion protéique des 3 parties Lpp-OmpA- β -lactamase. Adapté de la référence (30)

La stratégie a été appliquée par la suite avec succès à l'identification de peptides et d'anticorps d'intérêt au sein de collections (31). Par ailleurs, le « yeast display » ou la présentation à la surface de levures a eu plus de succès. La première réussite remonte à 1993 où Schreuder et ses collaborateurs ont pu présenter une α -galactosidase à la surface de *Saccharomyces cerevisiae* en fusionnant l'enzyme à la partie C-terminale d'une sous-unités de l' α -agglutinine (32). Le système a été mis au point et développé par l'équipe de Witturp. L' α -agglutinine est composée de deux

sous-unités l'Aga1p et Aga2p liées de manière covalente par des ponts disulfure. Aga2p tolère la fusion N- ou C-terminale avec une protéine d'intérêt qui se retrouve exposée à la surface de la cellule (Figure 8) (33, 34). Ce système a permis l'exposition de plusieurs protéines telles que des récepteurs de cellules T, des anticorps ou des fragment d'anticorps (35).

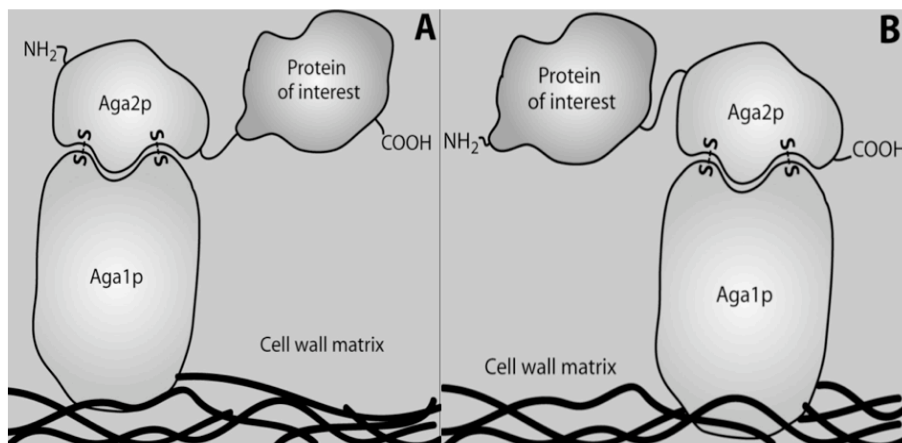


Figure 8: Le mode de fusion entre la protéine Aga2p et la protéine à exposer à la surface du la levure peut affecter son expression et sa fonction. (A) La construction qui fusionne la protéine d'intérêt par son N-terminus à la partie C-terminale d'Aga2p via un linker flexible permet une sécrétion extracellulaire guidée par le peptide signal d'Aga2p. (B) La construction qui fusionne la protéine d'intérêt par son extrémité C-terminale à la partie N-terminale d'Agap2 exige un peptide signal au niveau du N-terminus de cette construction. Figure tirée de la référence (35).

L'incubation d'une collection de levures ou bactéries présentant des protéines mutantes à leur surface avec une cible marquée par un fluorophore suivie d'un triage au FACS permet de récupérer les clones d'intérêt (liant la cible). Le débit de criblage est très élevé puisque les meilleurs FACS trient jusqu'à 100,000 cellules par seconde.

II.3.2. Micro-compartmentalisation *in vitro*

Une technologie prometteuse et ingénieuse dans le domaine de l'évolution moléculaire est la micro-compartmentalisation. De nos jours, il est possible d'imiter la compartmentalisation cellulaire en générant des microgouttelettes par émulsification d'un mélange eau-huile. Ces émulsions doivent être stabilisées à l'aide de surfactants appropriés. Elles peuvent être générées simplement à l'aide d'un homogénéiseur mais les microgouttelettes ne sont alors pas de tailles homogènes. Pour obtenir des émulsions monodisperses, il faut utiliser la technologie microfluidique plus laborieuse. Dans ces microgouttelettes, il est possible de réaliser de la traduction de protéines *in vitro*, mais aussi de la transcription d'ADN en ARN et de l'amplification d'ADN par PCR (36). Il est ainsi possible d'obtenir des micro compartiments dans lesquels des protéines sont exprimées au départ d'un ADN monoclonal et ce en l'absence de toute cellule vivante. Généralement, la protéine résultante est testée pour son activité sur un substrat fluorogénique et le criblage des bons mutants se fait à l'aide d'un trieur de gouttelettes (FADS pour « fluorescence activated droplet sorter ») ou d'un FACS classique s'il s'agit d'une double émulsion eau-huile-eau. Cette technique a été utilisée avec succès dans plusieurs études pour évoluer des enzymes telles que la beta-galactosidase ou une thiolactonase (29, 37). La compartmentalisation *in vitro* permet également d'envisager des sélections plutôt que des criblages. La propriété recherchée doit alors être couplée d'une manière ou d'une autre à la réplication préférentielle du gène correspondant. Cette approche a notamment été utilisée pour l'évolution d'une méthyle transférase (37) et d'une ADN polymérase (38). Dans le cas de la recherche d'une activité méthyle transférase, la méthylation de l'ADN encodant l'enzyme au sein du micro compartiment permet de protéger l'ADN d'une restriction ultérieure. *In fine*, l'ADN d'intérêt pourra donc être sélectivement amplifié par PCR. Pour l'ADN polymérase c'est encore plus

simple puisque la réplication compartimentée de l'ADN encodant la polymérase elle-même mène automatiquement à sa sélection.

III. Exemples de « success story » de l'évolution dirigée

III.1. Ingénierie d'une protéase stable en milieu organique

En 1993, l'équipe de Frances Arnold, grâce à l'évolution dirigée, a pu créer un variant de la Subtilisine E capable d'agir 256 fois plus efficacement que la protéine de départ en milieu contenant jusqu'à 60 % de diméthylformamide (DMF) (40). Des tours successifs de mutagenèse et de criblage ont permis de générer 4 mutants avec une activité améliorée en présence de solvant organique polaire. Le criblage était basé sur la formation de halots autour des colonies productrices et cultivées sur un solide milieu contenant de la caséine qui est un substrat insoluble de la subtilisine. Ces clones sont ensuite transférés dans un milieu qui contient le DMF et ceux dont la Subtilisine tolère la présence du solvant vont également former des halots. La largeur du halot donne une idée sur l'activité de la protéine. La combinaison dans un seul mutant (PC3) des 10 mutations identifiées lors du crible a permis d'obtenir le meilleur résultat. Toutes les mutations sont situées dans l'entourage du site actif et du site de fixation de substrat. Cette réussite a montré le potentiel de l'approche d'évolution dirigée pour changer les caractéristiques d'une protéine et la rendre active et efficace même dans un milieu non naturel. Ceci est très intéressant pour les industries en biocatalyse et en chimie fine qui utilisent des procédés où les conditions de réactions sont souvent extrêmes (pression et température élevée, solvants organiques).

III.2. Evolution dirigée d'enzymes énantio-sélectives

Un des grands challenges de la biocatalyse, est la déracémisation de composés chimiques. Bien que les enzymes soient des catalyseurs chiraux par excellence, leur capacité à ne transformer qu'un des deux énantiomères d'un mélange racémique n'est pas toujours garantie, surtout lorsque les molécules ne sont pas présentes dans la nature. Pour trouver des enzymes énantio-discriminantes, la stratégie ancienne consistait à cribler des milliers d'organismes producteurs d'enzymes, une tâche très laborieuse et dont le succès n'était pas garanti. Par évolution dirigée, Manfred Reetz et ses collaborateurs ont identifié, au laboratoire, une enzyme mutante de la lipase de *Pseudomonas aeruginosa* qui est devenue énantio-sélective (41). La création d'une banque de mutants (environ 1000 clones) via une PCR mutagénéisante a permis l'identification de 12 premiers clones énantio-sélectifs. Quatre tours de mutations et de criblage supplémentaires ont finalement mené à une lipase ayant un facteur de sélectivité 10 fois plus important que celui de la protéine de départ et menant à des excès énantiomériques (ee) proches de 90 % (Figure 9). Dans ce système le criblage a été effectué en microplaque en utilisant des esters chiraux de paranitrophenol comme substrats. L'hydrolyse enzymatique libère le paranitrophénol facilement détectable par spectrophotométrie et la comparaison des activités sur les deux substrats énantiomériques permet d'identifier les clones d'intérêt.

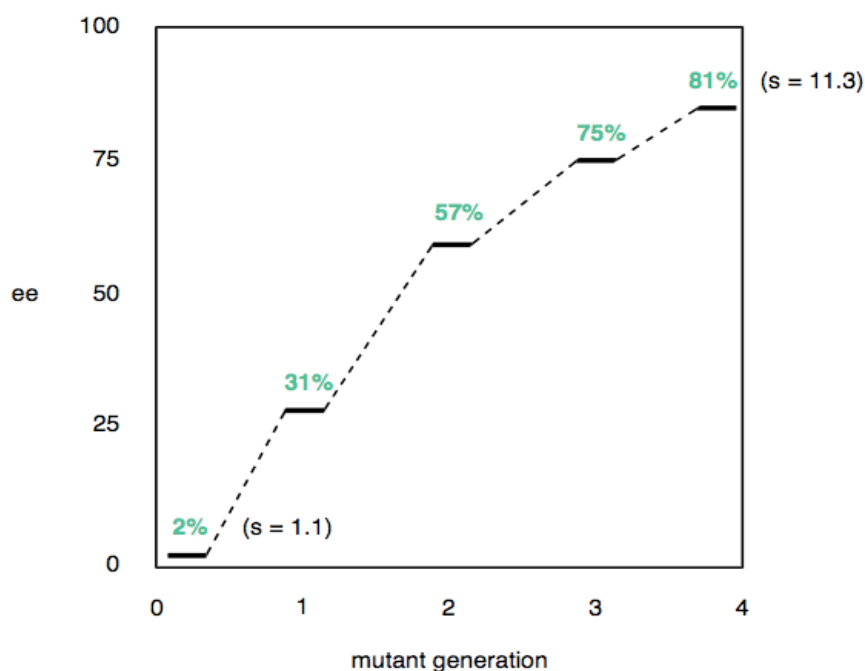


Figure 9: Amélioration de l'excès énantiomérique (ee) et de la sélectivité (s) d'une lipase au cours des tours de mutations et de criblage de petites collections de mutants.

Le succès de cette expérience a pleinement ouvert la porte à la chimie fine ou verte. D'autres projets visant à améliorer la régio- et la stéréo-sélectivité d'enzymes ont également été couronnés de succès (42, 43).

III.3. Création d'un nouveau repliement et d'une nouvelle fonction

L'ingénierie de nouvelles fonctions enzymatiques capables de catalyser des réactions chimiques est de très grand intérêt que ce soit pour les sciences fondamentales ou appliquées. Pour exploiter le potentiel de l'évolution dirigée et montrer que, sans informations mécanistiques ou même structurales, il est possible de créer une activité enzymatique à partir d'un repliement non catalytique, l'équipe de Szostak a créé une banque de mutants à partir d'un domaine stable de la protéine humaine Retinoid-X-Receptor (RXR α) et sélectionné des mutants dotés d'une activité ARN

ligase (43). Deux boucles ont été randomisées et, après 17 tours de mutations et sélections basées sur la stratégie de « mRNA display » (Figure 10), une ARN ligase est apparue.

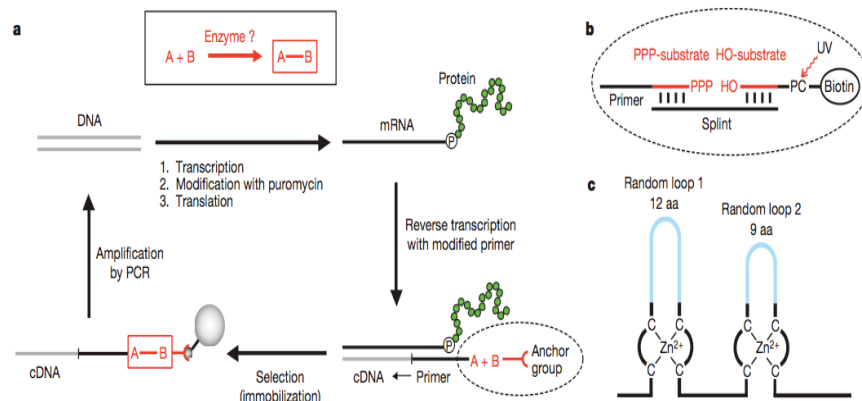


Figure 10: Schéma général de sélection des enzymes qui catalysent la ligation d'ARN : (a) Une banque d'ADN est transcrite en ARN lié à l'oligonucléotide 3'-puromycine, puis traduite in vitro. La banque d'ARNm liée à la protéine exposée subit une transcription inverse avec un primer lié au substrat A. Le substrat B qui contient un groupe d'ancrage est ensuite ajouté. L'enzyme capable de lier A et B attache le groupe d'ancrage à son ADNc, ce qui permet sa capture sélective. Les ADNc sélectionnés sont amplifiés par PCR et utilisés comme ADN de départ pour un tour suivant. (b) Sélection d'enzymes qui performent la ligation d'un ARN 5'-triphosphate activé (PPP) à un deuxième ARN avec un hydroxyle en 3' (OH) et une fonction biotine en 5' : Le substrat PPP est lié au primer et utilisé pour la transcription inverse. L'ADNc encodant la molécule catalytiquement active est immobilisé sur des billes de streptavidine via la biotine, est libéré après lavage par photoclivage du linker PC. (c) Les banques sont créées en randomisant les deux boucles représentées en bleu. Cette figure est tirée de la référence (44).

En présence de cette nouvelle protéine la ligation des deux ARNs est 2 millions de fois plus rapide qu'une réaction non catalysée. De manière fort surprenante, la protéine de départ et la protéine évoluée n'ont pas le même repliement (Figure 11) (44). Ceci montre qu'il est possible de générer de nouveaux enzymes sans être limité au même repliement. Dans ce travail, 23 substitutions et 17 délétions ont changé le repliement et la fonction de la protéine de départ. Ces résultats suggèrent qu'il y a un grand nombre de possibilités pour faire émerger une nouvelle fonction.

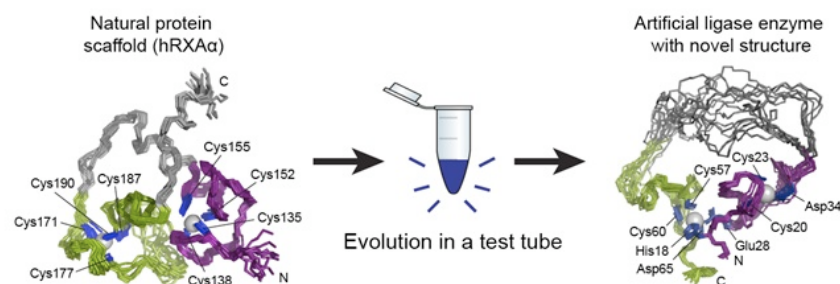


Figure 11: Changement structural après l'évolution dirigée de RXRα en ARN Ligase. Figure tirée de la référence (45).

III.4. Une nouvelle machinerie catalytique dans un ancien repliement

Une des grandes limitations de l'évolution dirigée est l'impossibilité de parcourir tout l'espace de séquences d'une protéine. Par exemple, pour une protéine de 300 acides aminés, il faut cribler plus de 10^{10} variantes pour avoir toutes les possibilités de triples mutants. D'autre part, à côté des nombreux succès dans le domaine de l'ingénierie de la spécificité (évolution dirigée d'enzymes) et de la reconnaissance moléculaire (évolution dirigée d'anticorps ou de protéines réceptrices), la capacité de prédire les mutations clés à introduire pour créer une nouvelle fonction catalytique rivalisant d'efficacité avec celles des enzymes naturelles reste un défi pour le futur (46). Dans ce domaine, un travail intéressant a été réalisé par Park et ses collaborateurs en combinant le design rationnel avec une approche évolutive. Leur stratégie a permis de créer une activité β -lactamase au départ de l'ossature metallohydrolase $\alpha\beta/\beta\alpha$ de la glyoxalase II (47). L'ingénierie de cette nouvelle activité a impliqué la délétion du domaine du site de fixation du substrat de la glyoxalase II suivie par l'introduction de boucles inspirées des structures des metallo- β -lactamases. Après des tours d'epPCR et de DNA shuffling la protéine résultante présente une activité céphalosporinase avec une efficacité catalytique (k_{cat}/K_m) de $1,8 \cdot 10^2 \text{ M}^{-1}\text{s}^{-1}$. Cette activité reste néanmoins inférieure de deux ordres de grandeur par rapport à celles des β -

lactamases naturelles mais le résultat souligne le potentiel de combinaison de l'évolution dirigée et du design rationnel.

III.5. Evolution dirigée d'une voie métabolique

Le succès de l'approche de l'évolution dirigée à générer des protéines avec des nouvelles caractéristiques a incité les chercheurs à l'extrapoler dans le domaine de l'ingénierie des voies métaboliques pour produire des composés à haute valeur ajoutée ou pour produire des souches intéressantes pour la bioremédiation. D'une manière simpliste, les mêmes stratégies appliquées à l'échelle d'une protéine sont adoptées pour toute une voie métabolique. Parmi ces tentatives, Cramer et ses collaborateurs ont réussi à faire évoluer une voie de détoxification de l'arsenate de la souche *Staphylococcus aureus* (48). L'opéron de résistance à l'arsenate contient les gènes *arsR*, *arsB* et *arsC* qui codent respectivement pour un répresseur, une pompe membranaire et une arsenate réductase. Cet opéron a été transféré dans *E. coli* et, après trois tours de DNA shuffling de ces gènes suivis de sélection pour une résistance à l'arsenate, des mutants 40 fois plus résistants ont été isolés. Cette évolution est associée à des mutations non silencieuses dans les gènes *arsB* et *arsC*.

VI. Le modèle général de l'évolution des protéines

Notre capacité à faire évoluer des protéines dans une direction souhaitée dépend également de notre compréhension du déroulement et des concepts de l'évolution naturelle car l'imitation des processus naturels en laboratoire peut s'avérer très utile. Il est maintenant établi que l'évolution et l'apparition des fonctions enzymatiques sont façonnées par la sélection naturelle. L'expansion fonctionnelle d'une superfamille d'enzymes devrait se faire d'une manière graduelle en accumulant des mutations adaptatives qui seront fixées au cours du temps. *In fine*, une superfamille d'enzymes forme un

réseau de fonctions reliées phylogénétiquement entre-elles par des ancêtres communs. Ainsi par exemple, les protéines de la superfamille de type TIM barrel ((β/α)₈) partagent un même repliement mais présentent des fonctions très différentes, allant des isomérases et kinases aux déshydrogénases et (dé)carboxylases (49, 50).

Le gain d'une nouvelle fonction pourrait être favorisé si la protéine ancestrale présente déjà la nouvelle activité mais à un niveau faible. La néofonctionalisation phénotypique pourrait alors simplement résulter de l'exacerbation de cette activité promiscuitaire secondaire. Certaines études confortent cette idée et suggèrent que les protéines ancestrales étaient des enzymes généralistes, c'est-à-dire actives sur une large gamme de substrats et/ou capables de catalyser des réactions diverses. Au cours du temps, avec les mutations et duplications de gènes, les protéines descendantes se sont progressivement spécialisées (51–53).

Ces activités secondaires sont également observées et décrites chez un nombre croissant de protéines modernes (54–56). Si une enzyme accepte des substrats alternatifs on parlera d'ambiguïté de substrat et si elle catalyse des réactions différentes on parlera de promiscuité catalytique. Ces activités secondaires constituent vraisemblablement un réservoir important d'activités de départ pour des processus évolutifs de subfonctionalisation et de néofonctionalisation.

L'évolution au départ de propriétés secondaires préexistantes n'est cependant pas le seul modèle de l'évolution moléculaire. Une propriété pourrait n'apparaître qu'après avoir libéré la protéine de la pression de sélection sur son activité originelle, suite à un changement des conditions environnementales ou suite à une duplication du gène correspondant. Cette « libération » pourrait être requise pour permettre l'apparition d'une autre fonction (57–60).

V. Contraintes de l'évolution des protéines

De nos jours, on peut représenter l'activité catalytique d'une protéine en fonction de la capacité adaptative (fitness) que cette activité confère à l'organisme. L'espace de séquences qu'une protéine peut explorer par accumulation de mutations sera associé à des variations de fitness de l'organisme. D'une manière générale le fitness est une caractéristique de l'organisme intimement liée à sa capacité adaptative, mais, de manière simplifiée, on parlera également du fitness d'une protéine comme d'une mesure de son activité spécifique *in vivo* (61). Lorsque le fitness est directement lié à l'efficacité catalytique d'une enzyme, l'accumulation de mutations menant à l'augmentation de cette activité sera bénéfique pour l'organisme. L'introduction de mutations au cours d'une trajectoire évolutive permet un voyage dans le paysage de fitness (« Fitness landscape ») et qui est un concept abstrait généralement représenté par une surface en trois dimensions où les deux dimensions horizontales représentent l'espace de séquences explorables par une protéine via des mutations et la dimension verticale indique le niveau de fitness de la protéine ou de l'hôte l'exprimant (62). Parcourir le paysage de fitness pour évoluer une protéine exige une compréhension approfondie des freins et des contraintes associés à l'évolution des protéines et plus spécifiquement des enzymes pour permettre de bien choisir et d'orienter les stratégies pour l'ingénierie de nouvelles activités tout en maximisant les chances de succès. Ces contraintes peuvent être décrites par deux caractéristiques du paysage de fitness. La première grande contrainte est liée à la forme du paysage de fitness et est généralement associée à la tolérance d'une protéine à accumuler des mutations (robustesse) ainsi qu'au nombre de trajectoires évolutives disponibles (Figure 12). Par exemple, si le paysage de fitness est lisse et présente un seul pic (modèle du Mont Fujiyama) l'évolution n'est pas freinée puisque tous les chemins empruntés à partir de chaque point de

départ vers le pic sont viables (Figure 12 A). Par contre, si le paysage de fitness est rugueux et présente plusieurs pics sub-optimaux séparés par des vallées, l'évolution adaptative et graduelle est jonchée d'obstacles (Figure 12B).

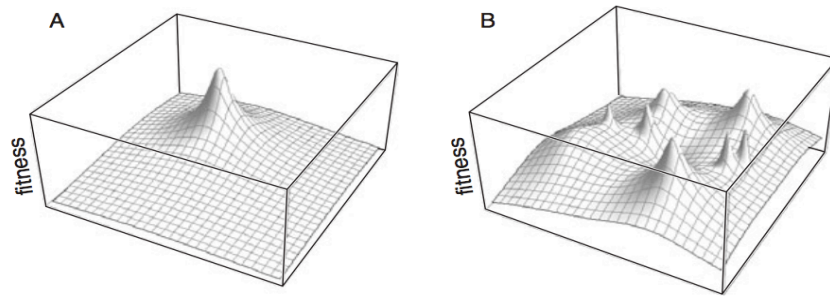


Figure 12: Paysages de fitness –Le Mont Fujiyama avec un seul pic global (A) et le paysage rugueux (B). L'axe des abscisses représente l'espace de séquences et l'axe des ordonnées représente le fitness. Figure tirée du livre « the logic of chance » de Eugene V. Koonin (63).

La deuxième grande contrainte est dépendante du chevauchement et de la connectivité entre deux paysages de fitness associés à deux activités catalytiques différentes (figure 13). Généralement, l'évolution ne peut pas établir un lien entre des activités ayant des pics de fitness non chevauchant (figure 13 A). De plus, le degré de chevauchement entre les pics façonne le type de compromis entre les deux activités échangées (figure 13 B, C et D).

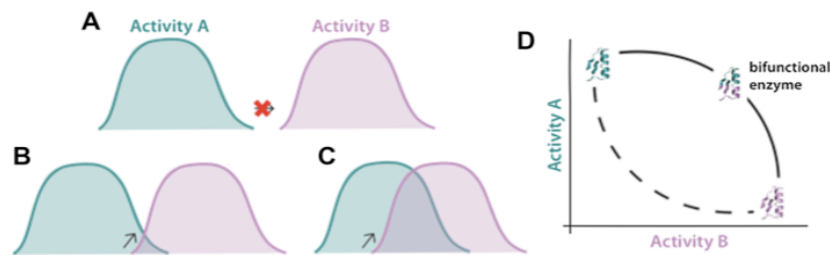


Figure 13: Connectivité entre les paysages de fitness associés à des activités différentes. Si les pics de fitness ne se chevauchent pas (un enzyme ayant l'activité A et pas l'activité B), il n'y a pas de point de départ évolutif (A). Quand les pics se chevauchent (l'enzyme ayant l'activité A et une activité B promiscuitaire), l'évolution d'un pic à un autre pourrait être accompagné d'un échange d'activités fort (B et ligne en pointillés dans D) ou d'un échange d'activités faible où un intermédiaire généraliste ayant les deux activités est viable (C et ligne continue dans D). Pour A, B et C, l'axe des abscisses respresente l'espace des séquences. Figure tirée de la référence (62).

Pour mieux comprendre ces deux grandes familles de contraintes, nous allons traiter un exemple illustrant chacune d'elles. Dans le premier exemple, le phénomène d'épistasie permettra de mettre en lumière les contraintes associées à la forme du paysage de fitness et, dans le second, nous verrons comment un exemple de connectivité entre fonctions facilite la compréhension des contraintes associées au chevauchement entre les pics de paysages de fitness.

V.1. L'effet de l'épistasie sur le paysage de fitness

Généralement on utilise le terme «épistasie» quand on étudie les effets d'interactions entre les gènes. Par définition, l'épistasie signifie que l'effet de la mutation d'un gène dépend de la présence d'une mutation dans un autre gène ou même de plusieurs. D'une manière similaire, cette définition peut être étendue au cas où les interactions entre les mutations sont observées au niveau d'un même gène (64). La figure 14 illustre la définition de l'épistasie de mutations. Supposons un gène ayant deux positions a et b qui sont les cibles de mutations A et B. Il n'y a pas d'épistasie dans le cas où $AB=A+B$, l'effet de la double mutation est égale à la somme des effets des mutations individuelles (Figure 14 A). On a une épistasie de synergie (ou épistasie positive) quand $AB>A+B$, c'est-à-dire quand l'effet de la double mutation est plus grand que la somme des effets des mutations individuelles. A l'opposé, l'épistasie est qualifiée d'antagoniste (ou épistasie négative) quand $AB<A+B$, c'est-à-dire lorsque l'effet de la double mutation est moins important que la somme des effets des mutations individuelles (Figure 14B). Dans les phénomènes d'épistasie négative, on distingue également l'épistasie de signe lorsque l'effet de la double mutation est plus faible que l'effet d'une des deux mutations individuelles (Figure 14C), et l'épistasie inverse lorsque l'effet de la double mutation est plus faible que chacun des effets des mutations individuelles (Figure 14D).

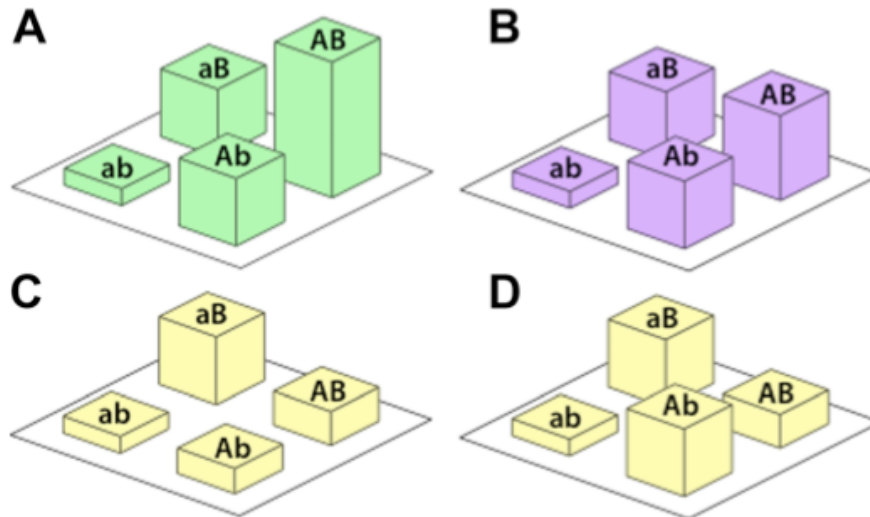


Figure 14: Epistasie. (A) Quand les mutations n'ont pas d'interactions, leur effet combiné AB est égal à la somme des effets individuels des simples mutations Ab et aB. (B) Epistasie antagoniste : l'effet dans le double mutant AB est moins important que la somme des effets des simples mutations. (C) Epistasie de signe : l'effet d'une simple mutation est plus grand que l'effet de la double mutation. (D) Epistasie inverse : chacun des effets des mutations simples est plus grand que l'effet de la double mutation. L'épistasie positive n'est pas représentée dans cette figure. Figure tirée de Kaltenbach et al., 2014 (62).

Au cours de l'évolution, si l'effet des mutations est synergique ou même légèrement antagoniste toutes les combinaisons de mutations avantageuses permettent de grimper vers un sommet de fitness sans tenir compte de leur ordre d'apparition. Par contre, l'épistasie de signe façonne un chemin évolutif accidenté dans le paysage de fitness (montées et descentes brusques). Comme certaines mutations ne sont favorables qu'en présence ou en absence d'autres mutations, le nombre de trajectoires possibles est réduit. Dans ce cas également, l'ordre et les combinaisons des mutations sont très importants (65–67).

Une étude de Weinreich et ses collaborateurs permet d'illustrer les difficultés de l'évolution des protéines causées par l'épistasie (68). Ils ont étudié l'effet de toutes les combinaisons de 5 mutations naturelles connues

pour améliorer de 5 ordres de grandeur la résistance de la TEM-1 β -lactamase aux céphalosporines de 3^{ème} génération. Contre toute attente, seules 18 trajectoires sur les 120 possibles mènent vers le pic de résistance en grimant graduellement (style Mont-Fujiyama). Les résultats indiquent donc que 102 trajectoires échappent à la sélection darwinienne et ne sont donc pas exploitables, l'épistasie apparaissant comme un frein évolutif important.

V.2. Connectivités entre les fonctions

Le second frein évolutif est lié à la connectivité entre les fonctions enzymatiques distinctes et leurs paysages de fitness respectifs. Comme vu précédemment, le chevauchement entre les pics de fitness représente une zone de promiscuité qui peut être un bon point de départ pour évoluer une fonction vers une autre (Figure 13 B-C : passage de l'activité A à l'activité B). Cependant, les mutations qui augmentent le niveau d'activité d'une nouvelle fonction pourraient affecter négativement la fonction initiale. Ceci provoque un problème évolutif si la perte de l'activité native dépasse le seuil de fitness nécessaire à la survie de l'organisme. Sans duplication de gène permettant de libérer une protéine de sa fonction originale (voir chapitre suivant), le chemin vers la nouvelle fonction ne sera donc pas tolérée et les mutations bénéfiques seront purgées.

Dans d'autres cas encore moins favorables, les paysages de fitness ne se chevauchent pas et l'activité de départ n'est pas liée à la nouvelle activité (Figure 13 A). Le passage par un intermédiaire qui peut faire le lien entre les deux activités distinctes pourrait être la solution à ce défi évolutif (Figure 15).

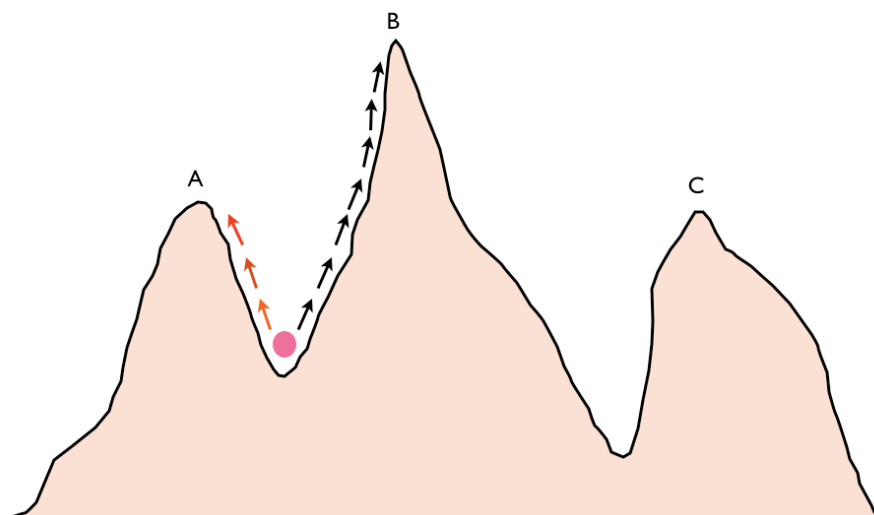


Figure 15: Schéma de 3 pics dans un paysage de fitness associés à 3 activités distinctes. Les activités A et B sont connectées et un intermédiaire (cercle rose) qui peut évoluer vers A ou B. L'évolution de cet intermédiaire vers une activité C est plus contrainte.

L'approche de la dérive neutre (Neutral drift) développée par Dan Tawfik et Frances Arnold permet également de connecter des activités qui sont a priori trop éloignées dans le paysage de fitness. Cette approche consiste à créer une diversité à partir d'un gène puis à imposer une pression de sélection permettant le maintien de l'activité initiale, ce qui permet de purger toutes les mutations délétères et non fonctionnelles (69–71). Les mutants sélectionnés forment un réseau (network) autour de la fonction initiale et sont proches ou à l'interface de l'espace de séquences d'autres fonctions (Figure 16). La population de mutants constitue dès lors un meilleur point de départ pour évoluer vers une nouvelle fonction car des nouvelles mutations générées à partir de cette population seront à même d'explorer un espace plus vaste dans le paysage de fitness.

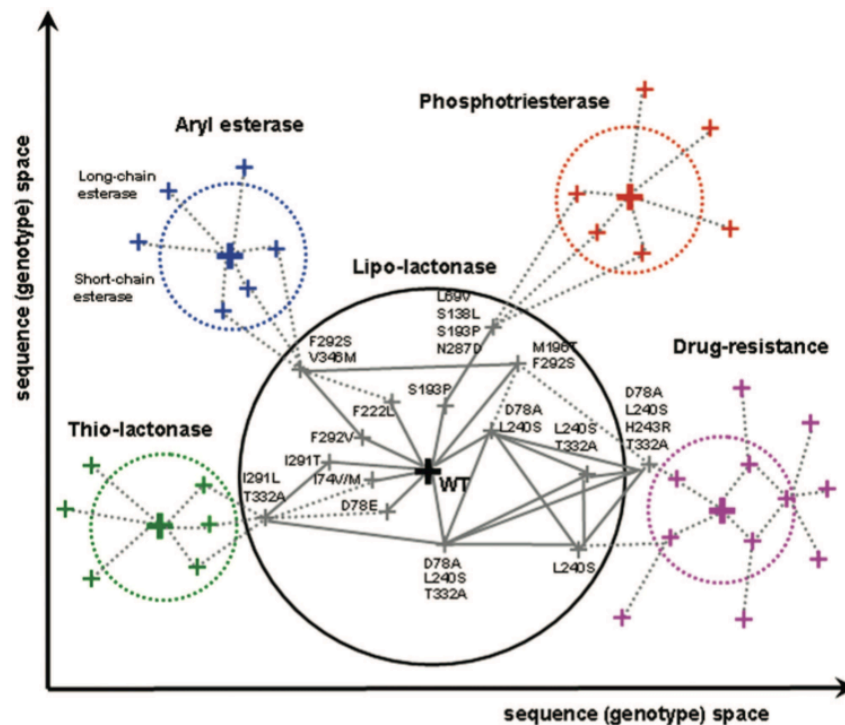


Figure 16: Représentation schématique du réseau neutre d'une lactonase (PON1). Les mutants générés par dérivation neutre forment un réseau aux alentours de PON1 wt en gardant l'activité initiale mais qui s'approchent des espaces de séquences d'autres activités. Le potentiel d'évolution de l'ensemble des variants est donc amélioré par rapport à PON1 wt. Figure adaptée de Amitai et al., 2007 (71).

VI. Evolution par duplication de gène et transfert horizontal

Le rôle de la duplication de gène comme force majeure dans l'évolution et l'innovation fonctionnelle a été le sujet de plusieurs études (72). La duplication de gène réfère à une duplication d'un segment d'ADN qui contient un ou plusieurs gènes. Le scénario le plus simpliste prédit que la redondance d'un gène permet la conservation d'une copie et la relaxation de la seconde copie de la pression de sélection associée à sa fonction. Durant

cette phase, l'accumulation de mutations va façonner le destin de ce deuxième gène et pourra mener à la naissance d'une nouvelle fonction.

Généralement, la duplication de gène peut se produire par trois mécanismes qui sont la rétroposition, la duplication du génome (ou chromosome) et les brassages inter-chromosomiques inégaux (crossing over) (73). La duplication ne se fixe pas facilement dans une population. En effet, la probabilité qu'une duplication de gène soit fixée est liée à l'effet de fitness apporté par cette duplication. Elle pourrait être délétère, neutre ou avantageuse pour un organisme. Vu le coût énergétique associé à la réplication, transcription et traduction d'un gène supplémentaire, Wagner a proposé que l'effet de la duplication est généralement délétère (74). Il est aussi possible que la duplication soit neutre si ce coût énergétique est négligeable par rapport à l'ensemble des dépenses énergétiques de l'organisme. Dans certains cas néanmoins, l'effet d'une duplication de gène peut être bénéfique. En effet, Perry et ses collaborateurs ont montré que la duplication de l'amylase salivaire est très avantageuse pour les populations qui ont un régime alimentaire riche en amidon (75). Le débat reste donc toujours ouvert pour expliquer si les événements de duplication se fixent par dérive génétique aléatoire ou par une sélection positive.

Plusieurs modèles théoriques ont décrit le processus de divergence qui suit la duplication de gène (57, 76). Le résultat le plus commun après une duplication de gène est la non-fonctionnalisation. Dans ce cas, une copie de gène accumule plusieurs mutations délétères et évolue en pseudo-gène. D'autre part, pour la duplication fonctionnelle, plusieurs possibilités sont décrites (Figure 17). Tout d'abord, la neofonctionalisation peut se produire quand une copie des deux paralogues accumule des mutations avantageuses et mène à l'émergence d'une nouvelle fonction. Le maintien de la fonctionnalité du gène avec les premières mutations ou le passage par un état non fonctionnel reste une question ouverte. D'autre part, la subfonctionalisation peut également se produire si le gène ancestral dupliqué

est multifonctionnel. Chaque copie du gène pourra alors évoluer vers une monofonctionnalité (77). Le découplage fonctionnel peut être avantageux du point de vue de la régulation car les deux fonctions ne sont pas forcément utiles dans les mêmes conditions physiologiques ou environnementales. Dans d'autres cas, le partage des fonctions sur deux copies peut se faire d'une manière neutre (dérive génétique) (78). Finalement, si le gène encode une protéine très régulée avec des fonctions vitales, la duplication peut mener à la conservation de la fonction ancestrale car la copie supplémentaire peut amener des possibilités de régulations de l'expression plus complexes. De ce fait, les mutations délétères ou celles qui engendrent la non fonctionnalité du gène sont contre-sélectionnées.

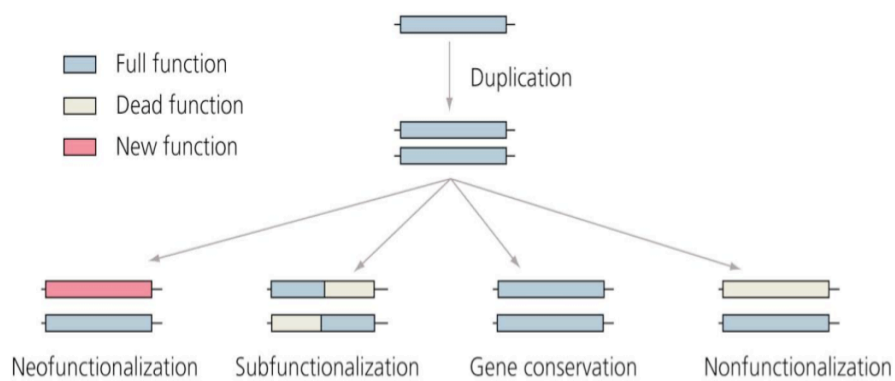


Figure 17: Destinée d'un couple de gène après duplication

Dans ce même contexte, le transfert horizontal de gènes (HGT), qui consiste à la propagation de la diversité génétique à travers les espèces, est, comme la duplication, une source d'innovation et d'évolution. Ce phénomène d'acquisition de gènes étrangers par un organisme, a été souvent décrit chez des procaryotes en réponse à des changements environnementaux (79, 80). Les exigences requises pour un transfert horizontal de gènes sont, premièrement, il faut des moyens pour que l'ADN de donateur soit livré dans la cellule réceptrice, deuxièmement, les séquences acquises doivent être bien

incorporées dans le génome de l'organisme récipient ou dans un élément réplcatif (plasmide, épisome), et troisièmement, ce nouvel ADN doit être exprimé d'une manière à être avantageuse pour l'organisme (81). L'acquisition de l'ADN étranger se fait par conjugaison ou par infection via un bactériophage (transduction) ou par transformation naturelle (certaines bactéries comme *B. subtilis* deviennent compétentes à un stade de leur cycle de vie et activent des mécanismes de capture d'ADN étranger et d'intégration dans le génome). L'introduction de l'ADN dans l'organisme récipient n'est pas une garantie d'un transfert réussi. Le maintien stable du nouvel ADN peut se faire par différents mécanismes : (i) recombinaison homologue, (ii) intégration par le biais d'intégrases de bactériophages ou de transposases des éléments mobiles (transposons), (iii) formation d'épisomes, ce qui nécessite une sélection pour éviter leur perte, (iv) intégration illégitime lors de la réparation de cassures doubles brins au niveau du génome.

Le transfert horizontal de gènes offre plusieurs avantages à l'organisme hôte. Un gène nouvellement acquis peut directement donner un avantage sélectif (résistance aux antibiotiques, nouvelles fonctions associées à l'amélioration du fitness..). D'autre part, l'ADN acquis qui, au départ, est neutre (pas d'avantage sélectif), peut subir des mutations ou des délétions ce qui pourrait générer de nouvelles fonctions. Une récente étude de Treangen et Rocha a pu illustrer cette idée en montrant que l'expansion des familles de protéines chez les procaryotes était à la base favorisée par Le transfert horizontal de gènes (82).

VII. Résistance aux antibiotiques et coût de fitness

Le fitness varie selon les conditions environnementales. Pour cette raison, certaines protéines peuvent présenter un avantage dans certaines conditions mais peuvent être un fardeau pour l'organisme dans d'autres. Ce comportement a été souvent détecté en étudiant les facteurs impliqués dans

la résistance aux antibiotiques. Comme le sujet de cette thèse concerne l'évolution d'une beta-lactamase, cette section de l'introduction présente les travaux qui ont mis en évidence un cout de fitness associé à l'expression de facteurs de résistance.

En 1988, Richard Lenski et ses collègues montrent qu'un plasmide exprimant des gènes responsables de résistance aux antibiotiques provoque une diminution du taux de croissance et que la souche hôte compense rapidement (quelques centaines de générations) ce coût de fitness par l'accumulation de mutations chromosomiques qui restaure le niveau de fitness initial (83, 84). Bien entendu, les gènes de résistance confèrent un avantage de croissance et parfois de survie pour les organismes qui les expriment en présence des antibiotiques correspondants. Par contre, en absence d'antibiotiques, la production des protéines de résistance devient une charge et va induire un coût de fitness (85, 86). A première vue, éviter l'utilisation des antibiotiques pourrait donc donner un avantage sélectif aux organismes non résistants. Mais le problème qui se pose est que les souches résistantes ont tendance à réduire les coûts de fitness associés par des mutations compensatoires. Ceci a également été illustré dans le travail de Marciano et ses collaborateurs (87). En effet, la transformation d'*E. coli* par un plasmide contenant une β -lactamase (SME-1) généralement chromosomique a engendré une chute de fitness. Des cultures successives de ces bactéries en absence d'ampicilline ont permis l'enrichissement du milieu par des clones ayant gardé la résistance mais avec un fitness comparable à la souche non transformée. L'analyse de ses clones a montré la présence d'insertions et de mutations dans le gène de la beta-lactamase. Dans la même optique, Björkman et ses collaborateurs ont eu la même constatation quand ils ont essayé de comprendre l'effet du fitness sur la virulence d'une espèce de *Salmonelle* (88). La mutation du gène *rpsL* mène à la résistance contre la streptomycine mais génère un cout de fitness en absence d'antibiotique qui peut être réduit par une simple réversion de la mutation.

Une récente étude de Fernandez et ses collaborateurs a également montré l'effet de l'expression de β -lactamases sur le fitness d'*E. coli* (89): l'expression, en absence d'antibiotiques, des β -lactamases OXA-10-like, OXA-24 et SFO-1 provoque un allongement de la cellule causé par des interférences avec la synthèse du peptidoglycane et donc une chute de fitness.

VIII. DD-peptidases et β -lactamases

VIII.1. Le peptidoglycane et les DD-peptidases

Le peptidoglycane forme un réseau tridimensionnel qui entoure la bactérie et la protège contre la pression osmotique interne de la cellule. C'est donc un élément essentiel à l'intégrité de la cellule qui de plus lui confère sa forme et sa rigidité. Comme la bactérie doit pouvoir croître et se diviser, le peptidoglycane est une structure dynamique en remodelage constant et les complexes enzymatiques impliqués fonctionnent de manière concertée avec les machineries intracellulaires impliquées dans la croissance et la division cellulaire (90).

Le peptidoglycane, schématisé à la figure 18, est constitué de chaînes linéaires de poly-diglycanes pontées par des chaînes peptidiques, ce qui forme un maillage sur une ou plusieurs épaisseurs. Les chaînes de glycanes, successions d'acides N-acétyl-muramique (NAM ou MurNAc) et N-acétyl-glucosamine (NAG ou GlcNAc) joints en β -1,4, sont modifiées par un pentapeptide sur le carbone 4 de l'acide N-acétyl-muramique, au moyen d'un lien amide. Ce pentapeptide se termine par deux résidus de D-alanine.

La production des unités di-glycanes fait intervenir une cascade de réactions cytoplasmiques qui aboutit à la formation du lipide I puis du lipide II. Ce dernier intermédiaire, constitué du monomère disaccharide-pentapeptide lié à un groupement undéca-prényle servant d'ancrage dans la membrane, est transféré grâce à la protéine flippase MurJ (91) vers la face externe de la

membrane, où il est incorporé dans le peptidoglycane via la formation d'un lien glycosidique catalysée par une transglycosylase. Les pentapeptides portés par des chaînes de glycanes voisines sont par la suite pontés entre-eux par les DD-transpeptidases qui assurent le maillage du polymère.

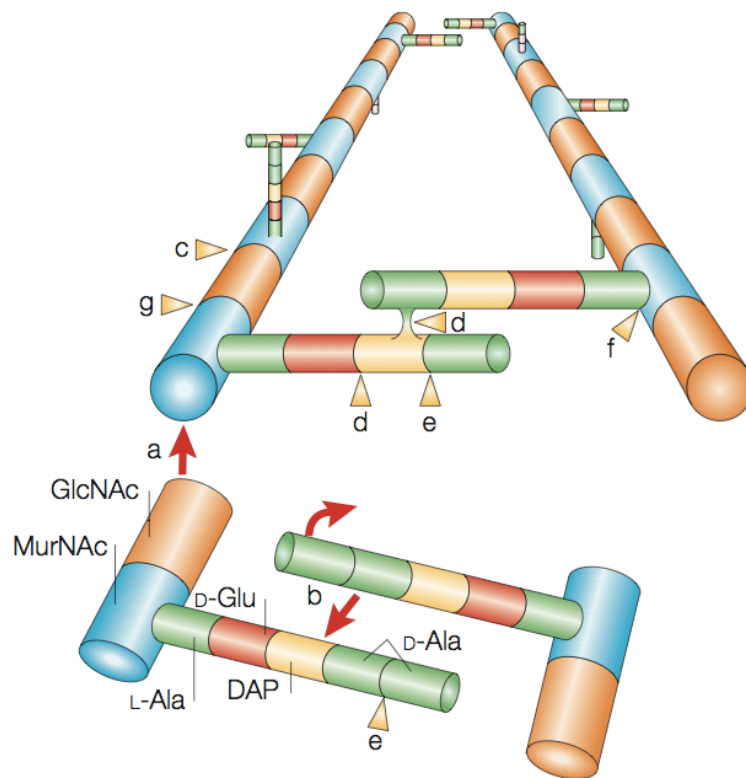


Figure 18: Synthèse du peptidoglycane. Les brins du glycanes sont formés par des répétitions de sous unités composées de l'acide N-acetylmuramique (MurNAc en Blue) et l'acide N-acétyl- glucosamine (GlcNAc en orange). Les pentapeptides sont attachés aux MurNAc. Les brins du glycanes adjacents sont liés par des pontages internes. La génération de ces pontages est dépendante d'un acide amine dibasique comme l'acide diaminopimelique (DAP). Les flèches rouges indiquent les réactions de synthèse ; les têtes de flèche jaunes indiquent les sites des activités de clivage. a : activité transglycosylase, b : activité transpeptidase, c : activité muramidase et transglycosylase lytique, d : activité endopeptidase, e : activité carboxypeptidase, f : activité amidase, g : activité N-acetylglucosaminidase. Figure adaptée de la référence (92).

La polymérisation et la rigidité du peptidoglycane sont donc tributaires de l'activité de transglycosylases pour l'élongation des chaînes de glycanes et

de transpeptidases pour la réticulation de ces chaînes. Par contre, le contrôle de la réticulation, la flexibilité et la dégradation du peptidoglycane résultent de l'action d'enzymes hydrolytiques (endopeptidases, carboxypeptidases et glycosidases). Les activités opposées des enzymes de synthèse et d'hydrolyse font du peptidoglycane une structure dynamique en continuel remodelage. Ainsi, à chaque génération d'*E. coli* la moitié du peptidoglycane est dégradée et recyclée dans sa quasi-totalité.

Les D-alanyl-D-alanine trans- et carboxy-peptidases ou DD-peptidases catalysent des réactions en deux étapes au départ des pentapeptides branchés sur les chaînes de glycanes. Premièrement, une acylation réalisée par une attaque nucléophile menée par l'hydroxyle de la sérine catalytique sur le carbonyle du pénultième D-Ala génère un intermédiaire covalent acyle-enzyme lié à la sérine par une liaison ester. Vu que la liaison scissile est acyclique, la réaction libère une D-alanine qui ne fait pas partie de l'acyle-enzyme et quitte rapidement le site actif. Puis une étape de déacylation par hydrolyse (DD-carboxypeptidase) ou aminolyse (DD-transpeptidase) fait intervenir une molécule d'eau ou la fonction amine d'un peptide accepteur qui, par un second déplacement nucléophile, libérera l'enzyme et le produit de la réaction (Figure 19). La réaction d'aminolyse permet la réticulation entre les chaînes de glycanes alors que la réaction d'hydrolyse empêche cette réticulation.

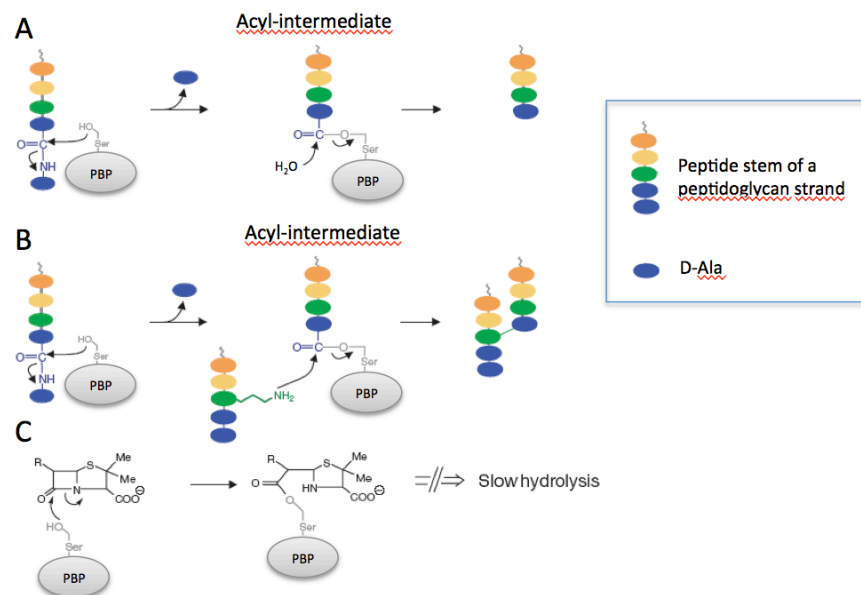
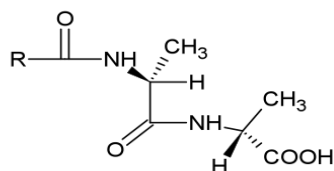
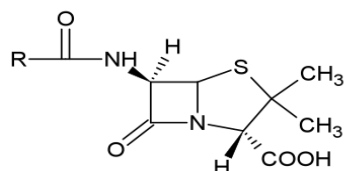


Figure 19: Réactions de carboxypeptidation (A) et de transpeptidation (B) catalysées par les DD-peptidases (aussi appelées PBPs) sur le pentapeptide du peptidoglycane. C. Réaction d'inhibition covalente des DD-peptidases par les pénicillines.

Ces enzymes font aussi partie de la famille des protéines à sérines reconnaissant les pénicillines ou PRE (Penicillin Recognizing Enzymes) (93). En effet, elles sont la cible des antibiotiques à noyau β -lactame comme les pénicillines (Figure 19). La ressemblance structurale entre ces antibiotiques et le substrat naturel des DD-peptidases (Figure 20) permet de comprendre pourquoi ces enzymes sont capables de reconnaître ces molécules. Comme avec leur substrat naturel, elles forment un acyle-enzyme, aussi appelé penicilloyle-enzyme, mais le site actif reste fort encombré car, à cause de la nature cyclique de la molécule, le groupe partant ne quitte pas le site actif. Ceci empêche donc une désacylation efficace. L'acyle-enzyme est stable, ce qui inactive l'enzyme et explique l'effet antibiotique des pénicillines. C'est aussi la raison pour laquelle ces DD-peptidases sont appelées PBPs (Penicillin Binding Proteins).



N-acyl-D-alanyl-D-alanine



Pénicilline (péname)

Figure 20: Similitude entre les peptides N-acyl-D-alanyl-D-alanine et les pénicillines (pénames). Graphe adapté de la référence (94).

En fonction de leurs tailles et de leurs domaines catalytiques, elles sont classées en PBPs de haut poids moléculaire (HMW pour High Molecular Weight) et PBPs de bas poids moléculaire (LMW pour Low Molecular Weight). Cette classification reflète également une classification fonctionnelle qui est reprise au tableau 2.

Tableau 2: Les PBPs d'*E. coli* : classification et fonctions.

Classe de PBPs		Fonction moléculaire	Fonction Physiologique
HMW class A	1a	Transglycosylase/transpeptidase	Synthèse du peptidoglycane
	1b	Transglycosylase/transpeptidase	Synthèse du peptidoglycane
	1c	Transglycosylase	inconnue
HMW class B	2	Transpeptidase	Elongation cellulaire
	3	Transpeptidase	Division cellulaire
LMW	4	Endopeptidase/carboxypeptidase	inconnue
	5	Carboxypeptidase	Forme de la cellule
	6	Carboxypeptidase	inconnue
	6b	Carboxypeptidase	inconnue
	7/8	Endopeptidase	inconnue

Tableau inspiré du la référence (92)

Les HMW-PBPs sont divisées en 3 classes A, B et C. Les classes A et B ont deux domaines fusionnés. Un domaine N-terminal de plusieurs centaines d'acides aminés est lié à un domaine transpeptidase (Penicillin Binding) C-terminal. Une région hydrophobe permet l'ancrage à la membrane et précède les deux domaines du côté N-terminal. Le domaine N-terminal des HMW-PBPs de classe A porte l'activité transglycosylase. Pour les enzymes de la classe B, le rôle du domaine N-terminal reste non élucidé.

La classe C représente les PBPs récepteurs des β -lactames comme BlaR1 de *Staphylococcus aureus*. Ces protéines ne sont pas directement impliquées dans le métabolisme du peptidoglycane. Elles n'ont pas d'activité DD-peptidase et sont impliquées dans l'induction de la synthèse des β -lactamases (95).

Les LMW-PBPs sont aussi classées en 3 sous classes A, B et C, sur base de comparaisons de séquences. Ces PBPs sont monofonctionnels endopeptidase ou carboxypeptidase. Ces enzymes sont exprimées sous forme de précurseurs fusionnés avec un peptide signal clivable qui permet leur exportation. Les LMW-PBPs sont exprimées sous forme de protéines solubles (PBP7 d'*E. coli*) ou liées à la membrane par une extension C-terminale qui permet soit l'ancrage soit l'interaction avec la surface membranaire (PBP4 et PBP5 d'*E. coli* par exemple).

VIII.2. Les β -lactamases à sérine

Les bactéries ont développé plusieurs stratégies pour résister aux antibiotiques: la réduction de l'affinité des protéines cibles, la diminution de la perméabilité cellulaire, l'efflux ou l'inactivation par modification chimique. Utilisant cette dernière stratégie, les β -lactamases (EC 3.5.2.6) sont capables d'hydrolyser les antibiotiques à noyau β -lactames telles que les pénicillines et céphalosporines. Les bactéries à Gram⁻ produisent ces enzymes dans leur périplasme, ce qui permet la dégradation de l'antibiotique

après avoir diffusé au travers des porines de la membrane externe. Par contre, les bactéries à Gram⁺ secrètent les β -lactamases dans le milieu extérieur. Les β -lactamases sont classées en deux groupes. Les β -lactamases à sérine et les β -lactamases à zinc (96). Comme leur nom l'indique, ces dernières nécessitent la présence d'ions Zn^{2+} au niveau du site catalytique pour activer la molécule d'eau hydrolytique qui attaquera le cycle beta-lactame. On les qualifie de métallo- β -lactamases ou de β -lactamases de classe B. Les β -lactamases à sérine procèdent selon un mécanisme ping pong en deux étapes. La première est l'attaque nucléophile du β -lactame par la sérine pour former un intermédiaire covalent de type acyle-enzyme (ou pénicilloyle-enzyme). Cette réaction est la même que celle qui mène à l'inactivation des PBPs (voir ci-avant). Les β -lactamases à sérine sont néanmoins capables de catalyser une deuxième étape qui consiste à hydrolyser l'acyle-enzyme de manière à compléter le cycle catalytique. Par rapport aux PBPs, les β -lactamases ont donc acquis la capacité à activer la molécule d'eau hydrolytique dans le site actif très encombré du pénicilloyle-enzyme. Une représentation plus détaillée des étapes de la réaction est donnée dans la Figure 21.

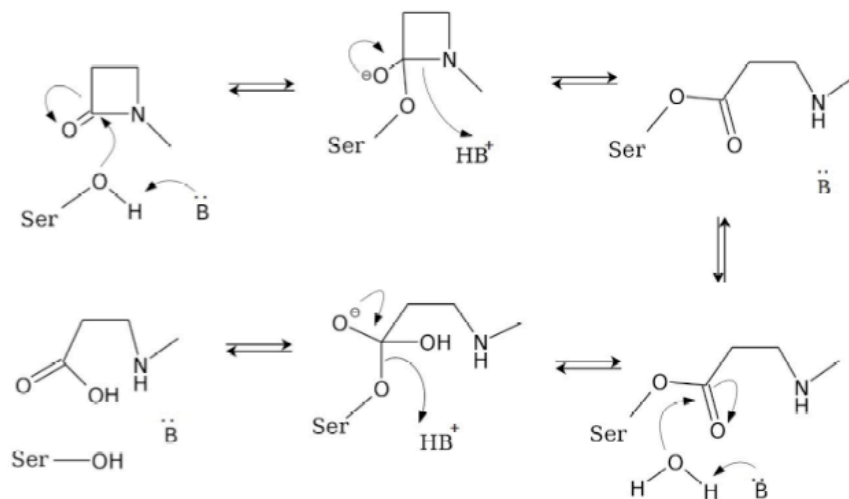


Figure 21: Mécanisme d'hydrolyse du noyau β -lactame par les β -lactamases à sérine.

Les β -lactamases à sérine sont subdivisées en 3 classes A, C et D sur base de leurs séquences en acides aminés. Les β -lactamases de classe A sont les plus répandues et les mieux étudiées. Ces enzymes sont aussi connues sous le nom de pénicillinases car elles présentent une bien meilleure capacité à hydrolyser les pénicillines que les céphalosporines. Les structures résolues des protéines de cette classe ont montré un repliement tridimensionnel similaire comprenant un domaine tout α et un domaine α/β , et la même organisation du site actif localisé à l'interface entre ces deux domaines. La TEM1- β -lactamase d'*Escherichia coli* est l'enzyme la plus étudiée de cette classe d'enzymes. Sa structure est illustrée à la Figure 22.

L'analyse détaillée des β -lactamases de classe A montre la conservation d'éléments du site actif: selon la numérotation de Ambler, on trouve le motif S₇₀XXK contenant la sérine catalytique au cœur du site actif, le motif S₁₃₀XN au niveau de ce que l'on peut appeler le « plafond » du site actif, le motif K₂₃₄TG situé au niveau du « plancher » et un glutamate à la position 166 porté par une boucle interfaciale appelée boucle Ω et flanquant un côté du site (Figure 22). Ces éléments forment, avec des molécules d'eau, un réseau dense de liaisons hydrogène qui permet les multiples réactions de protonation/déprotonation ayant lieu lors de la catalyse (97).

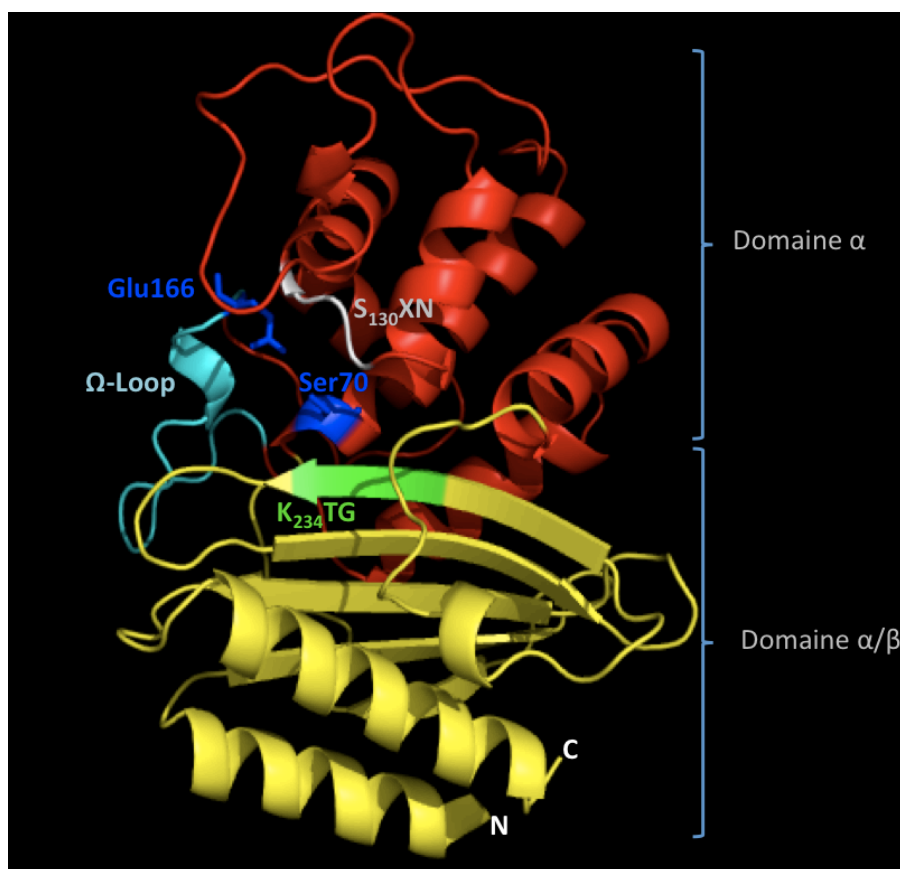


Figure 22: Structure tridimensionnelle de la TEM-1- β -lactamase (PDB code 1BTL). Les résidus du site actif : Ser70 et Glu166 sont présentés en bleu, S₁₃₀XN en gris et K₂₃₄TG en vert. Le domaine α est indiqué en rouge, le domaine α/β est en jaune et l'oméga loop est en cyan.

Le mécanisme des β -lactamases de classe A a été bien étudié. La sérine 70 est le résidu nucléophile principal qui sera acylé et désacylé lors de la catalyse mais d'autres résidus sont impliqués dans son activation. La substitution du résidu Glu166 par d'autres acides aminés a montré qu'il s'agit d'un résidu indispensable pour la désacylation mais qu'il est également important pour l'étape d'acylation (98, 99). Ce résidu forme un pont hydrogène avec une molécule d'eau du site actif elle-même en pont hydrogène avec l'hydroxyle de la sérine catalytique (Ser70), ce qui permet un relais de proton pour l'activation de la sérine (catalyse basique). La lysine

73 est également proposée comme résidu basique important pour l'activation de la sérine et il est possible que les deux résidus (K73 et E166) coopèrent lors de l'acylation. La nucléophilicité de la Ser₇₀ est donc exacerbée ce qui permet l'attaque de l'oxygène sur le carbonyle du β -lactame sur la face α . Un réseau de liaisons hydrogène impliquant les groupements amine des résidus Lys₇₃ et Lys₂₃₄ et l'hydroxyle de la Ser₁₃₀ permet le transfert d'un proton à l'azote du noyau β -lactame et la formation du pénicilloyle-enzyme. Pour l'étape de désacylation, le résidu Glu166 joue le rôle de base catalytique pour activer la molécule d'eau hydrolytique qui attaquera le carbonyle de la fonction ester du pénicilloyle-enzyme également par la face α . Le motif S₁₃₀DN est impliqué aussi dans un réseau de liaisons hydrogène pour stabiliser le substrat. Il semble avoir un rôle dans le maintien de la géométrie du site actif. Il est aussi avéré que le motif K₂₃₄TG est également impliqué dans la liaison au substrat en stabilisant le groupe carboxylique de l'intermédiaire tétraédrique. Le résidu Gly₂₃₆ hautement conservé est indispensable pour maintenir la géométrie du site actif.

Les β -lactamases de classe C sont aussi appelées céphalosporinases pour leur aptitude à mieux hydrolyser les céphalosporines. Ces enzymes partagent le même mécanisme pour l'hydrolyse du noyau β -lactame avec quelques différences au niveau de la désacylation. En effet, la molécule d'eau attaque l'acyle-enzyme par l'autre face (face β) de la fonction ester. Par ailleurs, cette molécule d'eau est activée différemment par d'autres résidus. La Tyrosine₁₅₀ était considérée comme responsable de cette activation mais des études ont montré que ce résidu reste neutre au pH où la réaction a lieu (100) et même sa substitution en phénylalanine affecte peu la vitesse de désacylation (101). Il a été donc proposé que, après l'ouverture du noyau β -lactame, l'azote du noyau dihydrothiazine du groupement acyle soit bien positionné dans l'acyle-enzyme pour l'activation d'une molécule d'eau (101,

102). Cette classe d'enzymes a un repliement similaire aux β -lactamases de classe A et les positions des motifs du site actif sont bien conservées mis-à-part la boucle omega qui est absente (103).

Les β -lactamases de classe D sont aussi appelées oxacillinasés car elles hydrolysent préférentiellement les oxacillines. Le mécanisme d'action de cette classe est également un peu différent. En effet, l'addition de dioxyde de carbone (CO_2) à l'amine de la chaîne latérale du résidu Lysine du motif SxxK permet la formation d'un carbamate en contact par liaison hydrogène avec la sérine catalytique, ce qui permet l'activation de cette dernière pour l'acylation (104). Ce carbamate est également impliqué dans l'activation de la molécule d'eau hydrolytique lors de la désacylation (105). Le repliement tridimensionnel des enzymes de cette classe est aussi similaire à celui des β -lactamases de classe A (106).

Les structures chimiques de substrats de β -lactamases sont présentées dans la figure suivante (Figure 23).

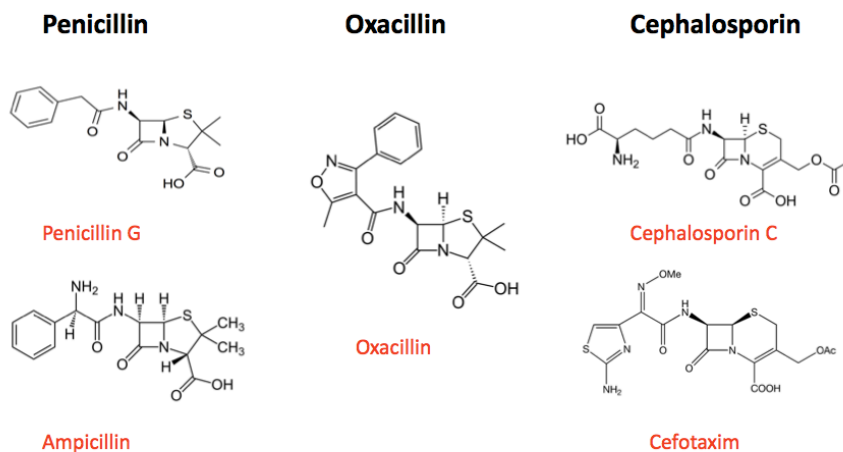
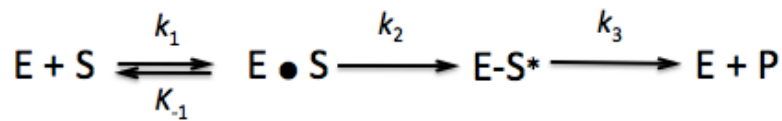


Figure 23: Structures chimiques de substrats des β -lactamases

VIII.3. Equations cinétiques de l'interaction entre les PBPs/ β -lactamases avec les β -lactames

Le mécanisme catalytique des PBPs et des β -lactamases peut être schématisé comme suit :



ou k_2 et k_3 sont les constantes de premier ordre d'acylation et de désacylation respectivement et k_1 et k_{-1} sont les constantes de vitesse d'association et de dissociation du complexe enzyme-substrat non covalent. Les paramètres cinétiques de Michaelis-Menten sont les suivants :

- le k_{cat} ou constante catalytique est, dans ce modèle, fonction de toutes les vitesses de premier ordre.

$$k_{cat} = \frac{k_2 \cdot k_3}{k_2 + k_3}$$

- le K_m ou constante de Michaelis qui équivaut à la concentration en substrat pour laquelle la vitesse égale à $V_{max}/2$

$$K_m = \frac{k_3}{k_2 + k_3} K_s$$

avec $K_s = \frac{[E] \cdot [S]}{[E \cdot S]} = \frac{k_{-1}}{k_1}$, soit la constante d'équilibre de dissociation du

complexe de Michaelis.

Dans le cas où $k_2 \gg k_3$ (ce qui est le cas des PBPs), $k_{cat} = k_3$; $K_m = K_s \cdot (k_3/k_2)$ et donc k_{cat}/K_m sera égal à k_2/K_s .

VIII.4. Liens évolutifs entre les DD-peptidases ou PBPs et les β -lactamases

Au cours de l'évolution naturelle, il semble logique que les DD-peptidases soient apparues avant les antibiotiques à noyau β -lactame qui ciblent ces enzymes, et que ces antibiotiques soient apparus avant les β -lactamases dont elles sont les substrats. Comme les DD-peptidases et les β -lactamases à sérine partagent un même repliement et des motifs conservés au niveau du site actif, il est donc généralement accepté que ces dernières aient évolué à partir des DD-peptidases sensibles aux pénicillines (107). L'analyse par alignements de séquences ne permet cependant pas facilement d'établir un lien phylogénétique direct, ou de tracer une trajectoire évolutive vu la diversité des séquences et surtout la faible similarité entre les familles. Certaines études d'alignements, comme celle menée par Kirby sur certaines séquences de β -lactamases de classes A et C en utilisant celle de la DD-peptidase R61 de *Streptomyces* comme référence, suggèrent néanmoins que les β -lactamases sont phylogénétiquement reliées aux PBPs (108, 109).

D'après Massova et Mobashery (110), l'évolution des PBPs vers des β -lactamases n'est pas un processus linéaire. En effet, différentes PBPs pourraient avoir évolué séparément et parallèlement vers différentes classes de β -lactamases en parcourant des trajectoires distinctes. Deux processus ont été proposés pour décrire la voie évolutive. Certains chercheurs pensent que des PBPs auraient initialement acquis l'aptitude à catalyser la désacylation du penicilloyle-enzyme. Ensuite, l'acylation et la désacylation auraient été optimisées par des substitutions et des insertions/délétions pour atteindre une activité catalytique adéquate. Selon ce processus, la PBP originale devient bifonctionnelle (DD-peptidase/ β -lactamase), ce qu'on qualifie également d'enzyme généraliste, résultat d'un faible "trade-off" négatif entre la nouvelle fonction et la fonction originale

(figure 24). Dans les dernières étapes de ce processus, les PBPs perdent leur fonction native par des réarrangements structuraux limitant ou bloquant les interactions avec le peptidoglycane, mais tout en gardant la capacité à hydrolyser le noyau β -lactame. D'autres scientifiques supportent l'idée que les activités PBP et β -lactamase sont exclusivement mutuelles (111). La co-existence des deux activités pourrait être impossible. Pour cette raison, la perte de l'interaction avec le peptidoglycane doit accompagner ou même précéder l'acquisition d'un site catalytique capable d'hydrolyser les β -lactames (trajectoire avec un "trade-off" négatif fort dans la figure 24). Cependant, si les deux activités sont mutuellement exclusives, il n'y aura donc pas de pression selective pour maintenir les intermédiaires non ou faiblement fonctionnels, ce qui pose question d'un point de vue évolutif.

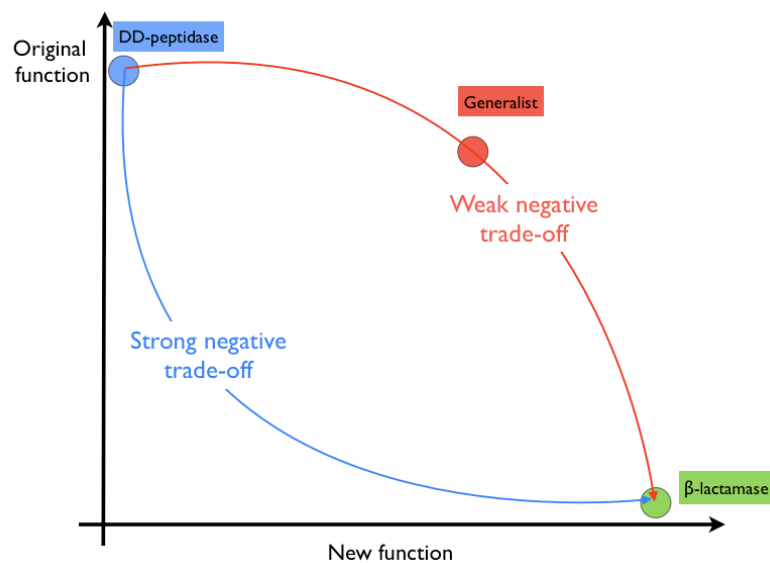


Figure 24: Schéma du "trade-off" entre deux fonctions au cours de l'évolution. Le gain-perte de la nouvelle contre l'ancienne fonction et la conversion d'une protéine spécialisée (cercle bleu) en une autre (cercle vert) pourrait suivre différentes voies. La trajectoire rouge représente un faible « trade-off » négatif : l'augmentation de la nouvelle fonction est accompagnée par une faible diminution de la fonction originale. Les intermédiaires de cette voix sont des généralistes (qui catalysent les deux réactions). La voie bleue représente une trajectoire avec un « trade-off » négatif fort: le gain d'un faible niveau de la nouvelle fonction est accompagné une grande perte de la fonction originale. Adapté de la référence (56).

IX. Tentatives de conversion d'une PBP en β -lactamase

L'évolution dirigée d'une PBP vers une β -lactamase à l'échelle du temps de laboratoire est un objectif qui semble réaliste au vu de la relation phylogénétique qui existe entre les deux superfamilles d'enzymes et de par la facilité à sélectionner l'activité β -lactamase *in vivo*. Cet objectif est intéressant à plus d'un titre : (i) il devrait permettre de comprendre les mécanismes d'évolution moléculaire permettant de faire émerger l'activité β -lactamase qui, par rapport aux DD-peptidases, nécessite une machinerie hydrolytique supplémentaire ou repositionnée ; (ii) il devrait également permettre de mieux comprendre les relations structure-fonctions des β -lactamases et les déterminants moléculaires à l'origine de l'efficacité catalytique de ces enzymes. D'un point de vue plus général, l'ingénierie génétique d'une nouvelle machinerie catalytique est également ambitieuse car la grande majorité des travaux dans le domaine concerne la modification de spécificité mais pas de la catalyse.

Au départ de cette thèse, peu de travaux visant cet objectif étaient publiés et tous rapportent des résultats modestes. Ceci reflète sans doute une difficulté intrinsèque à réaliser cette conversion enzymatique.

IX.1. Evolution de la PBP2X

En 2002, Chesnel et ses collaborateurs ont utilisé la PBP2X de *Streptococcus pneumoniae*, une des PBPs les plus étudiées, comme un point de départ pour essayer de générer une activité β -lactamase (112). Cette PBP montre peu de similarité de séquences avec les β -lactamases à sérine mais la comparaison des structures 3D montre une grande similarité de topologie et une conservation des motifs du site actif (98). Cette protéine comme la majorité des PBPs n'a pas de boucle Ω et donc pas de glutamate capable de catalyse basique. Cette position est occupée par une phenylalanine (Phe450).

L'idée de Chesnel consistait donc à muter ce résidu par plusieurs acides aminés basiques pour espérer améliorer la désacylation (Tableau 3).

.

Tableau 3: Taux de désacylation (k_3) de différents mutants en fonction du pH. Tableau tiré de la référence (112).

Proteins	k_3 ($s^{-1} \times 10^5$)				
	pH 5	pH 6	pH 7	pH 8	pH 9
S-PBP2x*	3.2 ± 0.6	1.9 ± 0.5	2.3 ± 0.6	3.3 ± 0.6	3.2 ± 0.7^a
Phe450 → Gly	3.0 ± 0.8	3.4 ± 0.6	2.6 ± 0.5^a	6.4 ± 0.9^a	6.4 ± 1.7^a
Phe450 → Ser	5.4 ± 0.9^b	2.5 ± 0.6	3.4 ± 1.0^b	6.1 ± 0.8	5.8 ± 0.7
Phe450 → His	15.3 ± 2.0	18.8 ± 4.9^b	21.1 ± 2.9	22.9 ± 2.1	14.8 ± 2.5
Phe450 → Glu	7.8 ± 0.8	7.4 ± 1.4	11.5 ± 2.2	29.0 ± 4.2	23.5 ± 3.2
Phe450 → Asp	22.8 ± 5.7	32.8 ± 9.2	89.6 ± 7.9	368.7 ± 29.1	210.2 ± 28.7

Ce travail, a permis une amélioration de la désacylation d'un facteur 110 à pH8 juste en mutant la Phe450 en aspartate. La dépendance du pH sur la constante de désacylation suggère que l'état de protonation des résidus impliqués dans la désacylation influence le déroulement de la réaction.

En 2003, Peimbert et Segovia ont rapporté un chemin évolutif qui mène à une activité β -lactamase au départ de la même transpeptidase PBP2X de *S. pneumoniae* (113). Ils ont adopté une stratégie qui marie le design rationnel et l'évolution dirigée. Pour anticiper les mutations avantageuses, les positions identifiées comme potentiellement importantes pour l'émergence de la nouvelle activité ont été rendues partiellement aléatoires (Figure 25). Les raisons pour lesquelles ces résidus ont été mutés sont énumérées dans le tableau 4.

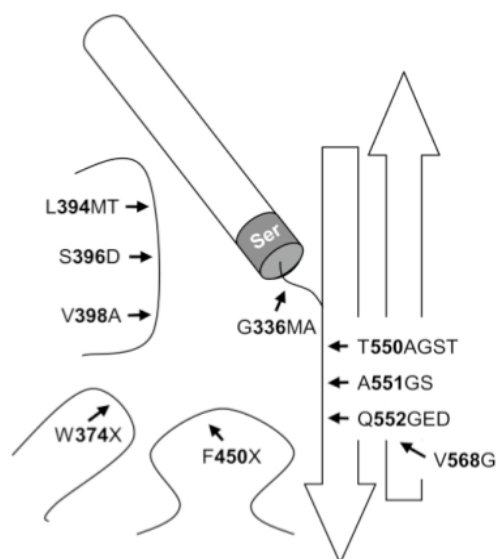


Figure 25: Diagramme des mutations ciblées représentées schématiquement autour de la sérine catalytique. Figure tirée de la référence (113).

Tableau 4 : Enumération des mutations ciblées.

Residu muté	Mutation	Homologie	Explication
Phe 450	X ^a	Glu166 de Classe A	Ces deux résidus bloquent l'accès au site actif d'une molécule d'eau indispensable pour l'hydrolyse.
Trp374	X ^a	Leu293 de Classe C	
Thr550	A/G/S/T	/	Ces positions se trouvent dans la poche de reconnaissance du substrat. Les mutations sont inspirées de la référence (114).
Ala551	G/S	/	
Gln552	G/E/D	/	
Leu394,	M/T	/	Ces résidus conservés chez les β -lactamases de classe A. Les analogues de Leu394 et Val398 sont généralement soit Met/Thr ou Ala respectivement.
Ser396	D	Asp131 de Classe A	
Val398	A	/	
Tyr568	G	Gly216 de Classe A	
Gly336	M/A		Proche de la sérine nucléophile. Des β -lactamases mutées dans des positions analogues ont

			montré une résistance à la clavulanate.
--	--	--	---

^a : Mutagenèse à saturation.

Après avoir introduit ces substitutions, le domaine codant pour la transpeptidase a subi une PCR mutagénisante pour construire une banque combinatoire de mutants. Ensuite, la sélection à faible concentration en céfotaxime (200 ng/ml) a généré des clones qui confèrent de 4 à 10 fois plus de résistance par rapport à la souche de départ. Ces clones ont des mutations dans 6 positions (Tableau 5).

Tableau 5 : les substitutions dans les clones résistantes au céfotaxime. Ce tableau est tiré de la référence (113)

	WT	Ala312	Gly336	Trp374	Phe450	Gln452	Ala554
C library	MutA	Thr	Ala		Leu	His	Thr
	MutB	Thr	Ala	Arg	Thr		
	MutC	Thr	Ala		Leu	His	
	MutD	Thr	Ala				

Les substitutions Gly336Ala, Trp374Arg et Phe450Leu sont issues de la mutagenèse ciblée, alors que Ala312Thr, Gln452His et Ala554Thr sont issues de la PCR mutagénisante.

Les auteurs ont créé par la suite, en se basant sur ces clones, de nouvelles banques par DNA shuffling et par mutagenèse de saturation. Ce nouveau tour de randomisation a permis la sélection d'un mutant montrant une accélération du taux de déacylation par 4 ordres de grandeur sur la céfotaxime tout en gardant son activité carboxypeptidase intacte. Il présente 3 substitutions dont 2 au niveau de la boucle Ω et 1 à côté de la sérine catalytique (Figure 26). Etonnament, la position 450 structuralement alignée avec le glutamate 166 dans les β -lactamases de classe A est mutée en leucine alors que la collection était construite de manière à y introduire un résidu basique. Par ailleurs, une histidine est sélectionnée un peu plus loin dans la

boucle en position 452 et pourrait prendre le rôle de catalyseur basique lors de l'étape de désacylation.

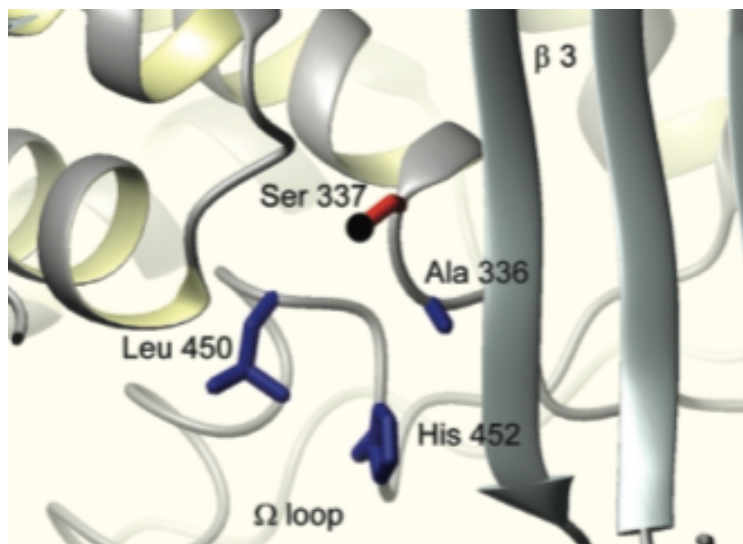


Figure 26: Modèle du site actif du meilleur mutant. Les trois mutations sont représentées en Bleu et la sérine catalytique en rouge. Figure adaptée de la référence (113).

Malgré la réussite dans le changement des propriétés de cet enzyme et l'augmentation importante de l'activité céphalosporinase, le k_{cat} est encore 5 ordres de grandeur inférieur à celui d'une beta-lactamase comme la TEM-1 et l'activité pénicillinase du mutant n'est pas améliorée. Par ailleurs, l'article n'est pas très convaincant car des expériences clefs ne sont pas présentées (construction d'un mutant catalytique, gel SDS-PAGE des protéines purifiées, recherche d'activité contaminante).

IX.2. Evolution de la DD-peptidase R61 de *Streptomyces*

Dans une étude focalisée sur l'importance de deux tryptophanes (positions 271 et 233) dans l'activité de la DD-peptidase R61 de *Streptomyces*, la mutation W233L a fortuitement généré une augmentation de la désacylation du benzylpenicilloyle-enzyme par 300 fois (115). Ce résultat inattendu a été accompagné par la perte de l'activité peptidase. Bien que le mutant soit

toujours environ 20000 fois plus lent qu'une beta-lactamase sauvage et ne confère pas de phénotype de résistance aux pénicillines, il s'agit néanmoins d'une amélioration importante d'activité résultant d'une mutation unique.

IX.3. Evolution de la PBP-A

Au sein de notre laboratoire, l'évolution d'une PBP en β -lactamase est un objectif de recherche depuis plus de 20 ans. Une première tentative visait à faire évoluer la PBP4 d'*E. coli* grâce à une stratégie de sélection *in vitro* par phage display utilisant des substrats suicides de la β -lactamase. Aucune amélioration de l'activité beta-lactamase n'a cependant été obtenue.

Une nouvelle famille de PBPs cyanobactériennes présentant la plus importante similarité de séquence avec les β -lactamases de classe A a ensuite été identifiée (109). Cette famille de protéines a la particularité de présenter une délétion de six acides aminés au niveau de la boucle Ω et dans la région qui porte le glutamate 166 chez les β -lactamases (Figure 27).

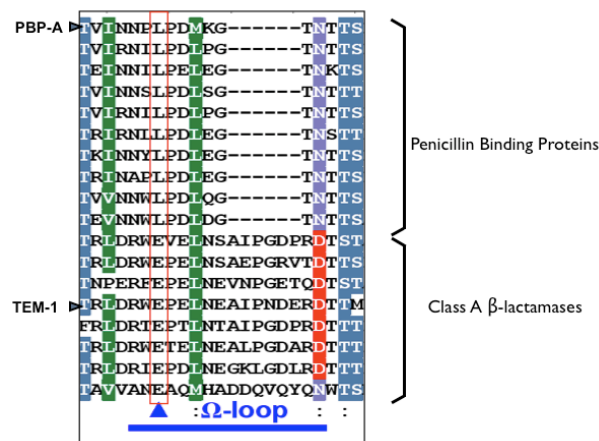


Figure 27: Alignements des boucles Ω des PBPs et des β -lactamases.

La PBP-A thermostable de *Thermosynechococcus elongatus* appartenant à cette famille a été choisie pour tenter l'évolution dirigée. Elle ne présente environ que 30 % d'identité de séquence avec les β -lactamases de classe A mais la détermination de sa structure tridimensionnelle a révélé un

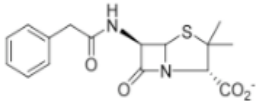
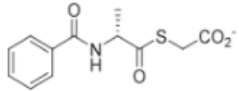
repliement quasi parfaitement superposable avec celui de la TEM-1 β -lactamase (Figure 28).



Figure 28: Superposition de la PBP-A (vert) et la TEM-1 β -lactamase (rouge). Figure adaptée de la référence (116).

L'introduction d'un glutamate à la position 158 à la place d'une leucine alignée avec le glutamate 166 des β -lactamases de classe A a permis l'augmentation de la vitesse de déacylation du penicilloyle-enzyme par un facteur 90 (Tableau 6). La structure du mutant montre que le glutamate 158 est bien positionné mais pas parfaitement superposé avec le Glu166 de la TEM-1. Par ailleurs, le mutant ne confère aucune résistance aux β -lactames à la bactérie hôte.

Tableau 6: Comparaison des paramètres cinétiques du mutant L158E et la PBP-A

	 Benzylpenicillin k_2/K_S ($M^{-1}s^{-1}$)k_3 (s^{-1})	 Phenyl-D-Alanyl-thioglycolate (D-Ala-D-Ala mimick) k_{cat}/K_M ($M^{-1}s^{-1}$)k_{cat} (s^{-1})K_M (mM)
PBP-A	1.7 10 ⁴ 1.5 10 ⁻⁴	3.5 10 ⁴ ≥ 52≥ 1.5
PBP-A L158E	1.6 10 ³ 1.4 10 ⁻²	3.9 10 ³ ≥ 5.8≥ 1.5

Suite à ce résultat encourageant, une banque a été créée par PCR mutagénisante au départ de ce mutant mais la sélection *in vivo* n'a pas permis d'isoler de clones résistants.

L'allongement de la boucle Ω a ensuite été testé car il est possible qu'une boucle plus longue soit nécessaire pour le positionnement précis de la base catalytique. Une banque a été construite par insertion d'une séquence peptidique partiellement dégénérée (Figure 29). La taille de la boucle observée chez les β -lactamases de classe A a été restituée en ajoutant six résidus fréquemment trouvés chez les β -lactamases et en forçant la présence du glutamate à la position 158. La sélection *in vivo* n'a cependant pas permis d'isoler des clones résistants aux pénicillines.

Omega Loop Grafting



Figure 29: Logo de séquence dégénérée ayant servi à la création d'une banque d'Oméga loops allongées.

A ce niveau, en profitant du succès de la cristallisation de PBP-A, une approche rationnelle a été envisagée. En effet, plusieurs mutations guidées par la structure ont été introduites mais aucun des mutants ne présentait une activité dépassant celle du premier mutant construit (PBP-A-L158E) (Tableau 7).

Tableau 7: Paramètres cinétiques des mutants construits sur base de la structure.

	$k_2/K_s (M^{-1} s^{-1})$	$k_3 (s^{-1})$	$k_{cat} (s^{-1})$
PBP-A	$1,07.10^4$	$1,55.10^{-4}$	$2,7.10^{-4}$
PBP-A-L158E	$1,56.10^3$	$1,44.10^{-2}$	$1,42.10^{-2}$
D160N	$2,9.10^4$	3.10^{-4}	$6,4.10^{-4}$
L158E-D160N	$7,8.10^3$	$4,9.10^{-4}$	$4,6.10^{-4}$
M161N	$6,1.10^3$	$5,5.10^{-4}$	$3,3.10^{-4}$
L158E-M161N	$1,7.10^3$	8.10^{-3}	$8,1.10^{-3}$
L158E-L205R	nd	nd	~ 0
L158E-V227R	nd	nd	$1.4.10^{-2}$

Le substrat utilisé est le benzylpenicilline (PenG).

Les échecs de ces tentatives multiples illustrent bien la difficulté à faire évoluer le mécanisme catalytique d'une enzyme même lorsque la structure du site actif de départ est quasi superposable à celle visée. Ils suggèrent des obstacles ou des freins inconnus qui complexifient les trajectoires évolutives permettant l'émergence de nouvelles enzymes. D'autres critères pourraient être pris en considération pour espérer un jour générer un niveau d'activité approchant celui des β -lactamases sauvages. La compréhension des freins évolutifs et les rôles physiologiques des deux types de protéines pourraient également aider à mettre en place une stratégie d'évolution.

Objectives and Strategy

I. Objectives

Enzyme engineers are creating biocatalysts endowed with new functions. Beside all the success recorded in the last decades in engineering enzymes with altered stability, specificity or regulatory properties, thanks to the impressive progress of molecular evolution tools and techniques, understanding how new catalytic machinery can emerge within an active site is still beyond the pareto front of our knowledge. This challenge push the scientists to dive into the unknown and try to find elements and clues leading us to unveil the mysteries of the evolutionary birth of enzymes.

The directed evolution between phylogenetically related enzymes that feature different catalytic machineries is the general objective of this thesis. More specifically, we will try to create a serine β -lactamase capable of hydrolyzing penicillins starting from a DD-peptidase involved in the biosynthesis and maintenance of the bacterial peptidoglycan (*107*). Because both enzyme families share a similar tridimensional fold and because the peptidoglycan is obviously a more ancient structure than penicillin, a realistic hypothesis is that serine β -lactamases evolved from some ancestral DD-peptidases. In the evolutionary birth of class A β -lactamases, the fixation of a glutamate in the active site could have been the main step undergone by the putative ancestral DD-peptidase. This additional catalytic element allows the activation of a water molecule in the crowded active site of the penicilloyl-enzyme, initiating the deacylation-limiting step. All the previous attempts to mimick this hypothesis has failed so far. One of the most significant improvement with regard to the engineering of a penicillinase activity was obtained in the work of Urbach (*109*), where a single amino acid substitution in a highly conserved loop resulted in a 90-fold increase in the deacylation rate. This jump in the evolutionary path was

encouraging but still not sufficient for conferring a penicillin resistant phenotype.

The aim of this PhD project is to test a new scenario for accelerating in a laboratory time scale the evolution of a DD-peptidase into a β -lactamase. Our starting point is the PBP-A from *Thermosynechococcus elongatus* used in the work of Urbach and her colleagues. This Penicillin Binding Protein is homologous to TEM-1 β -lactamase (27% identity) and its 3D structure has been solved (116) making this protein a good candidate to engineer a type A-like β -lactamase activity. However, both rational and random mutagenesis strategies have failed in generating enzymes conferring penicillin resistance to *E. coli*, suggesting that the straight evolution may be too simplistic. Since results from previous evolution attempts (Carole Urbach thesis) have shown that many faster growing colonies in a library of PBP-A mutants have mutations in important residues of the active site, we thought that PBP-A could be slightly toxic to the bacteria and that it could be an obstacle for its evolution. A potential origin of such toxicity would be an interference with the peptidoglycan metabolism. Furthermore, if the hydrolytic power of PBP-A is to be increased when the enzyme evolves toward a β -lactamase, such interference may even worsen if the enzyme does not lose its capacity to recognize its original D-Ala-D-Ala based substrate. Our idea is therefore to create a stepwise route of evolution in which we assess the possibility that the DD-peptidase initially evolves to become a strict penicillin binding protein, unable to recognize its original substrate but still capable of being acylated by penicillin, before attempting its neofunctionalization into a β -lactamase. In other words, our working hypothesis is that a first step of neutralization of the initial activity may be required to allow the second step of neofunctionalization. Indeed, the potential of neutralized mutants to further evolve towards a β -lactamase should be greatly increased if the toxicity problem is minimized.

II. Strategy

In this work, we want to test a scenario of evolution in which we don't consider the evolutionary route only as a trade-off between DD-peptidase and β -lactamase activities but we include the PBP function. If the DD-peptidase initially loses its native function towards the D-Ala-D-Ala substrates while keeping its aptitude to recognize penicillin (exclusive PBP), we speculate that its potential to evolve into a β -lactamase will increase (Figure 30).

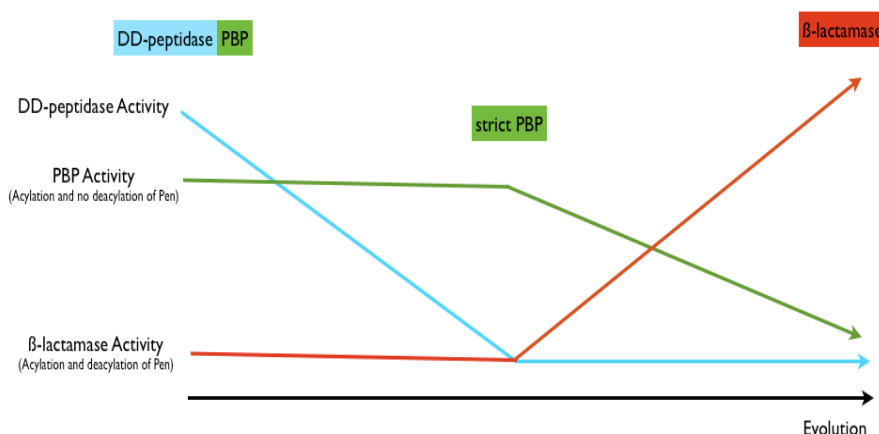


Figure 30: Schematic view of functions trade-off in our hypothesis of evolution (Pen=Penicillin).

The strategy adopted for this project will start by the creation of a diversity of PBP-A mutants using error prone PCR. We will initially evaluate the fitness problems associated with the expression of PBPs. Cultivating the libraries without any other selective pressure than optimized growth should allow the emergence of PBPs with reduced toxicity. We hope that this simple drift experiments will lead to the selection strict PBP variants. However, it does not guarantee that the selected proteins will keep their capacity to react with penicillin. Therefore, a phage display selection

strategy will be set up for the *in vitro* selection of PBPs that are acylated by penicillin. A biotinylated penicillin will be used for capturing the interesting PBP-A variants within the library (Figure 31). In combination with the *in vivo* selection for reduced toxicity, these approaches should afford the selection of the so-called strict PBPs. In the second phase of evolution, the selected variants will be subjected to new rounds of mutagenesis and *in vivo* selection for ampicillin resistance will be applied. Between mutagenesis rounds, the ampicillin concentration will be progressively raised from sub-lethal to lethal doses. The emergence of phenotypically active β -lactamases will be followed and the most interesting mutants will be characterized both *in vivo* and *in vitro*. We will also determine the eventual benefits from the neutralization step for the evolvability of the protein by evaluating the epistatic effects between the mutations selected in the first and the second phase of evolution.

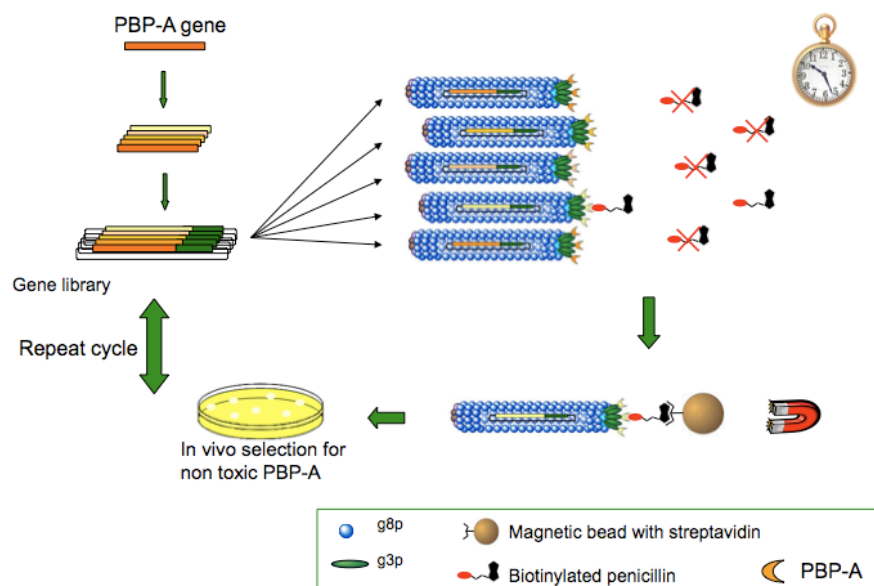


Figure 31: Phage display selection for penicillin acylation. Library of PBPs are phage displayed and panned on streptavidin coated beads after treatment with a biotinylated penicillin.

Results

Chapter I Foreword

According to our strategy, the first task in this project is to build representative collections of PBP-A mutants.

This first chapter covers the following points:

- Optimization of “error prone PCR” protocols allowing different rates of mutations.
- Building and characterization of the PBP-A libraries
- Assessment of the fitness problems within the libraries

Generation of PBP-A diversity and evaluation of fitness problems within the libraries

Heykel Trabelsi and Patrice Soumillon

I. Abstract

In a directed evolution project, the size and quality of the starting library is correlated to the success rate of the following selection experiments. In the first part of this chapter, we describe our method to introduce random mutations into the gene encoding PBP-A of *Thermosynechococcus elongatus* by error prone PCR (epPCR) and evaluate the diversity and the compositions of the libraries. In the second part, we focus on the toxicity assessment of the generated diversity by performing fitness competition experiments under different conditions of protein expression.

Key words: Directed evolution, Error prone PCR, Library, mutations, Drift, Fitness, PBP-A

II. Introduction

Investigating and studying the rules governing protein evolution and understanding how the divergence of protein sequences leads to new structures and functions should improve our knowledge in fundamental biochemistry and our capacity to evolve proteins in a desired way. More particularly, in the case of enzymes, engineering a new catalytic function within an existing scaffold is still very challenging. Beyond the rare cases of success reported in literature, there were many failures (*117*), reflecting our poor and limited understanding of enzymatic performances and mechanisms of evolution.

To bypass the lack of knowledge, directed evolution has arisen since three decades. This method is inspired from natural evolution. It mimics the Darwinian process based on the generation of diversity that undergo natural selection but in a laboratory time scale. This methodology relies on iterative rounds of mutagenesis and selection until the desired trait is reached (*118*). The first step of directed evolution is to build libraries of variants by introducing random mutations in the whole gene of interest or in specific regions of its sequence. The creation of genetic diversity is an important step in a directed evolution project. The size and the quality of the mutant library may greatly impact the success of the experiment (*119*). There are plenty of methods to randomize genes (*120*). Commonly, scientists use error prone polymerase chain reaction (epPCR), oligonucleotide-base mutagenesis and genetic recombination by DNA shuffling (*121–125*). Thanks to the impressive advances in molecular biology and to the need of designing smart libraries, these techniques were developed and optimized to improve the success rate of directed evolution experiments (*126*). The choice of the mutagenesis technique is generally depending on the initial knowledge and context of the protein to evolve. EpPCR, the technique that we used in this

project, is often chosen when there is no a priori on the residues that will have to change for evolving a specific protein in a desired way. The technique is simple, fast and cheap compared to the other techniques.

With the ultimate objective of engineering a β -lactamase activity from a DD-peptidase, we will start by describing, in this chapter, our detailed methodology for building large epPCR libraries starting from the wt PBP-A gene. Then, we will give an overview of the libraries compositions with comparison to other techniques. Finally, since we had suspicions of PBP-A toxicity from a previous work, we will monitor the evolution of the libraries during a genetic drift under different levels of expression and its impact on the following selection or screening procedures.

III. Results and Discussion

One of the crucial requirements for succeeding in this project is the building of libraries that are complex enough to contain a maximum of mutations while maintaining a large proportion of properly folded proteins. This is a matter of compromise since too many mutations per gene will be deleterious for the protein integrity. Additionally, these libraries should contain a good balance of all types of transitions and transversions in order to cover a maximum of possible substitutions and afford a good diversity. Finally, the construction of these libraries should be as simple and cheap as possible.

III.1. The vector: pHDiex

The vector, a gift from Philippe Minard (Université Paris Sud, Orsay), is a phagemid that was designed to allow different expression levels of proteins and the use of the phage display technology (Figure 32). The gene of interest is under the control of both the pLac and pT7 promoters. The second one allows more efficient transcription in strains expressing the T7 polymerase, typically the BL21(DE3) strain. In this study, the genes encoding the mature PBP-A variants will be cloned in frame with an OmpT signal peptide affording the periplasmic exportation. A C-terminal six Histidine (His₆) tag is also encoded for purification purpose. After this tag, a suppressible amber codon (TAG) is introduced and followed by a sequence encoding the C-terminal fragment of the pIII coat protein of the M13 filamentous phage. Therefore, when introduced in a suppressive strain, the phagemid is suitable for the phage display technology whereas, in a non-suppressive strain, the protein is produced as a free protein in the periplasm.

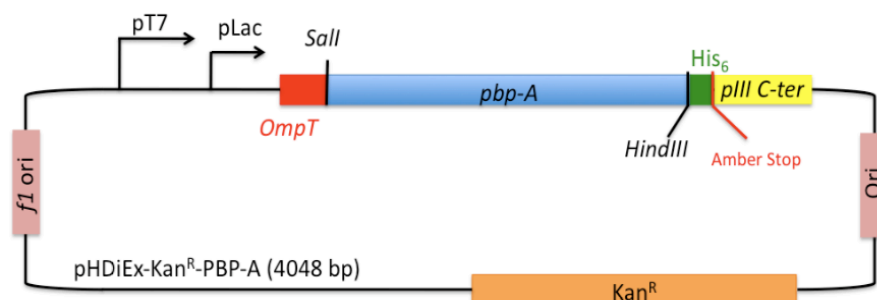


Figure 32: Schematic representation of the pHDiEX vector used for the expression of the PBP-A variants. Kan^R: gene encoding the kanamycin resistance. *pbp-A*: gene encoding the PBP-A. Ori: ColE1 origin of replication. pT7: promoter specific to T7 RNA polymerase. pLac: promoter of the lactose operon. OmpT: signal sequence for periplasmic targeting (from the OmpT protease). Amber stop: suppressible amber stop codon (TAG). pIII C-ter: sequence coding for the C-terminal domain of gene III from M13 phage. f1 ori: single strand origin of replication and packaging signal of f1 bacteriophage.

III.2. Error prone PCR optimization

Error prone PCR is a common mean for generating random mutations at any position of a gene of interest (GOI). This method is very simple, it is based on amplifying the GOI, like in a standard PCR, but under specific conditions allowing a decrease of the polymerase fidelity which promotes misincorporation of nucleotides (127). These specific conditions are: (i) the use of an enzyme with a relatively low fidelity like the Taq polymerase or the Mutazyme (a commercial polymerase engineered for low fidelity); (ii) the use of unbalanced dNTP concentrations; (iii) the presence of a low amount of Mn²⁺. Then, the mutation rate can also be leveled up by increasing the number of replication cycles of the PCR (54, 128–130). The mutagenesis rate can also be tuned by the concentration of Mn²⁺ in the PCR mix (18). After the epPCR, the resulting products must then be cloned into an expression vector for creating the library of randomized genes.

Our initial objective was to optimize the epPCR into two protocols allowing the randomization of the gene, one at a low mutation rate (around 4 mutations per gene) and the other at a high mutation rate (around 12

mutations per gene). The mutation rate is the mean number of mutations per Kb. Depending on the codon composition and the mutations spectrum, between one third and one fourth of the mutations will be synonymous.

A homemade approach was set up to reach the desired number of mutations. The GoTaq polymerase (Promega) was used to introduce mutations since it lacks the 3'→5' exonucleolytic editing activity. This polymerase features an average error rate of 10^{-4} mutation per incorporated nucleotide (*131, 132*).

All the parameters of the epPCR (number of cycles, DNA template concentration and the unbalanced dNTP concentrations) were fixed except for the amount of Mn^{2+} and Mg^{2+} where several concentrations were tested. The choice of reduced concentrations of purines was based on the results obtained by Shafikhani and his colleagues to reduce the bias in the type of base pair changes (A mutates more often to T than C to G) (*133*).

The table 8 summarizes these conditions. The conditions B and D generated an average of approximately 15 and 5 mutations per Kbp respectively. These conditions were used to build the high and low mutation rate libraries.

Table 8: Parameters optimization for epPCR.

epPCR Protocol	A	B	C	D	E
DNA (ng)	1				
[dATP] (mM)	0.2				
[dCTP] (mM)	1				
[dGTP] (mM)	0.2				
[dTTP] (mM)	1				
Nb of cycles	30				
[MgCl ₂] (mM)	6	6	10	12	14
[MnCl ₂] (mM)	0.2	0,5	0	0	0
Number of sequenced clones	6	8	5	7	6
Error rate (per Kb)	7 ± 2	15 ± 5	2 ± 1	5 ± 3	6 ± 3

Error was calculated as the standard deviation (the variance from the average).

Synonymous mutations were about 30 %.

The conditions B and D were used to construct libraries with high and low mutation rates respectively. The error rate (mutation rate) is determined by the number of mutations per Kb.

III.3. Generation of High and low mutations rate libraries

The PBP-A wild type gene served as a template to build two libraries with high and low mutation rates. The idea was to explore the maximum of the sequence space and to increase the chance to have advantageous mutations that are supposed to be rare (*134, 135*). The mutated DNA was ligated into pHDiex vector and transformed into JM109 or XL1Blue strains. Serial dilutions served to estimate the size of both libraries (see methods). 20 clones picked at random were sequenced and their sequences served to estimate the mutation rate (see the Libraries Analyses section (below)).

To estimate the number of different functional variants within a library built by epPCR, it is possible to use some statistical software such as PEDEL (Program for Estimating Diversity in EpPCR Libraries) (*136*). By introducing the template sequence length (L in base pair), the library size

(N) and the mean mutation rate (λ) (assumed to be much less than N) to the software, an estimation of the library composition could be computed (137). The probability in a library that a mutated gene contains exactly x mutations while the mean number of mutation is λ , can be described by the Poisson distribution in the equation below:

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

The table 9 shows that both libraries should contain between 10^8 and 10^9 full length variants without stop codons or indels. These diversities can be considered as high since obtaining more than 10^9 transformants after ligating a PCR fragment into a plasmid is more or less the upper limit.

Table 9: Diversity estimation using the PEDEL software

Property	Estimate	
	Low rate (LR)	High rate (HR)
EpPCR Library type		
Total library size ^a	$9.4 \cdot 10^8$	$1.2 \cdot 10^9$
Mean number of amino acid substitution per gene	3.43	11.57
Number of variant with no indels or stop codons	$8.8 \cdot 10^8$	$6.7 \cdot 10^8$
Unmutated (wt) sequences %	3.93	0.02
Number of distinct full length protein in the library (Poisson estimation)	$5.36 \cdot 10^8$	$6.7 \cdot 10^8$

^a The library size was estimated by performing serial 10X dilutions and counting the colonies then using the following formula: CFU x dilution factor x total volume. Except the library size and the mean error rate all the values were generated by PEDEL software online.

III.4. Minimizing the bias problem

During the epPCR, the polymerase randomly introduces some mutations all along the daughter strand but the frequencies of each possible transitions (GC→AT, AT→GC) and transversions (GC→CG, GC→TA, AT→TA, AT→CG) are not equal (128). There is also a sequence dependant context that favors mutations at specific positions (hot spots) but this is not predictable and the impact of this bias is not easily evaluated. These error-biases mean that libraries have non-random composition with respect to both the position and the identity of changes (127). A way to reduce biases consists in using a mix of two polymerases, Taq and Mutazyme, sharing complementary mutational spectra. However, in this project, our cheaper (compared to Genemorph Mutazyme II Kit) epPCR protocol allowed us to have a good balance between all type of mutations compared to the other approaches (see next part).

There is also another type of bias called the « amplification bias » explained by the fact that during the amplification process of the epPCR, if a mutation appeared in the first cycles and was copied during the next amplifications. This mutation will be over represented in the final PCR product. This could be a serious problem if this mutation is deleterious or affect the presence of advantageous ones. It is not a problem in our case because the epPCR reactions start with one or ten billions of DNA molecules (1 or 10ng of a 1000bp DNA fragment). The last bias, which is one of the limits of the epPCR approach, comes from the nature of the genetic code. Some amino acid substitutions require more than one nucleotide substitution. With a single nucleotide substitution, a given amino acid can be mutated in less than 6 others in average meaning that more than two thirds of all the possible non synonymous mutations require 2 or 3 substitutions in the codon. Increasing both the diversity and the mutation rate could generate more rare mutations. However, the maximum diversity is experimentally limited to the number of

transformants one could obtain (generally between 10^6 and 10^{10}) and a high mutation rate is deleterious for the protein integrity.

III.5. Libraries analyses

Some clones from the two different libraries were randomly chosen for characterization.

We first analyzed the restriction profiles of their plasmidic DNA. The extracted DNA were cut by *HindIII* and *SalI* and the agarose gel profiles of the clones (Fig 33) showed that 86% of the clones (25/29) shared the expected two bands corresponding to the vector and the insert.

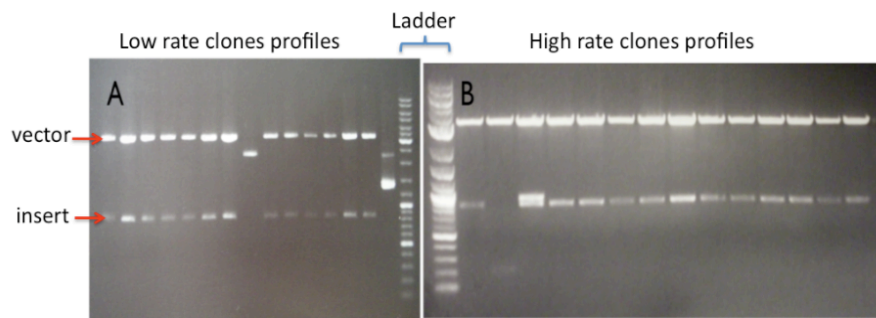


Figure 33: Restriction profiles of clones randomly chosen from the low mutations rate (A) and high mutations rate (B) libraries. DNA restricted by *HindIII* and *SalI*.

Then, 20 clones featuring the expected restriction profile were sequenced for evaluating the number and distribution of the mutations, their types (transversion, transition) and their impact on the gene (synonymous and non synonymous mutations, stop codons, insertions, deletions...).

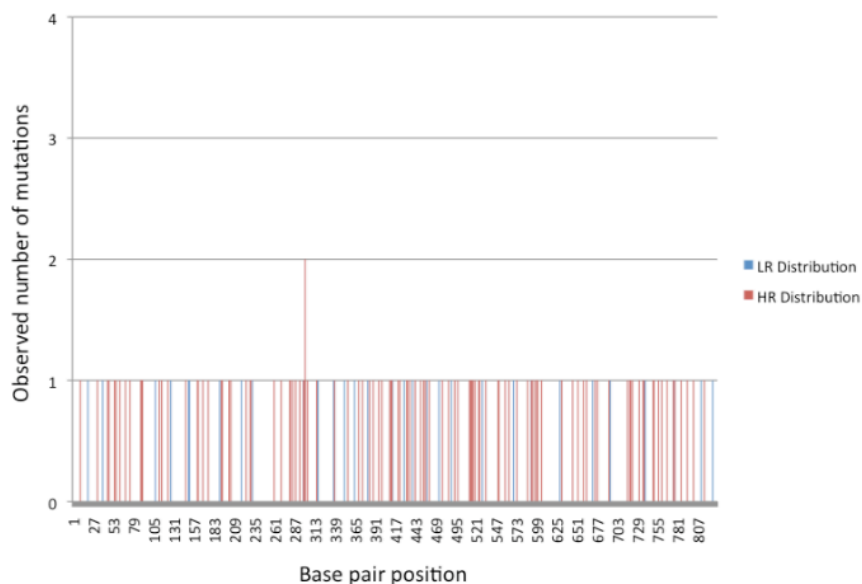


Figure 34: Mutations distribution along the gene. 10 clones from each Library were sequenced and scanned for mutations. HR (red) and LR (blue) stand for the clones analyzed from the high and low mutational rate libraries respectively.

180 mutations were found. The figure 34 shows that they are distributed all across the gene with no detectable hot spots, indicating that most of the PBP-A residues are probably easily accessible to mutation.

Inspired by the mutational spectrum comparison of the different common approaches used to generate diversity (119), the resulting mutations in our low and high rates epPCR protocols were compared to those from the theoretical ideal method, the standard epPCR (low fidelity Taq, $MnCl_2$ and imbalanced dNTPs) and the Mutazyme II (138, 139). The difference between our homemade method and the standard epPCR, is that the low mutation rate is tuned only with the high Magnesium concentration without Manganese (used only for High mutation rate).

It is also to be specified that the results shown in Table 10 should be taken with precautions because of high statistical variance associated with a limited sampling of mutations, they suggest that our two protocols afford equilibrated amounts of transitions and transversions. Among the 210

observed substitutions, 159 AT and 51 GC base pairs are modified and the GC->CG transversion is the most rare. Interestingly, with the High rate epPCR conditions, the mutational spectrum appears even more equilibrated than what is obtained with the commercial Mutazyme II.

Table 10: Mutational spectrum analyses. Low rate and high rate are the epPCR variants used for this project. All values are percentage (%).

Methods	Transitions		Transversions				Indels	Ts /Tv
	AT->GC	GC->AT	AT->TA	AT->CG	GC->CG	GC->TA		
Ideal method	16.7	16.7	16.7	16.7	16.7	16.7	0.0	0.5
Taq ,MnCl ₂ , biased dNTP	38.1	19	27.8	4.0	0.0	6.3	4.8	1.5
Mutazyme II	14	32	15.1	2.3	5.8	22.1	8.3	1
Low rate ^b	22 ± 9.3	7 ± 5.2	42 ± 12.9	14 ± 7.4	0 (<5) ^a	9 ± 6	6 ± 4.8	0.44
High rate ^c	20.1 ± 8.9 (24)	8.4 ± 5.8 (10)	30.25 ± 11 (36)	9.24 ± 6 (11)	10.08 ± 6.35 (12)	10.92 ± 6.61 (13)	10.92 ± 6.61 (13)	0.47

^a Statistically, no observation in 10 sequenced clones indicates that the true proportion is between 0 and 5 per 10 clones (with 95% confidence).

^b For the low rate library, 100 mutations were observed; the percentage is therefore equal to the number of observations.

^c The number between brackets indicate the number of observations.

Ts=Transition, Tv=Transversion , Indels=insertions and deletions. Statistical errors (e) are calculated using this equation: $e = \pm 2 \times \sqrt{n}$, with n as the number of observations.

III.6. Genetic drift of the PBP-A libraries

A previous work in the laboratory suggested that PBP-A is associated with a fitness problem in *E. coli*. A plausible hypothesis is that the activity of PBP-A is somehow interfering with the peptidoglycan metabolism.

To have a better understanding of the situation, the high and low mutation rate libraries were drifted by cultivation under repression, basal expression and induced expression of PBP-A. A number of bacteria corresponding to 10 times the diversity from LR₀ (low mutation rate library) and HR₀ (high mutation rate library) were cultivated in 100 ml of LB medium for 24 h under different levels of expression. 100 µl from these cultures served to inoculate fresh media (100 ml). The drift was performed for 10 successive cultures by reinoculating a fresh medium every day (Figure 35).

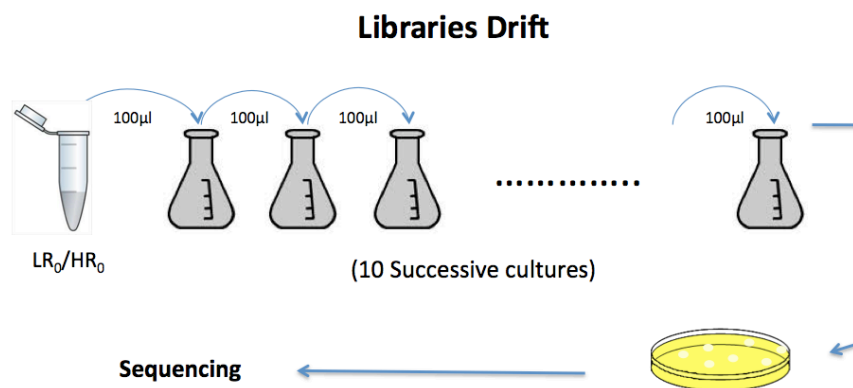


Figure 35: Schematic view of the library drift protocol. LR₀ and HR₀ are the first epPCR libraries built with low and high mutation rates respectively. The volume of the culture is 100 ml.

After plating serial dilutions on Petri dishes, colonies were randomly picked and their DNA were extracted and sequenced. As shown in Table 11, genes encoding truncated proteins are selected under basal and induced expression

conditions and there is a correlation between the expression level and the proportion of full-length genes that remain in the libraries after the drift.

Table 11: Sequences analysis of the drifted libraries under different conditions.

Cultivation conditions	Sequenced colonies ^b	Full length genes (%)	Truncated proteins (%)	
		(no stop codons)	Indels	Stop codons
Repression (LB-Kan-Gluc) ^a	25 (23)	88	12	0
Basal expression (LB-Kan)	20 (20)	50	30	20
Induction (LB-Kan-IPTG)	20 (19)	25	50	25

^a Gluc=glucose 1%; Kan=Kanamycin, ^b the sequenced clones are a mix from low and high mutation rates libraries, the number of different sequences are between brackets.

We then scanned all the mutations in the full-length genes resulting from the drifts in the three expression modes. Interestingly, some mutations were more represented than others indicating the selection of specific clones without any selection pressure for a proteic function. Since this selection appeared as a key phenomenon in the PBP-A conversion into a β -lactamase, it will be discussed in more details in the next chapter.

IV. Conclusion

The success of any directed evolution experiment depends on the quality of the starting Library. Thanks to the molecular biology techniques improvement, there are nowadays many approaches for creating large libraries.

In this project, we used the most common approach, epPCR, affording the introduction of mutations of all the residues of the target gene. However, the method generates a biased mutational spectrum. In this work, we have set up a « home made » epPCR protocol that appears to reduce this bias toward a more equilibrated ratio of transitions over transversions. We have tuned the conditions to control the mutation rate (4 and 12 for low and high mutation rate libraries respectively).

Libraries characterization after a genetic drift under different conditions have shown a positive correlation between the level of expression and the selection of truncated or rearranged clones. The experiments were performed using a high copy number vector and the PBP-A variant genes are under the control of a pLac promoter known to be leaky even when repressed (*140*, *141*) (see figure 1). In fact, a previous work in the lab showed that, using the same promoter under repression condition, the level of TEM-1 β -lactamase expression was 15% compared to the IPTG induced conditions (Valery Mathieu thesis). Therefore, in order to minimize biases upon cultivation, it has been decided to perform all the subsequent drifts and selections under repression conditions (with glucose). Supporting this choice, it was also reported in the literature that highly expressed proteins evolve more slowly than lowly expressed ones (*142–144*).

V. Materials and methods

General

GoTaq polymerase was purchased from Promega. The T4 DNA ligase was from New England Biolabs. Restriction enzymes were from Roche. Ampicillin and Tetracyclin were from Applichem, Kanamycin from Gibco.

Primers

The oligonucleotides were from Eurogentec.

Forward : 5'-TTGTGTGGAATTGTGAGC

Reverse : 5'-CCCCTTATTAGCGTTTGC

Bacterial strains

JM109: *endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB+ Δ(lac-proAB)*
e14- [F' *traD36 proAB+ lacI^q lacZΔM15*] *hsdR17(rK-mK+)*

XL1BlueMRF': *Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44*
thi-1 recA1 gyrA96 relA1 lac [F' *proAB lacI^qZΔM15 Tn10 (Tet^R)*]

DNA extraction and purification

DNA from plasmids was extracted using the High Pure Plasmid Isolation Kit (Roche) and the DNA resulting from PCR was purified using the High Pure PCR Product Purification Kit (Roche).

DNA electroporation and titration

DNA libraries were electroporated by aliquots of 100 ng in 100 μl of *E. coli* electrocompetent cells (JM109 or XL1BlueMRF') and suspended in 900 μl of SOC pre-warmed at 37°C. The cells were then incubated for 1 hour at 37°C, 180 rpm. The tubes were pooled together and 10 μl were sampled for titration. The cells were centrifuged for 15 min at 4000 rpm and the pellet

was resuspended in 1 ml of LB and spread on large LB-agar plates (250 ml) containing the adequate antibiotics. After an overnight incubation at 37°C, colonies were harvested with 5 ml/plate of fresh LB medium and stored at -80°C in 20% (v/v) of glycerol. 10 µl sample was used to estimate the library diversity. 10 fold serial dilutions were made in a final volume of 1ml LB. For each dilution, 100 µl were spread on LB-agar plates supplemented with the appropriate antibiotics. Plates were incubated overnight at 37°C and then, the colonies were counted and the total diversity was estimated using the following formula: CFU x dilution factor x total volume.

Preparation of electrocompetent cells

1 liter of fresh LB medium supplemented with the appropriate antibiotics was inoculated with 1ml of an overnight preculture. The whole was incubated at 37°C and 180 rpm shaking. When the optic density at 600 nm (OD_{600}) reached 0.7 (exponential growth phase) the culture was immediately centrifuged at 4000 rpm 4°C for 10 min. The bacterial pellet was washed successively twice with 400 ml of sterile MilliQ water and once with 20 ml of sterile 10% glycerol (v/v) by resuspending each time the pellet and centrifuging at 4000 rpm, 4°C for 10 min. Then, after the last centrifugation the pellet was resuspended in 1 ml of sterile 10% glycerol and 100 µl aliquots were immediately frozen in liquid nitrogen and stored at -80°C. Note that all the steps after the culture, were carried out at 4°C and the solutions and the centrifugation bottles were beforehand chilled and kept cold on ice. The competence of the cells was assessed by electroporating 2 ng of pUC19 Amp^R (New England Biolabs) and serial dilutions were plated in LB-Amp (100 µg/ml) Petri dishes. After an overnight incubation at 37°C, the competence was defined as the number of colonies obtained by µg of electroporated DNA.

Libraries construction

The wild type PBP-A gene cloned into a pHDiEx vector was used as a template for error-prone PCR. Random mutagenesis was performed using GoTaq Polymerase. The PCR was optimized to ensure a 0.5-1% error rate per gene (1 to 5 amino acids changes) for low mutation rate library and a 2-4% for high mutation rate library. Primers (Fwd and Rev), which flank the 5' *SalI* and 3' *HindIII* restriction sites, were used for amplification. Briefly, 1 ng of vector DNA was mixed with 0.2 mM of dNTP mix (2 mM dATP and dGTP and 10 mM dCTP and dTTP), 0.3 μ M of each primer, 1X GoTaq buffer, 6 mM MgCl₂ (or 12 mM for low mutation rate), 0.5 mM MnCl₂ (not used for low rate library) and 2.5 U of GoTaq polymerase in 50 μ l. The PCR was incubated at 95°C for 4 minutes followed by 30 cycles of 95°C for 1 min, 58°C for 1 min, 72°C for 1 min, with a final extension at 72°C for 7 min. The resulting PCR was purified using high pure PCR product purification Kit (Roche) and digested for 2 hours with *SalI* and *HindIII* at 37°C. The digested product was purified with Chromaspin 400 (Clontech) and ligated into *SalI/HindIII* digested pHDiEx vector using T4 DNA ligase over night at 16°C with a molecular ratio insert/vector of 3:1. The resulting library was directly transformed into electro-competent JM109 or XL1Blue MRF' cells and plated on LB-agar supplemented with Kanamycin (50 μ g/ml).

DNA sequencing and sequences analyses

The DNA samples extracted from the clones were sent to MacroGen company sequencing service. The PBPs in the pHDi-Ex vector were sequenced using pHDi-Ex Fwd and Rev primers.

The DNA sequences were analysed using the GENEIOUS software (downloaded from <http://www.geneious.com>).

Agarose gel electrophoresis

DNA samples were diluted in 6X loading dye (Thermo scientific) and then loaded in the wells of a 1% Agarose gel (w/v) supplemented with Ethidium bromide (final concentration of 0.5 g/ml) in a TAE buffer (Tris-Acetic acid-EDTA). The DNA fragments were then separated during 40 min by electrophoresis at 120 V and 70 mA.

Chapter II Foreword

After creating libraries of PBP-A with different mutation rates, the observation that specific mutants are selected without any selective pressure for catalytic activity drove the project into a new direction. The initial phage display strategy that was planned for selecting exclusive PBPs was abandoned and replaced by a so-called neutralization genetic drift.

This chapter explains our methodology and approach to select intermediates in the evolutionary path toward phenotypically active β -lactamase.

Aspects of protein fitness and evolution are also discussed.

Neutralization of an Enzyme Promotes Its Neofunctionalization

Heykel Trabelsi, Thomas Delcourt, Carole Urbach[†] and Patrice Soumillion^{*}

Unit of Biochemistry, Biophysics and Genetics of Microorganisms,
Institute of Life Sciences, Université catholique de Louvain, Place
Croix du Sud, 4-5, bte L7.07.06, B-1348 Louvain-la-Neuve, Belgium

[†]Present address: MedImmune, Cambridge, UK

^{*}To whom correspondence should be addressed.

Email: patrice.soumillion@uclouvain.be

This paper is in preparation

Contribution :

HT did the main experimental work and supervise the work of TD.
CU performed the biochemical characterization presented in Table 13.
TD contributed to the growth competition experiments

I. Abstract

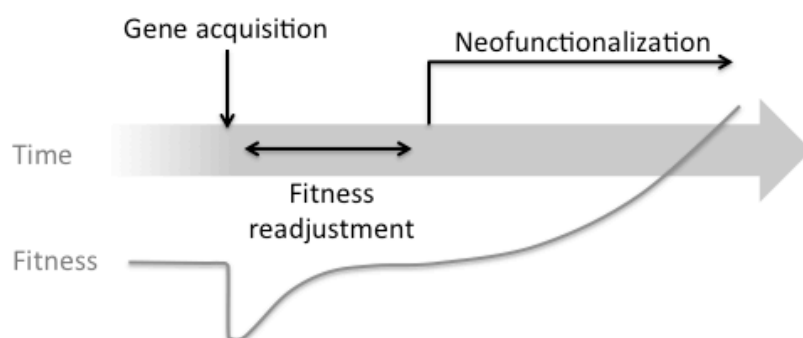
It is generally accepted that gene acquisition by duplication or horizontal transfer is preceding the evolutionary birth of enzymes endowed with new functions by mutations and selection. Just after gene acquisition and before phenotypic neofunctionalization, there may be a time for fitness readjustment that could impact the evolutionary potential of the encoded protein. Here, we show that fitness readjustment after horizontal transfer of a DD-peptidase encoding gene follows a specific trajectory leading to the neutralization of the original activity by active site reshaping without mutating the essential residues for catalysis. We also show that the neutralized mutants have acquired the potential to further evolve into phenotypically active penicillinases. Expression of newly born enzymes is associated with a fitness cost indicating that, besides antibiotic resistance, the new hydrolytic activity is creating a fitness problem to the host. Overall, our results are shedding light on an unexplored property of enzymes, i.e. their security against interfering activities. Securing enzymes may constitute an important obstacle for their natural evolution as well as for their engineering in the laboratory.

Keywords:

Directed evolution, Fitness, Neutralization drift, Penicillin Binding protein, β -lactamase

II. Introduction

Understanding the emergence of new catalytic functions in enzymes is an essential objective for unraveling the origins of these unsurpassed catalysts and their molecular mechanisms of evolution. The evaluation of evolutionary trajectories in the laboratory using Darwinian processes of mutation and selection is a way to provide insights in this field of investigations. Such approach has notably been used for highlighting the promoting effects of catalytic promiscuity (56, 145) neutral drift (69, 70) and chaperoning (146, 147) in enzyme neofunctionalization. The biochemical characterization of mutated enzymes at intermediate stages of neofunctionalization also allowed the description of epistatic effects and trade-offs in specific courses of evolution. Notably, the key role of generalist enzymes endowed with both the old and the new function has been pinpointed several times (54, 148). A common experimental scheme in these studies is to start from a functional enzyme for evolving it toward a new selectable function. The starting function is always maintained before the selection for the new one begins. However, at the start of a natural evolutionary trajectory, just after horizontal gene transfer or gene duplication, the function of the encoded protein is not necessary anymore and may even be associated with some fitness cost. Therefore, the gene may rapidly accumulate mutations for readjusting the original host's fitness (Scheme 1). Although the idea of such a functional neutralization incurring before neofunctionalization is counter intuitive from a Darwinian perspective, this process may facilitate the birth of a new enzyme, especially if the old function and the new one are somehow mutually exclusive.



Scheme 1. The hypothesis of fitness readjustment after gene acquisition (horizontal transfer or duplication) and prior to neofunctionalization.

III. Results and discussion

For testing this evolutionary scenario, we have chosen the conversion of a D-Alanyl-D-Alanine-peptidase into a serine beta-lactamase. DD-peptidases are involved in the biosynthesis and maintenance of the bacterial cell wall and catalyze the hydrolysis or aminolysis of the peptide bond between the two C-terminal D-Alanine residues of peptidoglycan precursors (94, 149, 150). In the mechanism, the essential nucleophilic serine of the active site is first acylated with the concomitant release of the terminal D-Alanine and a second nucleophilic displacement with a water molecule or an amine is releasing the free enzyme and completing the catalytic cycle. On the other hand, serine β -lactamases confer bacterial resistance to beta-lactam containing antibiotics such as penicillins or cephalosporins by rapidly hydrolyzing their cyclic amide bond following a similar mechanism of acylation and deacylation (94). The active site of the acyl-enzyme is however much more crowded since the cyclic nature of the substrate prevents the release of the leaving group upon acylation. β -lactam antibiotics are structurally similar to the D-Alanyl-D-Alanine group so they are also capable of acylating the active site of DD-peptidases but the deacylation step

is between 10^6 and 10^9 fold slower compared to beta-lactamases, making them act as covalent inhibitors. For this reason, DD-peptidases are called Penicillin Binding Proteins (PBPs). Because the peptidoglycan is most probably more ancient than the antibiotics targeting its biosynthesis and because DD-peptidases and serine β -lactamases are closely related, as evidenced by a common fold and conserved active site motifs (SxxK, SDN and K(T/S)G), it is generally accepted that present Class A β -lactamases evolved from one or more ancestral DD-peptidases by acquiring an additional catalytic element capable of activating a water molecule in the very crowded environment of the penicilloyl-enzyme. However, this evolution may represent a risk for the peptidoglycan integrity since DD-peptidases are considered as poorly active and tightly regulated enzymes (151) whereas beta-lactamase are very efficient hydrolases, some of them being almost diffusion controlled. A generalist enzyme may therefore be toxic for the bacterium. The fact that a fitness cost is associated with the expression of some beta-lactamases in the absence of beta-lactam antibiotics (87, 152) supports this hypothesis as well as the failures of many groups that have tried the conversion from a PBP.

The PBP-A from *Thermosynechococcus elongatus* was chosen as the starting enzyme. It belongs to a family of cyanobacterial PBPs that is the phylogenetically closest family to class A β -lactamases (153). An important difference between PBP-As and class A β -lactamases is the presence of a six-residues shorter omega-loop flanking the active site of PBP-As. In class A beta-lactamases, this loop carries an essential catalytic base (Glu166) responsible of water activation in the deacylation step. In PBP-A, the structurally aligned position is occupied by a leucine (Leu158) (116). Mutation of this leucine into a glutamate results in a 90-fold increase of the beta-lactamase activity. However, with a k_{cat} of 10^{-2} s^{-1} , this L158E mutant is

still about 10^5 -fold less active than a natural beta-lactamase and unable of conferring any penicillin resistance to *Escherichia coli*, although active site superposition with the TEM-1 β -lactamase is quasi perfect. Prior to this report, we have tried hard to further evolve PBP-A into a phenotypically active enzyme but all the selections performed on large libraries of mutants generated either by error prone PCR (epPCR) or insertions of randomized omega loops failed (Carole Urbach thesis).

Following our hypothesis of fitness readjustment after horizontal gene transfer or gene duplication and before neofunctionalization, we decided to perform drifting experiments on epPCR libraries of PBP-A without selective pressure for any catalytic activity. The genes were cloned in a high copy expression plasmid under the control of a pLac promoter. Two libraries of $9,4 \times 10^8$ and $1,2 \times 10^9$ transformants were created with, respectively, an average of 4 and 12 mutations per gene. The libraries were grown for 10 days with a 1000-fold dilution in fresh medium every 24 hours and under the lowest level of expression (without IPTG and with glucose for catabolic repression). Surprisingly, two active site mutations, E96V and L158Q, were particularly enriched after the drift with respectively 70% and 50% occurrence in the sequences of randomly picked clones. As surprisingly, no mutations in the essential active site motifs (SxxK, SDN and KTG) were observed, suggesting a conservation of the catalytic machinery. Moreover, the average number of mutations per gene as well as the percentage of stop codons and frameshifts dropped upon drifting (Table 12), suggesting that the expression of truncated or highly mutated proteins is associated with a significant fitness cost. The selected clones were subjected to a second round of mutagenesis and drift without activity selection. As shown in Table 12, the same trend was observed with both the E96V and/or L158Q mutations observed in all drifted clones. A third mutation more remote from the active site (V263A) was also further enriched. At this stage, we hypothesized that

the PBP-A activity was the source of a fitness cost and that active site reshaping was the best solution for correcting the problem. Interestingly, mutation of essential catalytic residues that should result in enzyme inactivation was not as optimal. A possible explanation is that such mutation will destabilize the protein by disrupting the network of hydrogen bond interactions involving conserved residues and water molecules in the active site of these enzymes (16). This may create a folding problem that is known to generate a stress in living cells. A similar stress can also be at the origin of the observed purging of truncated and highly mutated genes.

Table 12: Mutation analysis after 2 rounds of epPCR and Drift according to the low (A) and High (B) mutation rate process (indels=insertions and deletions).

A	Round1		Round2	
	Before the Drift	After the Drift	Before the Drift	After the Drift
% of E96V	0 (0/12)	70 (7/10)	87,5 (7/8)	100 (10/10)
% of L158Q	0 (0/12)	50 (5/10)	62,5 (5/8)	80 (8/10)
% of E96V and L158Q	0 (0/12)	50 (5/10)	50 (4/8)	80 (8/10)
% of V263A	0 (0/12)	40 (4/10)	62,5 (5/8)	80 (8/10)
% of Indels and stop codons	33,33 (4/12)	20 (2/10)	25 (2/8)	10 (1/10)
% of different clones	100 (12/12)	90 (9/10)	100 (8/8)	60 (6/10)
Average number of mutations	5	4	7	4
Number of sequenced codons	3300	2750	2200	2750

B	Round1		Round2	
	Before the Drift	After the Drift	Before the Drift	After the Drift
% of E96V	0 (0/10)	80 (8/10)	83,33 (10/12)	100 (10/10)
% of L158Q	0 (0/10)	60 (6/10)	66,66 (8/12)	70 (7/10)
% of E96V and L158Q	0 (0/10)	60 (6/10)	66,66 (8/12)	70 (7/10)
% of V263A	0 (0/10)	30 (3/10)	50 (6/12)	70 (7/10)
% of Indels and stop codons	30 (3/10)	20 (2/10)	25 (3/12)	20 (2/10)
% of different clones	100 (10/10)	80 (8/10)	75 (9/12)	70 (7/10)
Average number of mutations	12	9	17	10
Number of sequenced codons	2750	2750	3300	2750

E96 and L158 residues contribute in shaping PBP-A's active site (Figure 36). To investigate the origin of fitness optimization, an enzyme carrying both the E96V and L158Q mutations (PBP-A_{VQ}) was purified for biochemical characterization. As many other PBPs, PBP-A is not active on small D-Ala-D-Ala containing peptides but its hydrolytic activity can be evaluated with thiolester surrogates.

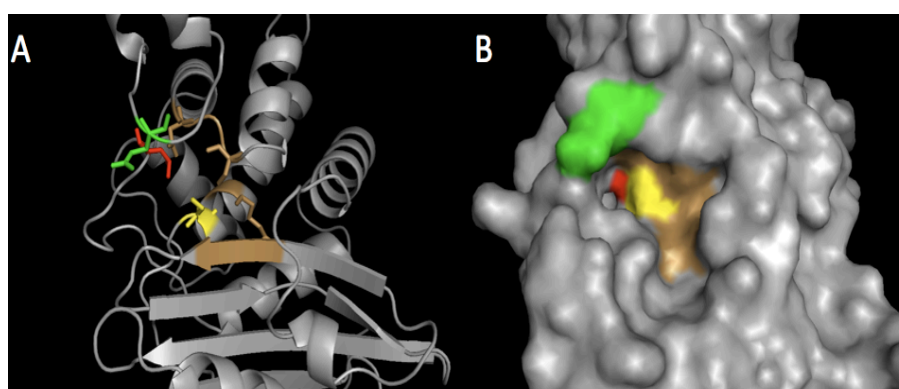
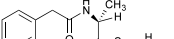
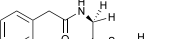
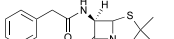


Figure 36: Localisation of the E96V and L158Q mutations. A and B represent respectively the ribbon view and a surface view of the active site. The conserved active site residues are in bright orange with the exception of the catalytic Serine highlighted in yellow. Leu158 and Glu96 are coloured in red and green respectively.

As shown in Table 13, the PBP-A_{VQ} thiolesterase activity on S2d substrate is reduced by around 300-fold. Interestingly, this effect is mainly due to a recognition defect at the level of the methyl side chain of the penultimate D-Alanyl residue. Indeed, with the glycyl-based S2a substrate, the activity drops by only a factor 9. Moreover, PBP-A_{VQ} is still fully active as a penicillin binding protein with an even faster acylation rate compared to wild type PBP-A, confirming the conservation of a functional catalytic machinery.

Table 13: K constants of PBP-A and of the neutralized PBP-A_{VQ} variant

 s2d	 s2a	 Penicillin G		
k_{cat}/K_M ($\mu M^{-1}s^{-1}$)		$k_{acylation}$ ($\mu M^{-1}s^{-1}$)	$k_{deacylation}$ (s^{-1})	
PBP-A	350 ± 80	26 ± 6	170 ± 40	(1.5 ± 0.4) × 10 ⁻⁴
PBP-A _{VQ}	1.1 ± 0.2	2.9 ± 0.7	> 900	(0.61 ± 0.15) × 10 ⁻⁴

When superimposing the structure of penicillin with a D-Alanyl-D-Alanine substrate (fig. S1), a hydrogen is occupying the position of the penultimate methyl side chain, confirming the fact that the mutations essentially affect the recognition of this group. In the structure of the penicilloylated PBP-A (2j8y), this hydrogen is pointing toward a small cavity partially built by the Glu96 and Leu158 residues (Figure 37). It is therefore not surprising that reshaping this cavity may impair D-Alanyl-D-Alanine binding while maintaining penicillin recognition.

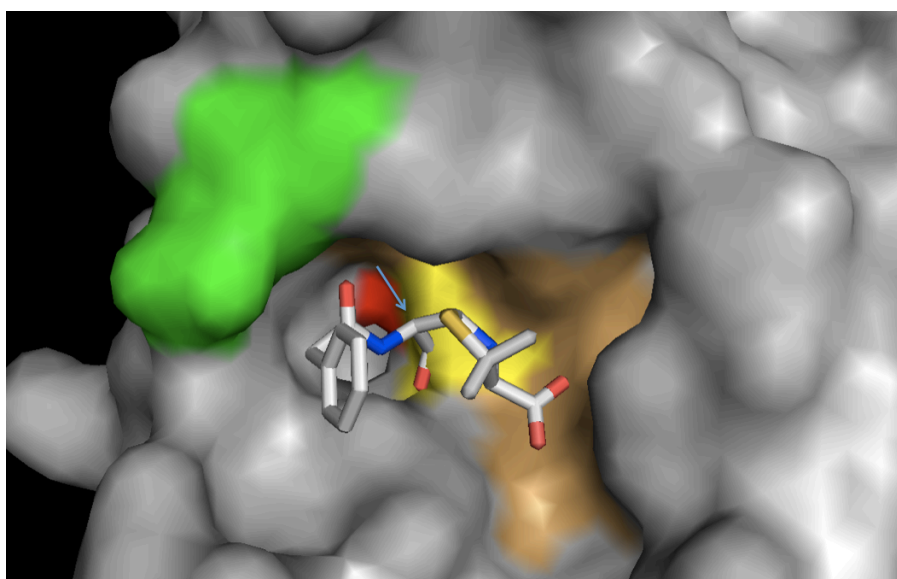


Figure 37: A surface view of the penicilloylated PBP-A. The active site residues are in bright orange, the catalytic serine in yellow, Leu158 in red, Glu96 in green and the penicilloyl group in sticks. The arrow points toward a hydrogen occupying a position that should be occupied by the methyl side chain of the penultimate D-Alanyl residue in the corresponding acyl-enzyme intermediate formed with the natural substrate.

Overall, our results demonstrate that the fitness readjustment has led to the selection of neutralized PBP-A mutants featuring a reshaped active site that prevents recognition of its natural substrate while preserving its full reactivity towards penicillin.

Are the neutralized enzymes more prone to conversion into a β -lactamase? To address this question, successive rounds of mutagenesis (high and low rates of errors) and selection were performed starting from the DNA of the drifted libraries. From round to round, the concentration of ampicillin was progressively increased from sublethal (1 $\mu\text{g/ml}$) to lethal (5 $\mu\text{g/ml}$) doses. The evolutionary trajectory is schematized in Figure 38.

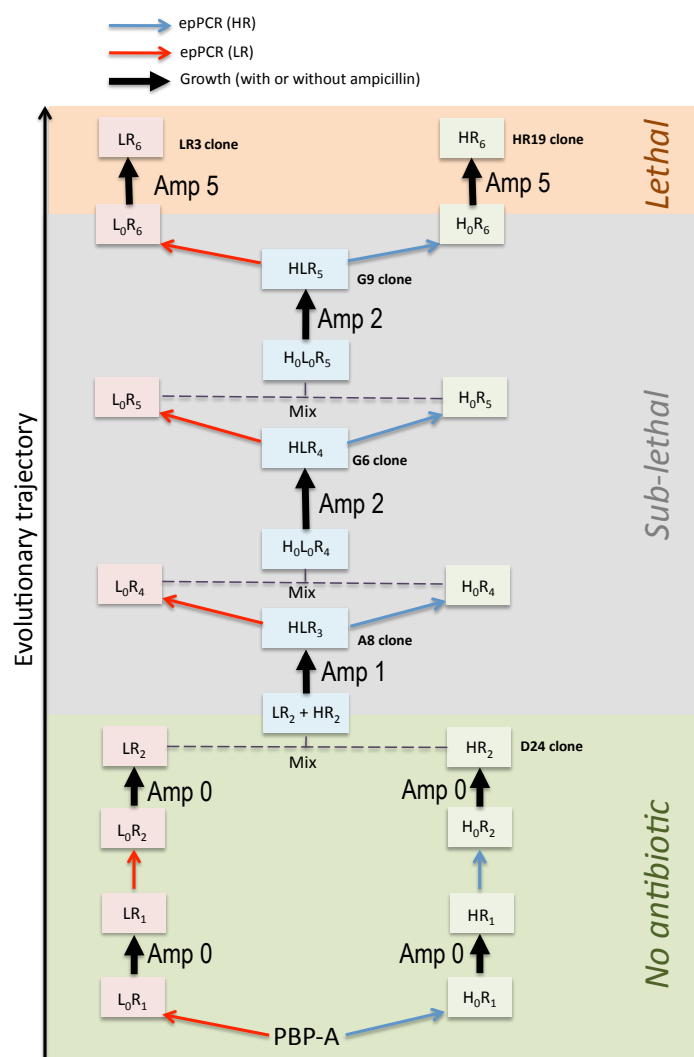


Figure 38: Evolutionary Process. Boxes represent libraries and unboxed names represent monoclones. LR and HR are for low and high mutation rate respectively; The index number (0) after L or H means library after the epPCR; L or H without index number means library after drift or ampicillin selection; the index number after R is for the round of evolution; HLR means mixed HR and LR library; some representative clones that were used for characterization are shown besides some libraries.

Using this selection process, we were, for the first time, able to generate ampicillin resistant transformants (Figure 39A) whereas ampicillin selections from similar libraries built on wt-PBP-A were unsuccessful. Back-transformations experiments indicated that the phenotype was brought by the plasmids.

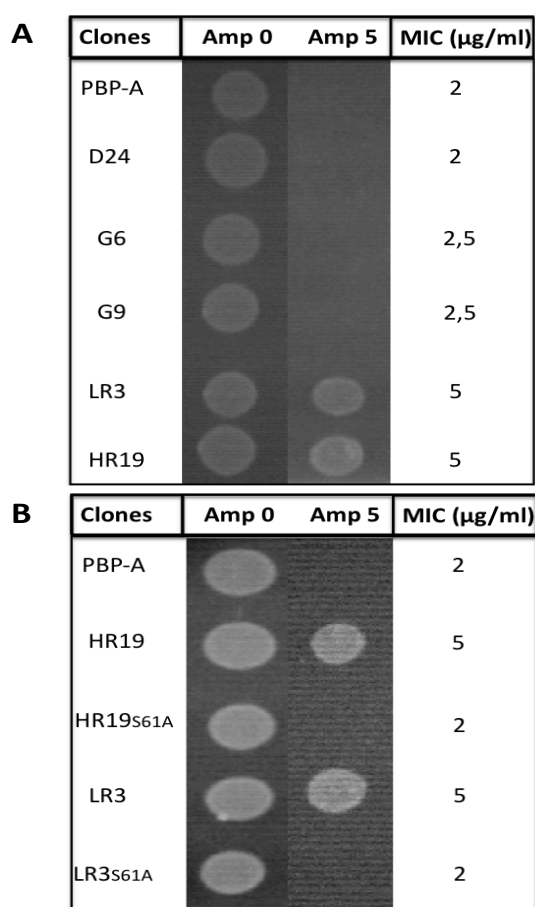


Figure 39: Ampicillin resistant phenotype evaluation of the representative clones from the evolutionary process (A) and the catalytic mutants (B).

The sequencing of selected clones indicated the enrichment of specific mutations (Table 14). The E96V, L158Q and V263A mutations remained conserved throughout the evolution and no other mutations in the region of the active site were observed. Nevertheless, two mutations (D171V and D234E) were frequently observed and progressively accumulated in the most active clones. For example, clones LR3 and HR19 share all of them although they emerged from two independent evolution experiments with low and high mutation rates respectively.

Tableau 14: MIC and mutations of the representative mutants compared to the PBP-A_{WT}.

Mutant	MIC (µg/ml)	[Amp] during the drift (µg/ml)	Number of mutations	Mutations
PBP-A	2	/	0	/
D24	2	0	3	<u>E96V</u> - <u>L158Q</u> - <u>V263A</u>
A8	2.5	1	7	F40Y- <u>E96V</u> -N151D-P169S-P250L-P254L- <u>V263A</u>
G6	2.5	2	4	<u>E96V</u> - <u>L158Q</u> -R190H- <u>V263A</u>
G9	2.5	2	5	<u>E96V</u> - <u>L158Q</u> - <u>D171V</u> - <u>D234E</u> - <u>V263A</u>
LR3	5	5	8	S45P- <u>E96V</u> - <u>L158Q</u> - <u>D171V</u> - <u>D234E</u> -M246I- <u>V263A</u> -Q269H
HR19	5	5	12	V56A-D91V- <u>E96V</u> -Q101H-Q145R- <u>L158Q</u> - <u>D171V</u> -L194Q- <u>D234E</u> - <u>V263A</u> -V267A-K273R

The V263A mutation was already enriched during the drifting experiment and is remote from the active site. This mutation was probably neutrally fixed since its reversion had no impact on the properties (ampicillin resistance and fitness) of the clones. The D171V and D234E mutations target residues that are located at the interface between the α and α/β domains of the protein. It is worth noting that, in the structural alignment between PBP-A and TEM-1 β -lactamase, the orientation of one domain

relative to the other is slightly different (Figure 40, Table 15). Hence, these mutations may affect this positioning.



Figure 40: Structural superimposition of the α domain of the PBP-A (yellow ; PBD code: 2J7V) and the TEM1- β -lactamase (green ; PBD code: 1BTL). Active site residues are presented in sticks mode. The α/β domains (cyan for PBP-A and red for TEM-1) are clearly not well superimposed when α -domains superimposition is optimal.

Tableau 15: rmsd between the α Domains, α/β Domains and the whole structures of PBP-A and TEM1- β -lactamase.

Type of Superposition	Global	α/β Domain	α Domain
Number of aligned residues	219	88	132
r.m.s.d (Å)	1.88	1.83	1.55

Rmsd= Root mean square deviation

Although the PBP-A variants from the most resistant clones could not be purified, we detected low penicillinase activities in periplasmic extracts (Figure S2) and the mutation of the essential serine resulted in the loss of the

resistance phenotype, confirming the involvement of the active site (Figure 39B).

To further characterize the evolution of fitness properties, we performed competition experiments (Figure S3) with several pairs of clones selected at different stages of the evolutionary trajectory. As already described, the initial drift without activity selection clearly led to a fitness optimization correlated with the introduction of E96V and L158Q mutations. Very interestingly, the progeny of these clones that were selected for ampicillin resistance feature a new fitness cost when grown in the absence of ampicillin, indicating a disadvantage associated with the expression of the newly born β -lactamases (Figure 41).

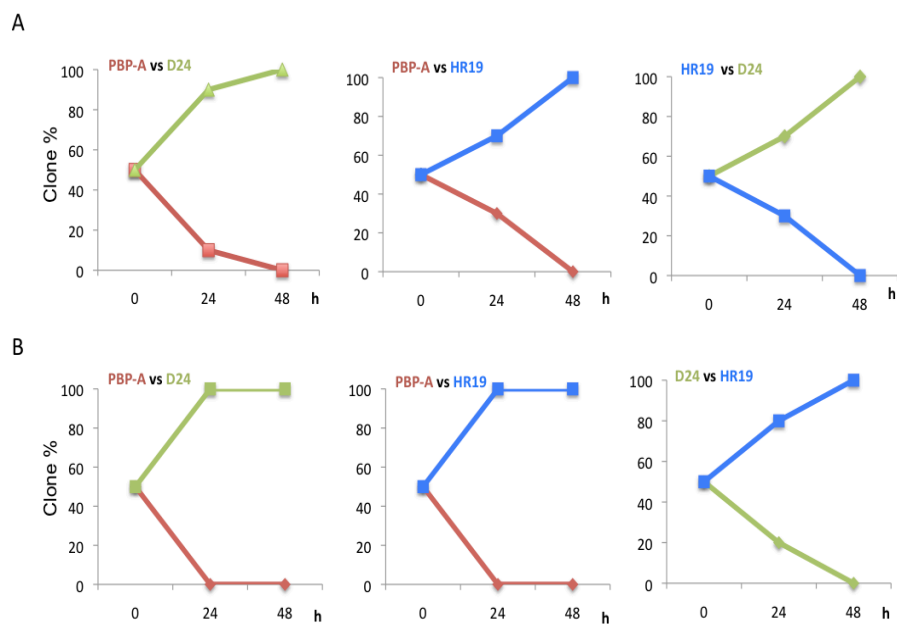


Figure 41: Fitness competitions between PBP-Awt, D24 and HR19 clones. D24 has been selected without ampicillin whereas HR19 was selected with a lethal concentration of ampicillin. The competitions were performed in media without (A) and with ampicillin (3 μ g /ml) (B).

Our results suggest that the initial fitness optimization in the absence of ampicillin was necessary for the subsequent neofunctionalization. We reverted the E96V and L158Q mutations in one of our variants conferring the highest ampicillin resistance (HR19V : this mutant has the same characteristics of HR19 but Ala263 was reverted to valine) and compared their resistance and fitness properties with the parent clone (Figure 42). Both mutations led to a major decrease in both ampicillin resistance (Figure 42B and Figure S4) and fitness (Figure 42A), confirming their key role in the evolutionary history.

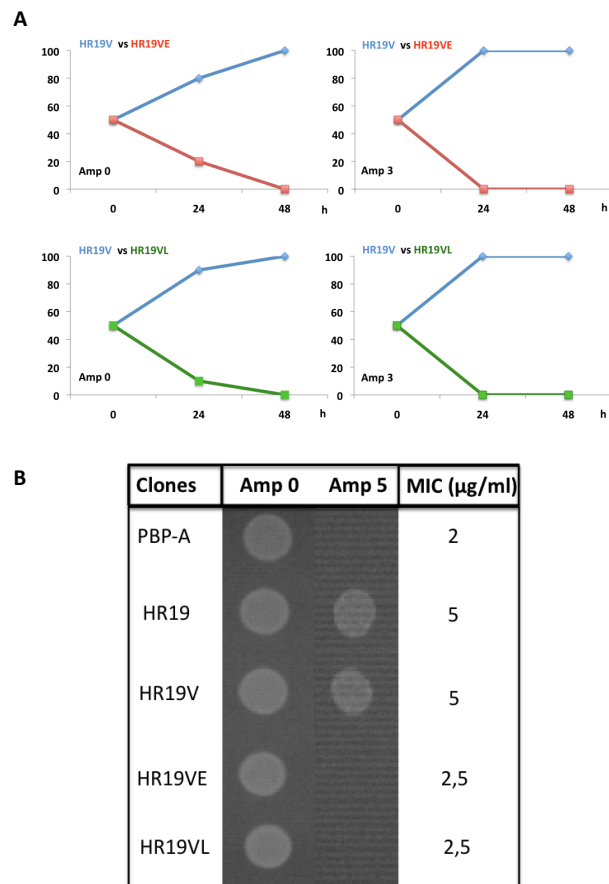


Figure 42: Fitness competitions (A) and analyses of the ampicillin resistant phenotype of the mutants with Val96 reverted to Glu (HR19VE) and Gln158 reverted to Leu (HR19VL) (B).

IV. Conclusion

In conclusion, we have simulated a horizontal gene transfer by cloning a cyanobacterial gene encoding a PBP into *E. coli* and observed a fitness readjustment leading to enzyme neutralization. This neutralization does not proceed anarchically by the selection of any kind of mutations leading to gene or protein disruption. A specific trajectory is favored that selects full-length proteins featuring a reshaped active site while catalytic residues are not mutated. Hence, the protein integrity and the catalytic machinery of PBP-A are somehow evolutionary protected in the absence of functional selection. The stress associated with the expression of unfolded, misfolded and fragmented proteins should explain this surprising observation. Similar experiments with other enzymes will be necessary to evaluate the extent of this new and potentially important concept for molecular evolution. We have then shown that the neutralization drift is necessary for the subsequent emergence of penicillinases that are capable of conferring an ampicillin resistance phenotype indicating that the new function is incompatible with the original one. The fitness associated with the expression of the newly born enzymes is consistent with this conclusion and suggests that the real challenge in this enzyme conversion is the avoidance of interfering activities rather than the engineering of the new catalytic function *per se*. Although catalytic promiscuity was shown as an important trait for the evolvability of some enzymes, our results highlight an alternative and counter-intuitive scenario where promiscuity is potentially dangerous and generalist enzymes are forbidden. Natural enzymes are not just very efficient and specific biocatalysts. They should also be considered as secured by evolution and fitness issues may be underestimated when trying to engineer new enzymatic functions. We hope that the neutralization strategy can be extended to other conversions and will push forward our capacity to create new biocatalysts and our understanding of the driving forces of molecular evolution.

Supporting information

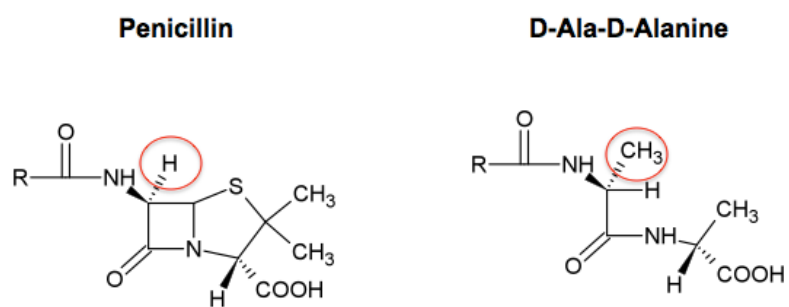


Fig. S1: Structural similarity between Penicillin and D-alanyl-D-alanine

Mutant	Activity x 10 ⁻⁸ Ms ⁻¹
PBP-A	5,4
D24	2.1
A8	4.9
G6	5.3
G9	6.6
LR3	39,6
HR19	44.8

Fig. S2: β -lactamase activities in the periplasmic extracts of the representative mutants from the evolutionary trajectory assessed on benzyl penicillin.

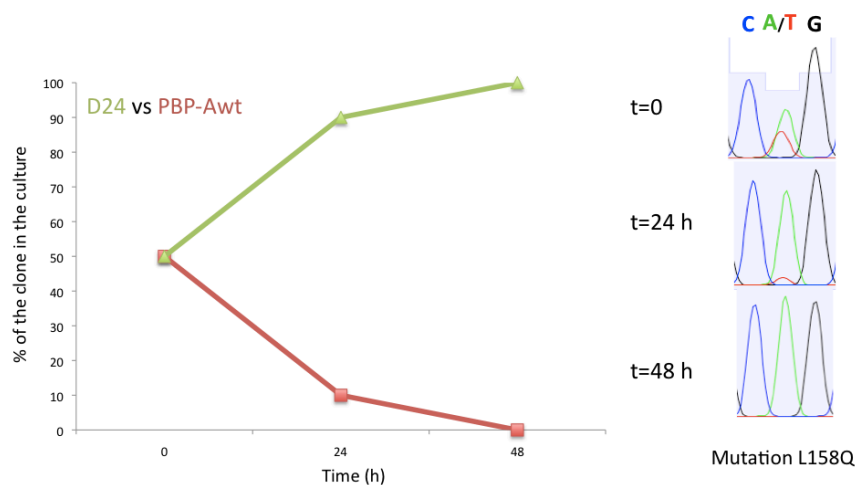


Fig. S3: Growth competition experiment between two clones. The percentage of each clone in the culture at a given time is estimated from the sequence analysis by measuring the relative fluorescence signal of a discriminating base in the electropherogram.

Mutant	Activity x 10 ⁻⁸ Ms ⁻¹
HR19	44,8
HR19V	45,2
HR19VL	2,45
HR19VE	4,7

Fig. S4: β -lactamase activities in the periplasmic extracts of the HR19 clone and three variants for which mutations are reverted to wt residues. V, L and E stand for Val263, Leu158 and Glu96. Activity was assessed on benzyl penicillin.

V. Methods

General

T4 DNA ligase, Taq polymerase and restriction enzymes were provided by Roche. Ampicillin was from AppliChem, kanamycin from Gibco, benzylpenicillin from Sigma.

Primers

The oligonucleotides are from Eurogentec, Liège

Forward PBP-A 5'-TTGTGTGGAATTGTGAGC

Reverse PBP-A: 5'-CCCCTTATTAGCGTTTGC

Frwd Val263 : 5'-AGTGAAGTCATTCGCCAGGTCTCGCGGATGG

Rev Val263 5'-CCATCCGCGAGACCTGGCGAATGAGTTCAC

FWD S61A 5'-TTTCCTGCCGCCGCGACGATTAAGTTTCCCA

Rev S61A 5'-TGGGAACTTAATCGTCGCGGCGGCAGGAAA

FWD V96E 5'-CTCATAGCCCCTGAGGCTGGCACATTG

Rev V96E 5'-CAATGTGCCAGCCTCAGGGGCTATGAG

FWD Q158L 5'-GATTAACAATCCGCTGCCGGATATGAAGG

Rev Q158L 5'-CCTTCATATCCGGCAGCGGATTGTTAATC

Bacterial Strains

JM109: endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB+ Δ (lac-proAB) e14-[F' traD36 proAB+ lacIq lacZ Δ M15] hsdR17(rK-mK+),

XL1BlueMRF': Δ (mcrA)183 Δ (mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F' proAB lacIqZ Δ M15 Tn10 (Tetr)]

Libraries construction

See chapter 1

Site directed mutagenesis

The site Directed Mutagenesis Kit from Stratagene was used to mutate the targeted DNA bases. Briefly, 2 forward and reverse primers containing the new mutation are used to amplify the DNA by PCR. Then, the DpnI is added to the DNA to digest the parental methylated strands (1h at 37°C). The resulting mix is used to transform fresh competent cells. The growing colonies with the new mutation are checked by sequencing.

DNA electoporation and titration

See chapter 1

Library genetic Drift

At least 10 times the size of the library were grown in 100 ml of LB medium containing Kanamycin (50 µg/ml), Tetracyclin (7.5 µg/ml), Glucose (1%) and Ampicillin (0 to 5 µg/ml) at 37°C. After 24 hours, 100 µl were used to inoculate a new 100 ml of fresh medium. At the end of 10 successive cultures, the DNA was extracted and served as a template to generate new libraries that underwent new Drift process. 100 µl at the end of every Drift round were plated in the same LB-Agar medium. Selected clones were sequenced and analysed.

MIC determination

The tested clone was cultivated until an $OD_{600} \approx 0.6$ then serial dilutions were spread on LB dishes containing different concentrations of Ampicillin. The same test was done in liquid culture. A MIC of 3 µg/ml means that the clone can not grow in solid or liquid culture supplemented with 3 µg/ml of the appropriate antibiotic.

***In vivo* selection for ampicillin resistance**

Escherichia coli JM109 or XL1BlueMRF' bacteria transformed with the different PBP-A mutants plasmids, were grown in LB-Kan medium (containing 50 μ g/ml kanamycin) at 37°C, 180 rpm until they reached the exponential phase. Serial dilutions were performed in fresh LB, bacteria were plated on LB-Kan medium containing different ampicillin concentrations.

Fitness competition assay

The fitness was determined by growth competition experiments between pairs of clones expressing different PBP-A mutants in LB media supplemented with 1% glucose and the appropriate antibiotics (Kanamycine, Tetracyclin with or without Ampicillin). For this purpose, XL1blue MRF' *E. coli* strains carrying out two different mutant genes were grown separately overnight from a single colony, diluted 1/100 and regrown to exponential phase ($OD_{600} \approx 0.6$) at 37 °C. Cultures were then mixed in a 1 : 1 ratio in 100 mL of medium (50 μ l each clone at the same OD_{600}). The growth was performed for 24 h at 37 °C. Then, 100 μ l of the culture were used to inoculate 100 ml of fresh media for a second day of growth under the same conditions. This step was repeated for a second day. The DNA of a sample from the initial mix and the growth after day 1 and 2 was prepared and sequenced. The quantification of each competing clones was performed by measuring the normalized-surface ratio of the signals from two different nucleotide at a chosen position in the chromatogram. Normalization of surfaces was performed by imposing the 1 : 1 ratio of the initial mix.

Periplasmic extract preparation

10 ml of bacterial culture (growth under the appropriate conditions) were centrifugated for 10 min at 4000 rpm and 4°C. The pellet was resuspended

in 350 μ l of resuspension buffer (20 mM Tris, 20 % Sucrose (w/v), pH 8). Then, 7 μ l of 0,5 M EDTA solution were added and the mix was incubated for 10 min at room temperature. After a centrifugation for 10 min at 6000 rpm and 4°C, the supernatant was removed and the pellet was resuspended in 500 μ l of a chilled solution of MgSO₄ 5 mM and incubated for 15 min at 4°C. The supernatant was recovered after a final centrifugation for 10 min at 10000 rpm and 4°C and constituted the periplasmic extract.

Activity measurements

The activities of the periplasmic extracts of the mutants were determined at 25°C in 50 mM Hepes buffer (pH 7). The hydrolysis of benzylpenicillin was followed by measuring the change in absorbance as a function of the time at 232 nm ($\Delta\epsilon_{232(\text{Product-Substrate})} = -1030 \text{ M}^{-1} \text{ cm}^{-1}$). 100 μ l of the periplasmic extract were added to 1,9 ml of PenG (0,5 mM) in Hepes Buffer (50 mM, pH 7.5).

Activity towards thiolesters (s2d and s2a) and the determination of the acylation and deacylation rates were described in the reference (15).

Notes

This work was supported by the “iPros” Interuniversity Attraction Pole P7/44 from the Belgian Science Policy (Belspo) and by the Belgian National Fund for Scientific Research (FNRS). H.T. acknowledges a Marie Curie fellowship from the EU-FP7-ENEFP Integrated Training Network.

Chapter III Foreword

The evolutionary process adopted in this project and described in the chapter II allowed the emergence of some interesting intermediates representatives of the different steps (drift and selection).

To better understand the biochemical aspects leading to phenotypically active mutants, it is important to purify and characterize the selected variants.

This chapter covers all the attempts of purification and the estimation of the kinetic parameters of the most evolved mutant.

Phenotypically active PBP variants : Attempts of purification and characterization

Heykel Trabelsi and Patrice Soumilion

I. Abstract

Protein engineering experiments usually give rise to new evolved proteins sharing new features and traits. Further characterization of the engineered variants allows a better understanding of the biochemical and biophysical changes that led to the emergence of these new characteristics. In certain cases, the selected intermediates from an evolutionary trajectory are neither stable nor properly folded. This makes the purification and the characterization of the evolved mutants more complicated. In this chapter, we highlighted our attempts to purify some phenotypically active variants emerged from the previously described selection. We also estimated the K_m and the k_{cat} of the X1 mutant indicating an increase in the deacylation rate by approximately 10^3 - 10^4 fold.

Key words : Purification, Western blot, Mass-spectrometry

II. Introduction

DD-carboxypeptidases and DD-transpeptidases are the key enzymes in peptidoglycan maintenance and biosynthesis (154, 155). These enzymes are the target of β -lactam antibiotics. Actually, they are reacting with the beta-lactam ring to form a stable acyl-enzyme adduct. For this reason these proteins are also called Penicillin Binding Proteins (110) (PBPs). These PBPs could be periplasmic or membrane bound and are involved in different cellular processes such as septation, cell elongation or shape determination (156, 157). To go into more depth in the understanding of their role and mechanism of action, many PBPs were purified and characterized.

Since decades, researchers had been purifying and characterizing PBPs from different organisms such as *E. coli* (158, 159), *Pseudomonas* (160), *Mycobacterium* (161), *Streptococcus* (162), *Cyanobacteria* (109) and others. Besides biochemical characterization, many 3D structures have been solved. In 2014, 244 PBP structures are available in the Protein Data Bank. Interestingly, all the characterized PBPs are wild type proteins or some site directed mutants conceived to understand some structure-function relationships. Despite the several endeavors to increase the very weak β -lactamase activity observed in PBPs or to find an evolutionary pathway leading to the emergence of a phenotypically active beta-lactamase, only a few mutants of such PBPs were biochemically and/or structurally described.

In this chapter, we describe our attempts to purify and characterize some intermediates selected from the neutralization drift and ampicillin selection (see chapter 2). Different approaches were used to detect and purify the mutants. The kinetic parameters of the most evolved variant were also estimated. The objective was to support the phenotypical results obtained *in*

vivo during the evolutionary process by biochemical data. However, we faced serious obstacles in protein expression, production and purification that could explain how difficult it is to get a hand on evolutive intermediates and why this issue is poorly described in the litterature.

III. Results and Discussion

As previously described, the PBP variant genes were cloned into Sall /HindIII digested pHDiEx vector. This construction placed the mutated genes under the control of the dual pLac-pT7 promotor and in frame with a C-terminal six Histidine tag to allow a one step affinity purification. The variants used in this chapter are highlighted in the table 16.

Table 16: Characteristics of the variants used in the purification attempts.

Variant	Number of Mutations	MIC (µg/ml)	Types of mutations	Type of Ampicillin selection
G6	4	2.5	EpPCR	Sublethal (Amp2)
G9	5	2.5	EpPCR	Sublethal (Amp2)
LR3	8	5	EpPCR	Lethal (Amp5)
HR19	12	5	EpPCR	Lethal (Amp5)
HR19V	11	5	EpPCR and reversion of A263V	Lethal (Amp5)
X1	31	5	EpPCR	Lethal (Amp5)

The HR19V clone represents the HR19 mutant with reversion of the Valine 263. This clone share the same characteristics with HR19.

III.1. Origin of the X1 mutant

The X1 mutant is not described in the previous chapter because chapter 2 was written as a manuscript for publication and, since we were looking for a well focused message, we decided not to present the X1 story there.

The X1 clone has been selected from an epPCR library that was generated from the DNA encoding the HR19 clone, i.e. the most active variant.

The HR19 clone's plasmid served as a template for a new round of error prone PCR under high mutation rate (8 mutations per gene for this experiment). The resulting library (1.7×10^7 transformants) was drifted in

repression mode and in presence of ampicillin ($5 \mu\text{g}/\text{ml}$). After a drift of 100 generations, 6 over 10 sequenced clones were corresponding to the X1 clone featuring 31 amino-acid substitutions (Table 17).

Table 17: X1 mutant characteristics.

Mutant	MIC ($\mu\text{g}/\text{ml}$)	[Amp] during the drift ($\mu\text{g}/\text{ml}$)	Number of mutations	Mutations
X1	5	5	31	Q21L-Q23L-V28G-F40Y-V56A-A70T-V80A-T81P-E84D-M88L-D91V-E96V-Q101H-N106Y-Q145R-L158Q-K162R-D171V-I179V-E183G-L194Q-M197S-R198W-D234E-N237Y-Q239K-Y241S-D253G-V263A-V267A-K273R

Compared to its HR19 parent, X1 confers almost the same resistance to ampicillin to the host bacteria (MIC= $5 \mu\text{g}/\text{ml}$) and the activity in periplasmic is slightly increased (rate vs rate for HR19). Interestingly, X1 clone features an improved fitness in the absence (Figure 43A) and the presence of ampicillin (Figure 43B).

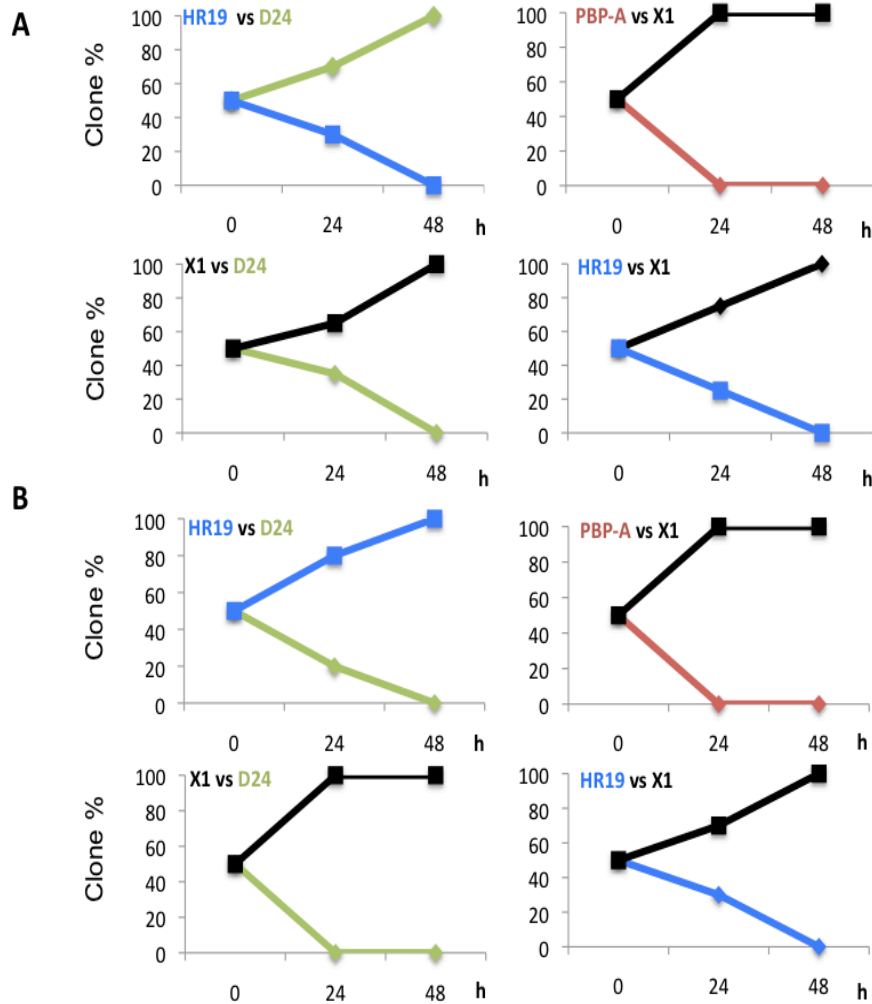


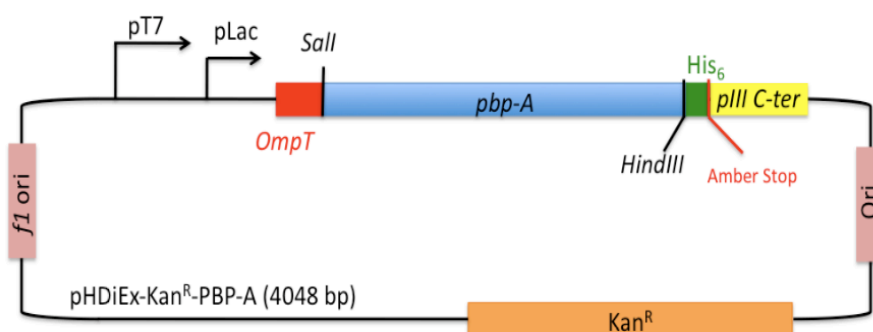
Figure 43: Fitness competitions between PBP, D24, HR19 and X1 clones. D24 has been selected without ampicillin whereas HR19 and X1 were selected with a lethal concentration of ampicillin. The competitions were performed in media without (A) and with ampicillin (3 µg /ml) (B).

This result indicates that the fitness cost associated with the emergence of the phenotype (HR19 birth) can be reduced in the next evolutionary phase. Surprisingly, X1 has 20 additional mutations compared to HR19. The presence of X1 in the starting epPCR library is highly improbable since the average mutations per gene is 8. We suspect that the stress caused by HR19

and the presence of ampicillin has triggered a mechanism of *in vivo* mutagenesis that has been described in the work of Matic group, where it was shown that at the presence of subinhibitory concentration of beta-lactam antibiotic in culture promotes the decrease in replication fidelity (163).

III.2. Expression and purification of mutant proteins

First, the expression of LR3, HR19 and X1 mutants in *E. coli* BL21-DE3 was carried out in the presence of different concentrations of IPTG. In BL21-DE3, the strong pT7 (see the genetic construction below) promoter should be active. It was found that for each mutant varying the IPTG concentration from 0.1 mM to 1mM result in the absence of any detectable β -lactamase activity in the cytoplasmic or periplasmic extracts.



Scheme 2 : The pHDiEX vector used for the expression of the PBP-A variants.

SDS-PAGE analysis of periplasmic extracts after 4 hours of induction at 30°C with 1mM of IPTG were also performed. As shown in Figure 44, it was not possible to distinguish a band corresponding to PBP-A (around 30 KDa) between the induced and the not induced conditions suggesting that the levels of expression are very low or the proteins were lost by proteolysis.

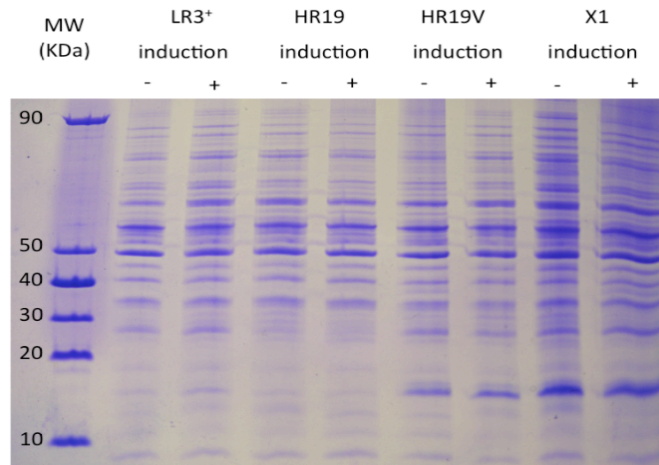


Figure 44: SDS-PAGE analysis of the periplasmic extract of BL21 transformed with the HR19, LR3, HR19V and X1 expression plasmids. (+) For induced with 1mM IPTG for 4h at 30°C, (-) non induced but cultivated for 4h at 30°C.

Further attempts to produce the mutants by varying the temperature and the time of incubation after the induction were unsuccessful.

We then decided to change the strain. The same plasmids were introduced into *E. coli* Top10 where only the pLac promoter should be active. We were able to detect activity toward penicillin (PenG) in the periplasmic extracts (Fig.45).

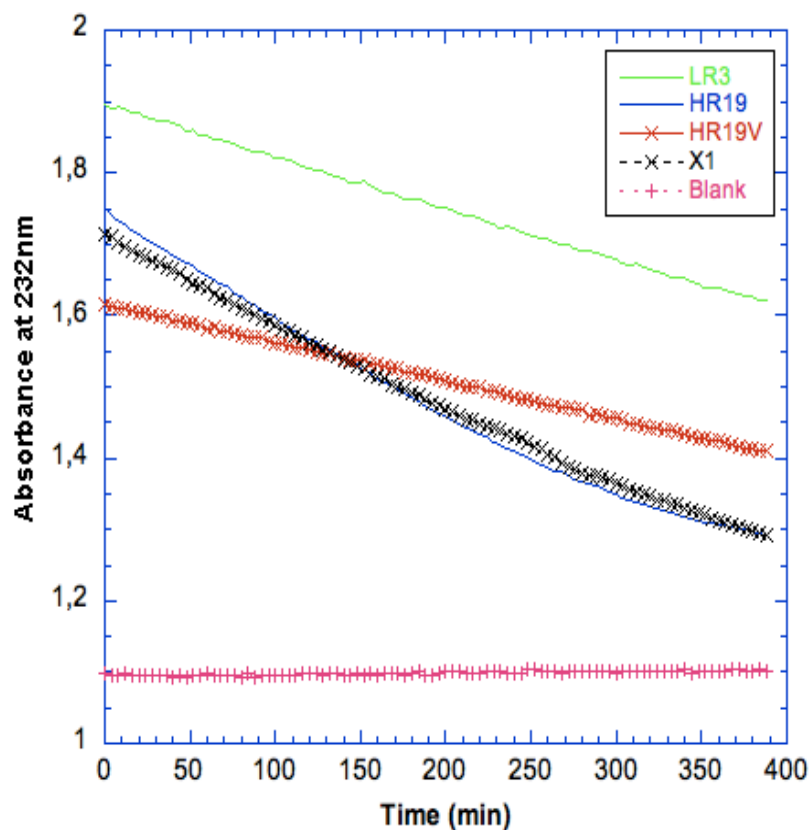


Figure 45: Hydrolysis of PenG by the PBP variants expressed in Top10 strain. For the measurement, 100 μ l from the periplasmic extract were added to 1,9 ml of PenG (0,5 mM) in Hepes Buffer (20 mM, pH 7,4). The absorbance at 280nm of the tested extracts were different (Different protein amounts). The blank was the periplasmic extract of the strain transformed with the empty vector.

Then, after an overnight dialysis in the binding buffer (see methods) the different periplasmic extracts were loaded onto a Ni-NTA column. After several Wash (10 times the column volume), a step by step gradient of imidazole (100 mM steps) was performed to elute the variants. The activities in the flow through, the wash solution and the elution fractions were checked to evaluate the binding of the mutants. We found that almost all the activity was recovered in the flow through, which means that the variants were not able to bind to the Ni-NTA via their His tag tails. This indicates that the tail

is either removed or not accessible. We varied the NaCl concentration in the binding buffer and also tested the addition of 0.1% of Triton X100 but no improvement in the binding was shown.

Since our mutants have many mutations (see table 16), one possibility is that the proteins are mainly unfolded or in a molten globule state with a buried histidine tag. To test this hypothesis, the X1 mutant was produced and the IMAC column was performed in the presence of guanidinium (6M) (see methods). After washing, the guanidinium was removed and imidazole was added for eluting bound proteins. Unfortunately, no activity was detected in the eluate.

We also tried cytoplasmic expression of the HR19 mutant after recloning the gene into a pBAD vector without its export signal. This was previously tested in the laboratory and was successful for the purification of the wild type PBP-A (109, 116). After arabinose induction, the cytoplasmic extract was checked and assayed for penicillin hydrolysis (0.5 mM), but no activity could be detected.

At this time, we decided to put more effort into the detection of the PBP-A variants. The experiments were carried out in XL1Blue strain under repression mode (glucose 1%) in order to keep the same conditions as during the selection. The periplasmic extracts of all these variants were prepared and labelled with BOCILLIN™ FL Penicillin. This fluorescent derivative of penicillin allows the detection of PBPs by fluorescence scanning of SDS-PAGE gels (164).

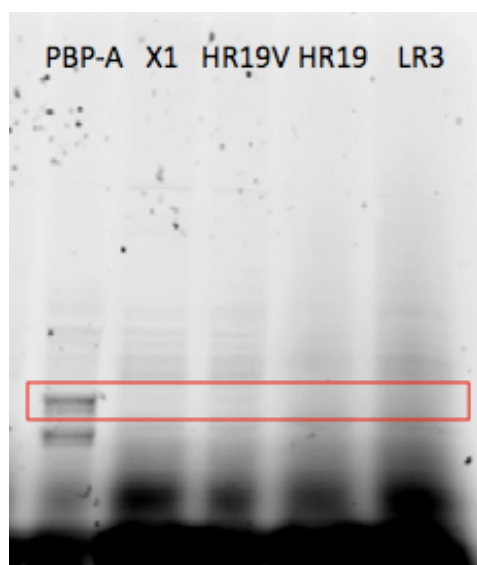


Figure 46: BOCILLIN labeling of the periplasmic extract of PBP-A, X1, HR19V, HR19 and LR3⁺ : the fluorescence detection was performed using the CY2 filter.

The Figure 46 shows that after the fluorescence revelation, only the wtPBP-A was detectable. Obviously, this result suggest that our mutants were not produced at a sufficient level or not labelled by Bocillin. Since these evolved PBPs have shown a selectable phenotype (penicillin resistant), an increase in their deacylation rate is expected. In this situation, if deacylation is faster than acylation, penicillin labelling is not feasible because the acylenzyme does not accumulate.

Bocillin labelling was also tested with two other intermediates that were selected on sublethal concentrations of ampicillin (G6 and G9 variants). For these variants, the deacylation is expected to be lower. As shown in figure 47, bands at the expected size are detected but with a weaker intensity compared to the wild type protein, suggesting that the expression level of the intermediates is reduced or a possible increase in the deacylation rate. The presence of two forms of the protein (with or without the signal peptide) could explain the presence of the two bands at the expected size. The bands

corresponding to the lower molecular weights are probably due to the proteolysis of some other PBPs of *E. coli* also recognized by Bocillin.

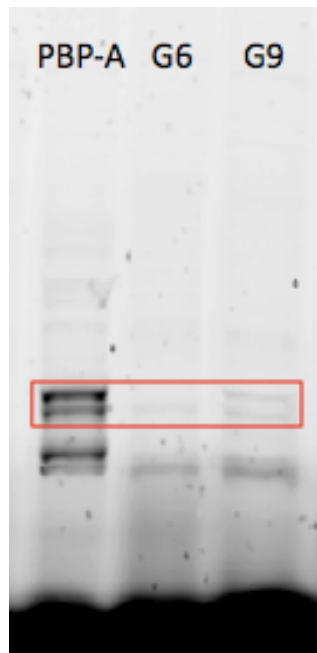


Figure 47: BOCILLIN labeling of the periplasmic extract of PBP-A, G6 and G9: the fluorescence detection was performed using the CY2 filter.

As another attempt to detect the PBP-A mutants, an immuno-detection assay was performed using an antibody directed against the Histidine tag. A purified his-tagged β -lactamase was used as positive control.

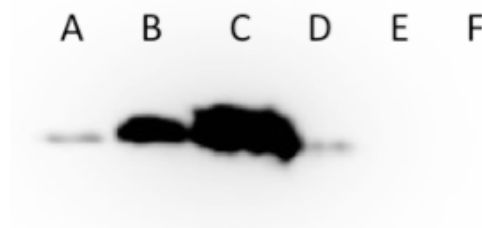


Figure 48: Western Blot using an anti-histag antibody. A, B, and C : his-tagged TEM1 β -lactamase respectively at 100ng, 500ng and 2 μ g. D, E and F : periplasmic extract of PBP, X1 and HR19 respectively. The expected weights for PBP-A variant and TEM1 are around 31 KDa.

The western-blot result showed the antibody is not very sensitive because only a faint band is revealed with 100 ng of β -lactamase. Another faint band is visible in the wt PBP-A lane although it is probably due to the leakage of the neighbouring slot because PBP-A should migrate a little slower than the TEM-1 β -lactamase (see figure 48). Anyway, if the expression levels of the mutants are lower, the absence of signal for them is not surprising.

We then decided to perform an analysis by mass spectrometry. Slices of the gel corresponding to the size of the two mutants were cutted and analysed by MALDI-TOF. Interestingly, this method allowed the detection of some peptides containing specific mutations and belonging to HR19V and X1 (Figure 49), thus confirming the expression of the mutants.

Mass Spectrometry results

PBP-A Matched peptides shown in **Bold Red**

1 PAPEAPTSTL PPERPLTNLQ QQIQQLVSRQ PNLTAGLYFF NLDGASLNV
 51 GGDQVFPAAS TIK**FPILVAF FKAVDEGRVT LQER**LTMRPD LIAPEAGTLQ
 101 YQKPNSQYAA LEVAELMITI SDNTATNMII DRLGGAAELN QQFQEWGLEN
 151 **TVINNPLPDM** KGTNTTSPRD LATLMLK**IGQ GEILSPRSRD RLLDIMRRTV**
 201 TNTLLPAGLG KGATIAHK**TG DIGIVVDAG MVDMPNGQRY VAAMMVKRPY**
 251 **NDPRGSELIR QVSRMVYQAF EKLSPISFN HHHHHH**

X1 Matched peptides shown in **Bold Red**

1 PAPEAPTSTL PPERPLTNLL QLIQQLSRQ PNLTAGLYFF NLDGASLNV
 51 GGDQAFPAAS TIKFPILVTF FKAVDEGRAP LQDRLTLRPV LIAPVAGTLH
 101 YQKPNSQYAA LEVAELMITI SDNTATNMII DRLGGAAELN QQFREWGLEN
 151 TVINNPQPD M RGTNTTSPRV LATLMLK**VGQ GGILSPRSRD RLQDISWRTV**
 201 TNTLLPAGLG KGATIAHK**TG DIGIVVDAG MVEMPY GKRS VAAMMVKRPY**
 251 **NGPRGSELIR QASRMAYQAF ERLSPPTSFN HHHHHH**

HR19V Matched peptides shown in **Bold Red**

1 PAPEAPTSTL PPERPLTNLQ QQIQQLVSRQ PNLTAGLYFF NLDGASLNV
 51 GGDQAFPAAS TIK**FPILVAF FKAVDEGRVT LQER**LTMRPV LIAPVAGTLH
 101 YQKPNSQYAA LEVAELMITI SDNTATNMII DRLGGAAELN QQFREWGLEN
 151 TVINNPQPD M KGTNTTSPRV LATLMLK**IGQ GEILSPRSRD RLQDIMRRTV**
 201 TNTLLPAGLG KGATIAHK**TG DIGIVVDAG MVEMPNGQRY VAAMMVKRPY**
 251 **NDPRGSELIR QVSRMAYQAF ERLSPPTSFN HHHHHH**

Figure 49: Mass spectrometry coverage of the PBP-A, X1 and HR19. The MALDI –TOF was performed on the protein extract from slices cut from an SDS-PAGE gel. In BOLD Red are the identified peptide.

All the previous analyses were performed on soluble fractions. We also wanted to check if the proteins are in the insoluble fractions of the cell. Therefore, the recovered pellets, after consecutive periplasmic and cytoplasmic extractions from 4 hours induced (0.1 mM IPTG) cultures of XL1blue strain transformed with pHDiex-PBP-A and pHDiex-X1 plasmids, were solubilized and tested by westernblot using an anti-histag antibody. Figure 50 shows that both the wt PBP-A and the X1 mutant protein are essentially insoluble when over-expressed. With wt PBP-A, two bands can be distinguished. It probably corresponds to the preprotein with the signal peptide and the processed PBP-A, suggesting that inclusion bodies are formed both in the cytoplasm and in the periplasm after export. For the X1 mutant, the lower band is not present, indicating that it precipitates only in the cytoplasm as a preprotein.

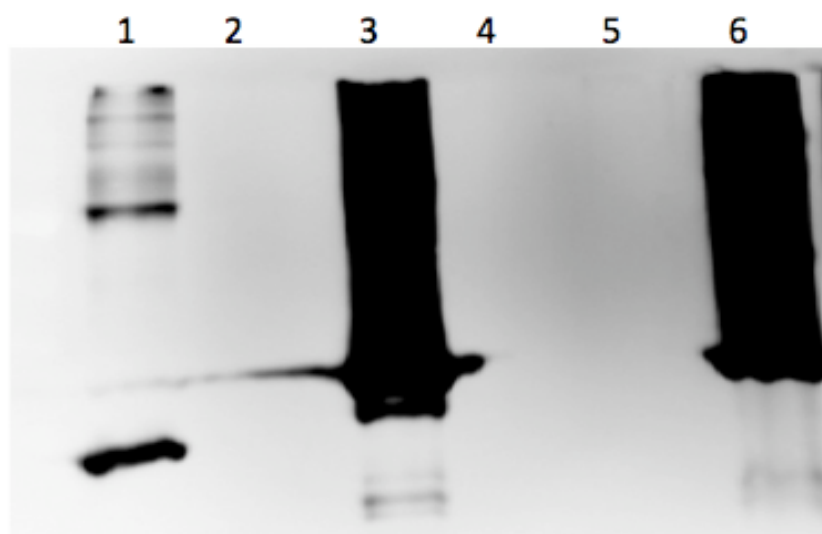


Figure 50: Western-blot using an anti-Histag antibody. 1 is the total soluble extract (mix of periplasmic and cytoplasmic fractions) of an *E. coli* culture expressing a Histagged β -lactamase, 2 is the total soluble extract of the PBP-A ; 3, 5 and 6 : insoluble fractions from an XL1 blue culture transformed with the expression plasmid for wtPBP-A (3) and X1 mutant (6), or untransformed (5). 4 : protein ladder.

An attempt to purify and refold the X1 variant from the insoluble fraction (pellet) was performed. After the solubilization of the pellet in denaturing conditions (in presence of 6M guanidium chloride), the sample was loaded into a Nickel column already equilibrated by the solubilization buffer. Then, the column was washed with guanidinium then without guanidinium to allow refolding on the column. The concentration of the imidazole in the elution buffer was increased (100 mM imidazole step gradient elution) but the column became clogged, indicating that the protein reprecipitate upon elution.

III.3. Purification attempt of the X1 mutant by ion exchange chromatography

Our initial purification attempts indicate that the active protein cannot bind to the nickel column. This pushed us to try classical ion exchange chromatography for the purification of the X1 mutant. Ion exchange also allowed to minimize the problems of contamination by the TEM1 β -lactamase routinely used in our laboratory. This can be a serious problem when working with poorly active and poorly expressed enzymes since traces of beta-lactamase are present in the purification equipments (tubings, columns, pumps...) in the laboratory. Moreover, there are two pairs of histidines at the surface of the TEM-1 so that it binds well to the nickel columns. Nevertheless, the beta-lactamase has a theoretical pI (isoelectric point) of about 5.7 compared to 9.5 for our mutant (estimated via the software Protparam from ExPasy). Working with ions exchangers at pH 7.4 was therefore an interesting alternative.

On the other hand, our results indicate that the induction or the over-expression is generating inclusion bodies. Even, the wild type PBP-A precipitates when it is induced. We also know from the previous results

(chapter 1) that overexpression has a high fitness cost. Therefore, a decrease in the expression level could be a solution for obtaining a small amount of soluble and folded protein. We therefore avoided the induction of the XL1Blue cells harboring the vector encoding this mutant. When the culture reached an optical density of 0.6 the incubation temperature was reduced to 30°C and we kept the same agitation overnight. The periplasmic extract from 10 litres of culture was concentrated up to 20 ml and then dialysed overnight against the equilibration buffer (see methods). The sample was then loaded into a Capto S cation exchanger column.

To check if our mutant is bound to the column we evaluated the activity toward penicillin G (PenG) in the Flow through (20 ml) and in the Wash (50 ml). The hydrolytic activity can be followed by the change in the absorbance at 232 nm since the product of hydrolysis absorbs less light than the PenG substrate ($\text{De}_{232_{\text{S-P}}} = 1030 \text{ M}^{-1}\text{cm}^{-1}$). 200 μl were added to 1.8 ml of PenG (0.5 mM) in Hepes buffer. The Figure 51 shows the expected observation. No activity can be detected in the flow through and the Wash, suggesting that the X1 variant is bound to the cation exchanger.

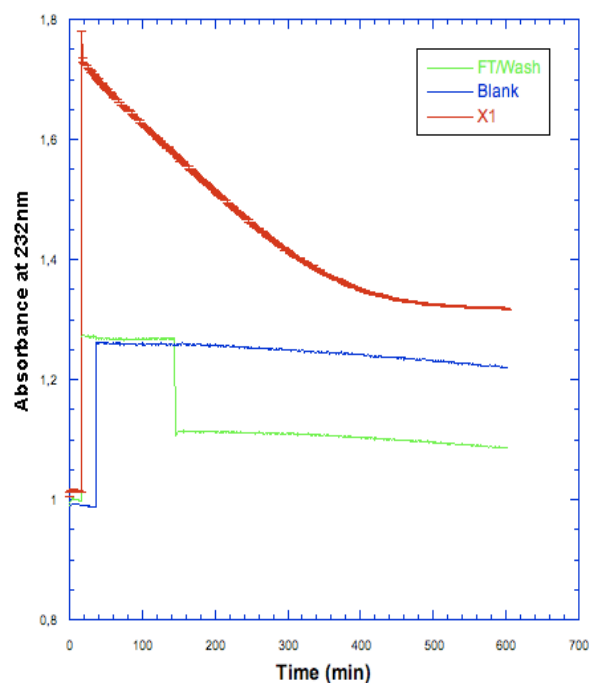


Figure 51: Activity Assessment when loading the periplasmic extract on the Capto S column. The Flow through (FT) and the Wash (W) fractions were mixed and assayed toward penicillin compared to the loaded sample (X1). The decrease of the green curve at 150 min is due to the dilution in the cuvette when 100 μ l of the wash were added to the FT.

Then, the fixed proteins were eluted by a NaCl step gradient (Starting from 100 up to 400 mM and with 100 mM steps) in 1 column volume for each fraction (5 ml). The activities (Figure 52) in all the fractions were measured to identify the presence of the X1 enzyme.

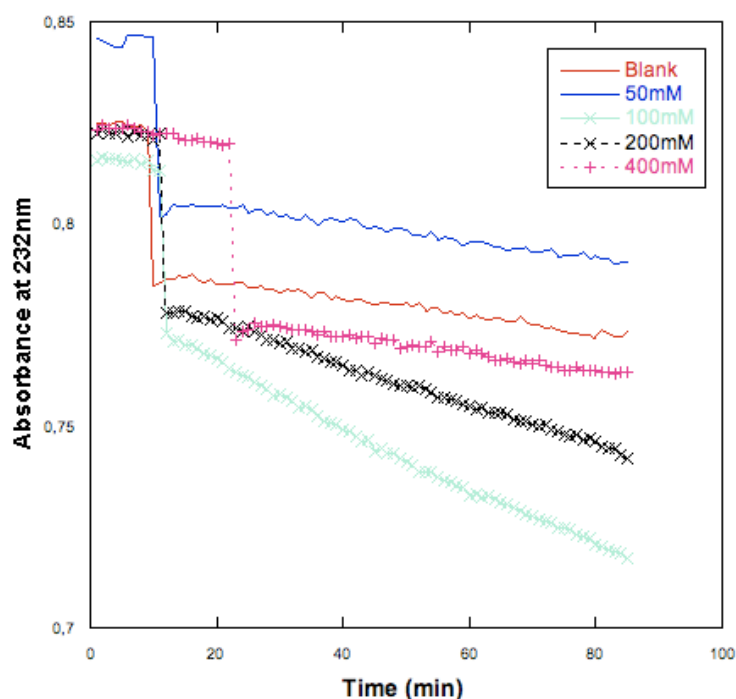


Figure 52: Capto S elution using a 100 mM stage NaCl gradient : the Blank is the PenG in the Hepes Buffer. The fraction 50mM was an additional wash of 1 column volume (5ml). 100 μ l of each fraction was assayed on PenG.

The active fractions (100 mM to 200 mM) were pooled and dialysed again against the equilibration buffer (50 mM Hepes, 10 mM NaCl and pH 7.4) and then loaded into the Capto Q anion exchanger. As expected, the activity was in the flow through. This step allowed to confirm that the activity is not associated with a TEM-1 β -lactamase contamination since this enzyme (PI= 5.58) should bind to the Capto Q at pH 7.4.

After concentration of the sample to 3 ml using a Millipore ultracentrifugation unit with a cut-off of 10 KDa, it was loaded onto the Superdex S200 gel filtration column (310 ml column). This step allowed the separation of two proteins corresponding to two bands in SDS-PAGE gel : the molecular weights were around 60 KDa and 30 KDa for the first and

second protein respectively. The second protein was active on Nitrocefin (rapid detection) and on PenG.

The figure 53 represents the 3 steps purification in SDS-PAGE gel.

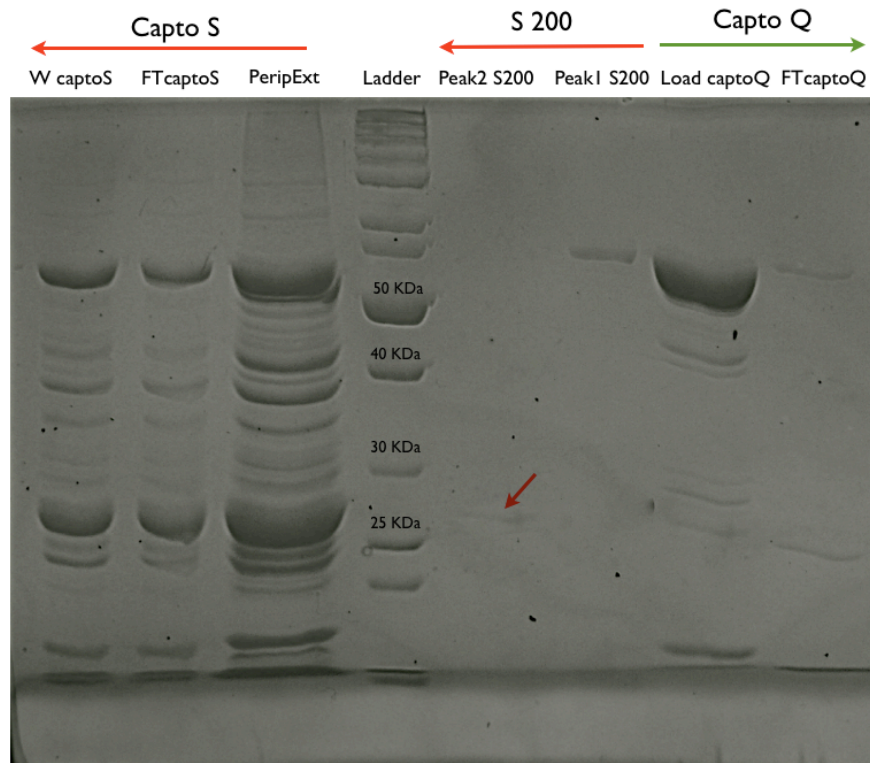


Figure 53: SDS-PAGE showing the different step of the purification of X1 mutant. FT=flow through, W=Wash, PeripExt= periplasmic extract. the purification was performed in 3 steps, CaptoS, CaptoQ and S200 gel filtration.

The band corresponding to the second peak (active fractions) after gel filtration (showed by the red arrow) was cut from the gel and analysed by Mass Spectrometry for identification. The identified peptides covered 40 % of our mutant protein sequence. A liquid sample (from the second peak) was also analysed and confirmed the first results (Figure 54).

Liquid fraction detected by Mass-Spectroscopy

```
1  PAPEAPTSTL PPERPLTNLL QLIQQLGSRQ PNLTAGLYYF NLD SGASLNV
51 GGDQAFPAAS TIKFPILVTF FKAVDEGRAP LQDRLTLRPV LIAPVAGTLH
101 YQKPYSQYAA LEVAELMITI SDNTATNMII DRLGGAAELN QQFREWGLN
151 TVINNPEPDM RGTNTTSFRV LATLMLKVGQ GGILSPRSRD RLQDISWRTV
201 TNTLLPAGLG KGATIAHKTG DIGIVVDAG MVEMPYGKRS VAAMMVKRPY
251 NGPRGSELIR QASRMAYQAF ERLSPPTSFN HHHHHH
```

Figure 54: Mass-spectrometry identification of the peptides (red) corresponding to the X1 mutant.

This classical ion exchange and gel filtration approach allowed us to purify X1 to some extent, but the yield is very poor. From 10 litres of culture we were able to recover only 1 ml of a protein solution with a very low concentration ($< 1 \mu\text{g/ml}$). Since the band is very faint, the purity cannot be estimated with precision.

III.4. Kinetic parameters estimation

The purified fraction (1ml) was sufficient to estimate the K_m of the X1 mutant on Benzylpenicillin. We varied the PenG concentration from 0,1 to 500 μM and the experimental rates were fitted to the Michaelis-Menten equation. Since the enzyme concentration is not known, we were only able to estimate the K_m value (Figure 55).

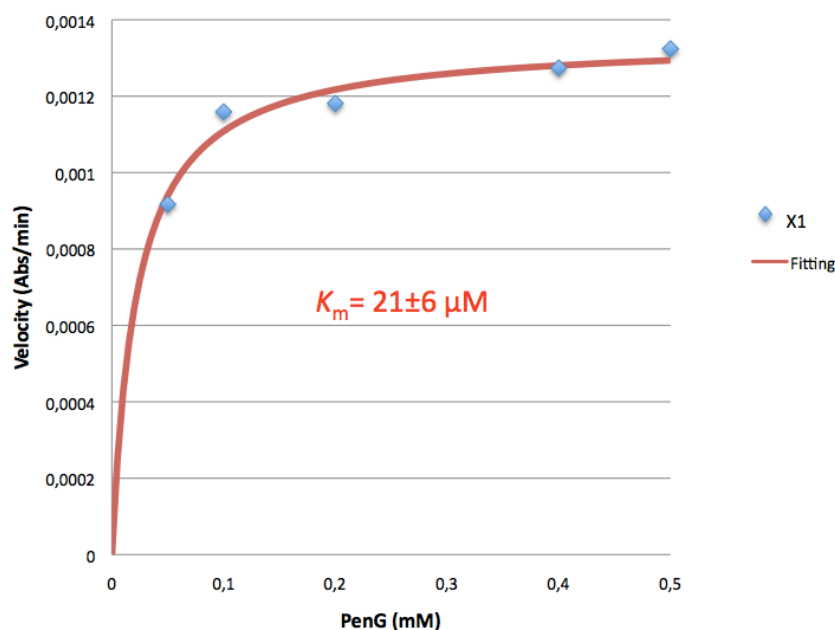


Figure 55: K_m estimation by fitting the experimental rates to the Michaelis-Menten equation.

The value of $21 \pm 6 \mu M$ is much higher than the K_m of wtPBP-A (around 15 nM) and in the same order of magnitude compared to those of PBP-A-L158E mutant and TEM1 β -lactamase known to be around $10 \mu M$ and $35 \mu M$ respectively (116, 165). PBP-A-L158E is a mutant featuring a 90-fold increase in the deacylation rate of the acyl-enzyme.

Then, we tried to estimate the k_{cat} of the mutant toward PenG. Since it has been really difficult to fully purify the protein and to evaluate its amount, we followed another experimental approach. First, the wild type PBP-A was purified and its recognition by a polyclonal antibody directed against the TEM1 β -lactamase was assessed by WesternBlot. Interestingly, while PBP-A and TEM-1 share only 24% of sequence identity, PBP-A could be detected by the anti-TEM-1 antibodies. A faint band was also detected in the periplasmic extract containing the X1 mutant. Second, we compared the intensity of the mutant band to those of the PBP-A at different

concentrations (See Figure 56). We estimated the X1 mutant quantity to be around 25-50 ng in the 20 μ l loaded in the SDS-PAGE gel.

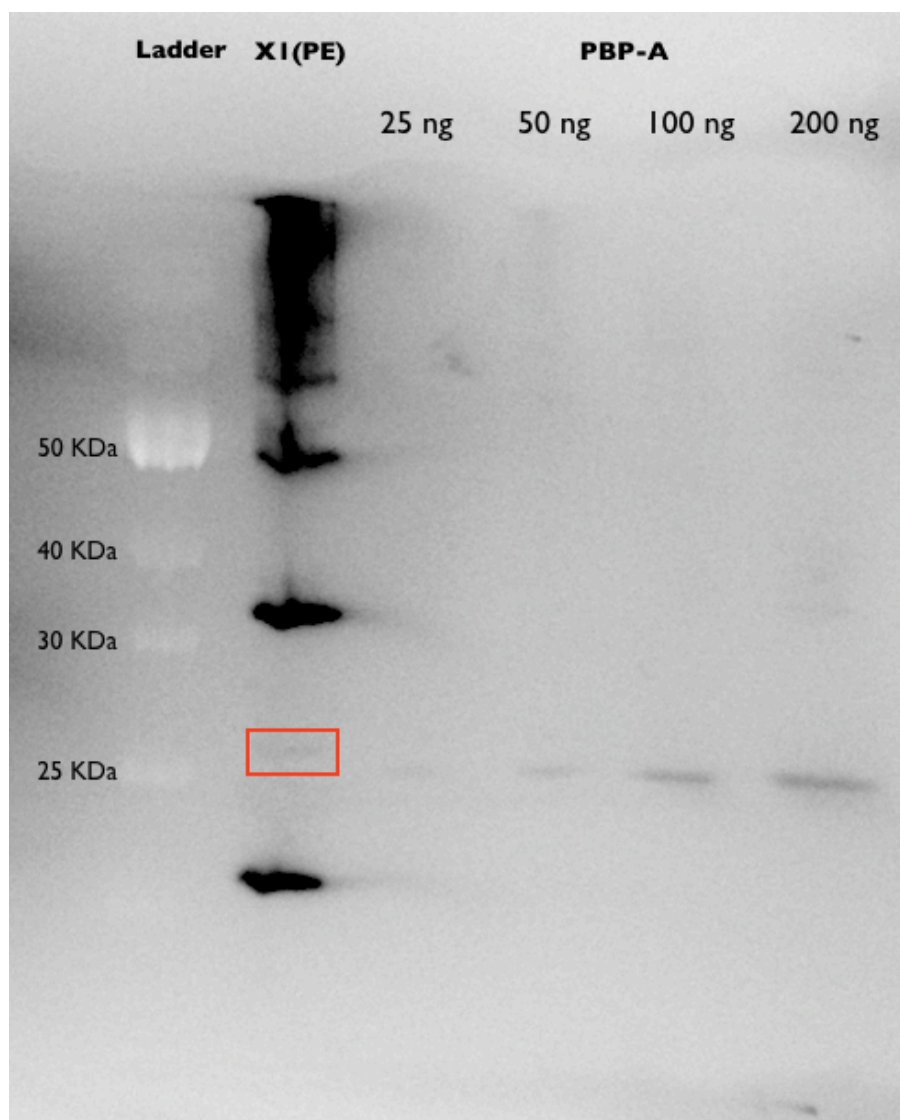


Figure 56: Western-Blot using a polyclonal antibody directed against the PBP-A and X1
Concentrations of PBP-A were determined by the Bradford Kit (Roche), PE= periplasmic
extract. The band in the red box was checked by Mass spectrometry and was confirmed to
be the X1 variant.

Then, 100 μ l from the same periplasmic extract were used to assess the activity toward the PenG (0.5 mM) in a final volume of 2 ml. Hence, the

concentration $[E]_0$ of the X1 mutant is estimated between 2.2 and 4.5 nM. The $V_{\max}$ was estimated by fitting the slope at the initial linear phase of the PenG hydrolysis (Figure 57) corresponding to the maximal velocity. The slope value was then divided by the PenG absorbance constant that gave at the end a value which is approximately equivalent to a $V_{\max}$ of about $21 \times 10^{-9} \text{ Ms}^{-1}$.

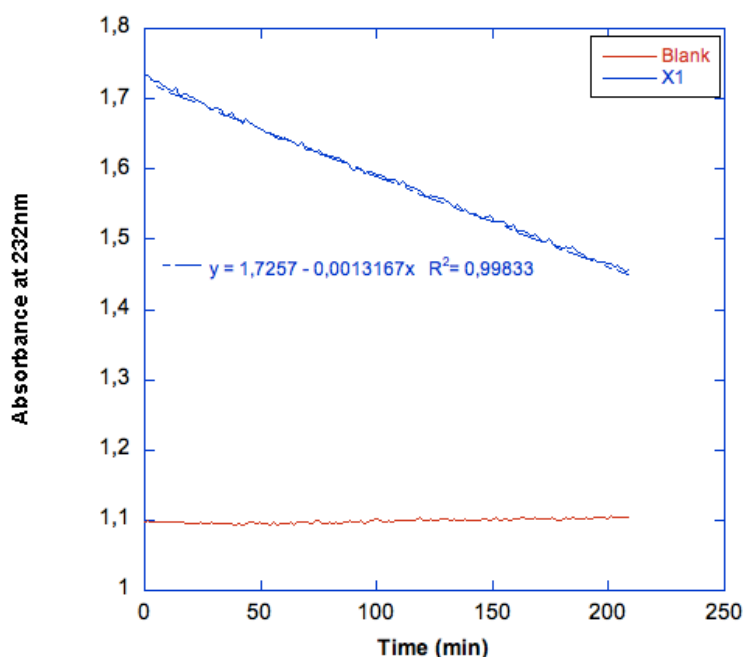


Figure 57: Estimation of X1 $V_{\max}$ by fitting the slope in the linear phase of PenG hydrolysis corresponding to the maximal velocity. The blank is PenG (0,5 mM) in Hepes Buffer.

Using the equation : $V_{\max} = k_{\text{cat}} \times [E]_0$, we estimated the k_{cat} to be around $5\text{-}10 \text{ s}^{-1}$. According to this estimation, the deacylation rate of our best selected mutant is increased by at least 35000-fold. This is a remarkable improvement although this experiment need to be reproduced with the proper controls. If confirmed, this improvement would be much better than the 90-fold factor obtained by Carole Urbach after introducing a glutamate in position 158 of PBP-A, a position corresponding to the essential Glutamate 166 found in class-A β -lactamases.

IV. Conclusion

In summary, we confirmed the production of the selected mutants in spite of their very low expression level and we showed that a classical method of purification combining ions exchanger columns and gel filtration could be used to purify the X1 mutant but we failed to obtain a concentrated solution so the purity cannot be estimated. However, a kinetic evaluation, using Western Blot for estimating the protein concentration, demonstrated that the evolutionary trajectory followed by PBP-A allowed a big jump in the functional space leading to an increase in the deacylation rate by at least 4 orders of magnitude.

The X1 mutant accumulated 31 amino-acid substitutions and showed a good fitness and a phenotypically relevant β -lactamase activity toward ampicillin. However, the enzyme is poorly expressed and most probably highly unstable. Because of this, it is very difficult to characterize the protein *in vitro* and structural investigations are not possible.

V. Materials and methods

Bacterial Strains

XL1BlueMRF': Δ (mcrA)183 Δ (mcrCB-hsdSMR-mrr)173 endA1 supE44
thi-1 recA1 gyrA96 relA1 lac [F' proAB lacIqZ Δ M15 Tn10 (Tetr)]
BL21 (DE3) :F⁻ ompT gal dcm lon hsdS_B(r_B⁻ m_B⁻) λ (DE3 [lacI lacUV5-T7
gene 1 ind1 sam7 nin5])
Top10 : F- mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80lacZ Δ M15 Δ lacX74 nupG
recA1 araD139 Δ (ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ ⁻

Cloning into pBAD vector

The gene encoding the PBP Variant in pHDiex was amplified by a standard PCR using the pBAD-Fwd5'-CATGGCCGCTCCTGAGGCACC and the pBAD-Rev5'-CTAGAGGCGGCGAAAGTT as primers. The amplified DNA was purified (PCR product purification Kit from Roche) and then ligated over night to the NcoI/XbaI digested pBAD vector at 4°C using the T4 DNA ligase (New England Biolabs). The resulting clone was confirmed by sequencing.

SDS-PAGE

Analytical polyacrylamide gel electrophoresis of proteins in the presence of sodium dodecyl sulphate (SDS/PAGE) was performed by the method of Laemmli (1966).

Periplasmic extract preparation

See chapter 2

Western Blotting

The western blotting was carried out with the periplasmic extracts, the pellets and purified proteins. After 10 min heating in the presence of

reducing SDS buffer containing β -mercaptoethanol (710 mM) the samples were loaded and allowed to migrate on 12 % (w/v) polyacrylamide gel at 200 V and 40 mA in electrophoresis buffer (25 mM Tris, 250 mM glycine, 0.1 % SDS pH 8). The separated proteins were then transferred from the gel to a nitrocellulose membrane (Sigma) at 14 V for 50 min in transfer buffer (25 mM Tris, 200 mM Glycine, 20 % methanol pH 8). The membrane is blocked with MPBST (140 mM NaCl, 2,5 mM KCl, 10 mM $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$, 2 mM KH_2PO_4 , 0.05 % (v/v) Tween20 and 5 % (w/v) skimmed milk powder) for 1 hour at room temperature. Then it was incubated with the primary antibody (Anti-Histag or Anti- β -lactamase in the MPBS (PBS with Milk, 30 ml)) for 1 hour at room temperature and washed 3 times with PBST (PBS with Tween, 30 ml) for 15 min each. After that, the membrane was incubated with the secondary antibody (bovine anti rabbit IgG-HRP) in MPBS for 1h at room temperature and washed again 3 times with PBST. The Horse raddish peroxidase (HRP) was then detected by luminescence using the Immuno-Star HRP Chemiluminescent TM kit (BioRad).

Activity measurements

The activities of the periplasmic and cytoplasmic extract of the PBP variants were determined at 25°C in 50 mM Hepes buffer (pH 7.5). The hydrolysis of the benzylpenicillin was monitored at 232nm by measuring the decrease in absorbance as a function of time (PenG: $\Delta\epsilon_{232(\text{P-S})} = -1040 \text{ M}^{-1} \cdot \text{cm}^{-1}$) The Nitrocefin was also used for rapid detection of the active fractions. 20 μl of the sample were added to 100 μl of Nitrocefin (30 μM) and the appearance of the red color after few minutes is an indicator of activity.

The kinetic parameters K_m and $V_{\max}$ were estimated by fitting experimental rates to the Michaelis-Menten equation. The k_{cat} was also estimated by dividing $V_{\max}$ by the estimated protein concentration determined indirectly via a western blot analysis.

PBP fluorescence detection

A mix containing 40 μ l of the periplasmic extracts, 20 μ l of BOCILLIN FL (Invitrogen) at 100 μ M and 140 μ l of Hepes buffer at pH 7,5 was prepared and incubated for 10 min at room temperature. After this step of acylation, all the PBPs should be blocked. 2 μ l of TEM-1 β -lactamase (10 μ M) were added for hydrolyzing the excess of BOCILLIN. Then, 6 μ l of SDS Dye were added to 18 μ l from each mix. The samples were separated on 12 % SDS-PAGE and the Bocillin-labeled proteins were visualized using an Ettan DIGE Imager (GE Healthcare). The machine uses a CY2 filter (excitation at 492 nm and emission at 510 nm)

Cytoplasmic extract preparation

The pellet from the last centrifugation in the periplasmic extract preparation was resuspended in 500 μ l of the lysis buffer (50 mM Tris, 100 mM NaCl, 1 mM EDTA, pH 8). Then, 10 μ l of lysosyme (10 mg/ml) were added and the mix was incubated for 30 min at room temperature and for an additional 45 min at 37°C. After the addition of 20 μ l of DNase (1 mg/ml) and a 15 min incubation at room temperature, the mix was sonicated 5 times for 20 s (separated by 20 s cooling steps on ice) and then centrifuged for 30 min at 10000 rpm and 4°C. The cytoplasmic extract was recovered in the supernatant.

Protein production

After several attempts to over-express the PBP-A variants, the IPTG-induction after reaching the exponential phase was avoided. The culture in LB-medium supplemented with 1% og glucose (XL1 Blue strain harboring the variant plasmid) with the appropriate antibiotics was simply incubated for 24 h at 180 rpm and 37°C. For the X1 mutant the culture volume was scaled up to 10 litres.

IMAC affinity purification

The periplasmic or cytoplasmic PBP variant extracts after an overnight dialysis against the equilibration buffer, were applied to 1 ml Ni^{2+} -nitrilotriacetate (NTA) resin (GE Healthcare) previously equilibrated with 50 mM Hepes, 50 mM NaCl and 10 mM Imidazole buffer (Equilibration Buffer; pH 7.4). After washing the column with five volumes of equilibration buffer the his tagged proteins were eluted with a 50 mM stage imidazole gradient (starting from 50 up to 450 mM) in the same buffer. All the fractions were checked for activity with Benzylpenicillin (PenG) by following the absorbance at 232nm or visually with nitrocefin that turns red upon hydrolysis (see Activity measurements).

Affinity purification in denaturing conditions

This method was used in the case of inclusion bodies (in the pellet) and of the total extract of the X1 mutant culture.

In both cases, the pellet (resuspended directly in 5ml of the equilibration buffer) or the total extract (2 ml mixed with 10 ml of the equilibration buffer) was incubated overnight in the equilibration buffer at pH 7.4 containing 50 mM Hepes, 50 mM NaCl, 10 mM Imidazole and 6 M guanidinium chloride. In the case of the inclusion bodies (pellet), a centrifugation at 10000 rpm for 20min at 4°C was done for removing the remaining insoluble fraction. The solutions were then loaded into the Nickel column (1 ml) beforehand equilibrated with the same buffer. After several Washes with the equilibration buffer (5 ml), an in-column refolding step was performed by washing the column with 50 ml using the native buffer without guanidinium (50 mM Hepes, 50 mM NaCl, 10 mM Imidazole and pH 7.4). Then a 50 mM stage imidazole gradient (Starting from 50 up to 400 mM) was performed for eluting the protein.

Cation exchange chromatography

The X1 mutant's periplasmic extract (20 ml from 10 litre culture) was first dialysed over night against the equilibration buffer (50 mM Hepes, 10 mM NaCl, pH 7.4). Then it was loaded into the 5 ml HiTrap Capto S (GE Healthcare) beforehand equilibrated in the same buffer. The Capto S is a strong cation exchanger that allows the fixation of all the positively charged proteins at pH 7.4 (the case of X1 mutant). After washing with 3 column volumes (3 X 5ml of equilibration buffer with 50 mM NaCl), the fixed proteins were eluted by a 100 mM stage NaCl gradient (Starting from 100 up to 400 mM). The active fractions toward PenG were pooled and then dialysed over night against the equilibration buffer before the next step of purification.

Anion exchange chromatography

The dialysed sample from the last step was loaded into a 5 ml HiTrap Capto Q column (GE Healthcare) already equilibrated in 50 mM Hepes, 10 mM NaCl (pH 7.4). This Capto Q is a strong anion exchanger that allows the fixation of all the negatively charged proteins at this pH. After washing with 3 column volumes (3 X 5ml of equilibration buffer with 50 mM NaCl) the fixed proteins were eluted by a 50 mM stage NaCl gradient (Starting from 100 up to 400 mM). The active fractions toward PenG were only in the Flow through as expected and then pooled and dialysed over night against 50 mM Hepes, 150 mM NaCl (pH 7.4).

Filtration on Superdex 200

The analytical exclusion column (HiLoad 26/60 Superdex™ 200 from Amersham Biosciences) was equilibrated with 3 volumes of 50 mM Hepes, 150 mM NaCl buffer (pH 7.4). The active fractions (X1 mutant) collected after Capto S chromatography were concentrated up to 3 ml using Amicon Ultra centrifugal filter (Millipore) with a cut-off of 10 Kda. The sample was

loaded to the column and the Äkta Prime System (GE Healthcare) was used to pass the proteins through the S200 column with another volume of buffer. The peaks were monitored by detector at 280 nm. All the technical instructions of the supplier (pressure, flow rate,...) were respected. The active fractions were detected by the red coloration when incubated with Nitrocefin.

Conclusions and Perspectives

I. General conclusions

This thesis project aimed to evolve a cyanobacterial DD-peptidase into a β -lactamase. Despite a very similar fold and the presence of most of the active site motifs, all the previous attempts based on rational design have failed to generate an activity that is sufficient to confer penicillin resistance. We then, have decided to mimick the natural evolution in a laboratory time scale by using random mutagenesis to create diversity that undergo selection. The adopted strategy based on the Darwinian process was to allow the emergence of the new activity through several rounds of mutations and selections with a gradually increasing pressure for beta-lactamase activity. Without any need for structural informations, just by generating a relatively high diversity and using an efficient selection or screening method the objective could be reached.

Since the beneficial mutations are probably rare, a key step for a successful directed evolution experiment is to design and generate a large and representative diversity.

The chapter 1 describes the use of our « home made » error prone PCR to build large libraries with different mutation rates (4 and 11 mutations per gene) to target the maximum of the protein sequence space of the PBP-A. The characterization of the libraries after successive cultures under different expression conditions have shown that the increase of the expression level was counter productive. In fact, in high expression condition, the diversity was almost lost and we have shown a dramatic decrease in the proportion of full length variants indicative of a toxicity or fitness cost associated with PBP-A expression. It has been also reported that highly expressed proteins evolve slowly (*144, 167, 168*).

More interestingly, two mutations in the active site region (E96V and L158Q) were selected in the full length mutants. Under repression mode and

without any other selection pressure than growth, these two mutations were preferentially selected. They potentially reduce the toxicity of the protein due to any interference with the peptidoglycan metabolism. Here, it was expected to enrich the culture with truncated mutants (caused by indels) supposed to be non toxic. It seems like there is a selection for properly folded mutant but with remodelled active site, promoted by the burden of the fitness cost generated by the truncated mutants.

The preliminary investigation of the role of these two mutations showed that they are involved in reducing the fitness cost of the variant.

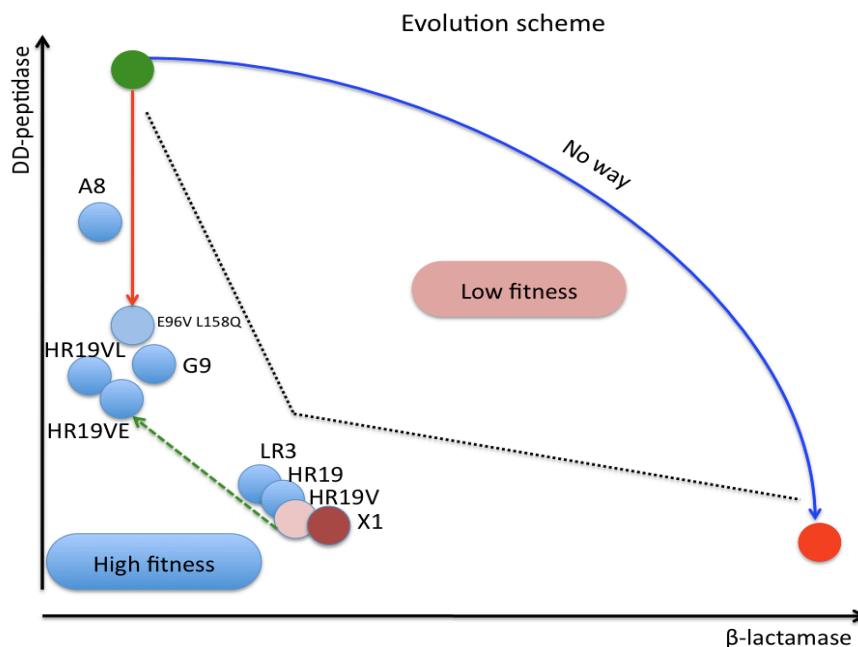
The chapter 2 describes all the evolutionary process and the trajectories followed to reach phenotypically selectable variants. Every drifted Library served as a template to build a new one with respect to the mutation rate adopted for the trajectory. The drifts were performed in sublethal than lethal concentrations of ampicillin. Back transformations were performed to confirm the appearance of new phenotype. Also catalytic mutation of the essential serine have confirmed that PBP-A activity is responsible of the phenotype. This chapter also describes the evolution of the fitness cost during the evolutionary trajectory. Competitions between selected variants have shown the role of fitness cost in shaping the fate of the process.

Our experimental set up mimicks a horizontal gene transfer. The acquisition of the new gene has a fitness cost. The drift served to select for variants with readjusted fitness. In fact, the early selection of variants with the E96V and L158Q mutations without any kind of selection pressure, shaped the destiny of the evolutionary process. An equivalent observation was highlighted in the work of Salverda and his colleagues (169). These two mutations in the active site were shown to be responsible in decreasing the recognition of the native substrate by 300 fold and an increase in the fitness. The following rounds of mutations, drift and selection in gradually increased concentrations of β -lactam antibiotic allowed the emergence of variants with selectable

phenotype. The new activity was accompanied by a loss of fitness. In fact, the selected mutant has a fitness advantage only in presence of ampicillin. Reverting the mutations at the positions 96 and 158 had surprisingly provoked a loss of the phenotype which comforts their indispensability for the new activity emergence.

Our work shows that after the horizontal gene transfer, to discard a non advantageous gene or even a nonfunctional gene could be at some cases not an easy task. Keeping the new gene as a pseudo-gene or neutralized gene has less fitness cost. This non functional gene could serve as a template to emerge a new function when the environmental conditions change.

The scoop of this thesis work is the demonstration for the first time that the neutralization of an enzyme promotes its neofunctionalization, which would be an interesting approach in evolving proteins where the « generalist » intermediate is nonviable or toxic. Here, we also give new insights describing the step that preceed the neofunctionalization of the PBP.



Scheme 3 : Evolutionary process between the PBP-A and the β-lactamase.

In the Chapter 3 we were interested to confirm the biochemical properties of the evolved PBP-A variants along the evolutionary trajectory. Despite the dramatic decrease in the expression level, we were able to detect all the intermediates. The partial purification of the most active variant was a hard task to perform. Since the affinity purification was not successful, a classical method combining ion exchangers and exclusion chromatography was adopted. Many hypotheses could explain the difficulty to purify the selected intermediate. It's possible that by accumulating mutations, the protein became unstable, not properly folded or even denatured. Also, the protein could embrace a molten globule shape and in presence of the substrate it restore the native fold. In this case it will be interesting to try to purify the PBP evolved variants in presence of ampicillin. One other possibility, is that the protein is very diluted and any attempt to concentrate it will provoke an immediate denaturation. Alternatively, the PBP variant could be as inclusion bodies. In this situation, it is possible that a small amount of the protein still active. Nowadays many proteins were successfully purified from inclusion bodies. An attempt with the X1 mutant is eventually possible.

The estimation of the kinetic parameters of the evolved mutant X1 have shown a K_m around 21 μM which is in the same range compared to those of PBP-AL158E mutant and TEM-1 β -lactamase and an increase in the k_{cat} by at least 4 orders of magnitude which explain the emergence of the new phenotype.

II. Perspectives

It is now clear that Directed evolution is an interesting alternative to rational design. In some cases, even with an available 3D structure, it is not easy to predict the target residues to convert the protein activity to a new one. By using directed evolution, we can see beyond the protein structure to engineer new functions.

The promising results obtained in this thesis project opens doors to new possibilities. Actually, the number of hurdles, we were forced to deal with, during the evolutionary process namely the problem of toxicity, contaminations and low expression level, push us to rethink our strategy to evolve the DD-peptidase.

If we focus on emerging the new catalytic property, an *in vitro* alternative using droplets and the microfluidic technology could be the solution. This alternative strategy is being under assessment since a new thesis project has started by Julia Wessel. The idea is to work in a cell-free environment which will reduce the noise caused by non specific interference within the cell. Every droplet contains a single mutated gene, all what is needed for an *in vitro* translation, and a Penicillin-like substrate becoming fluorescent when hydrolysed. The active variants will be selected according to the level of the droplet fluorescence. This new project is free from all the physiological pressures and limitations. In the droplet, even a toxic protein could be assessed and characterized (if the toxicity is not due to the protein translation).

On the other hand, understanding protein evolution and the emergence of new functions by mimicking some evolution scenarios in the lab, is still very challenging. Our scenario of evolution could be improved after slight modifications.

Attempts to evolve the X1 mutant were not enough. This variant could be situated in a local optimum at the fitness landscape, and more probably its high number of mutations is responsible of its dropped tolerance to new ones (robustness problems). The ideal strategy will be to remove the deleterious mutations by a back crossing DNA shuffling against the wild type. The generated variants with only the interesting mutations will serve as template for new rounds of mutagenesis and selection.

In addition, these variants with the responsible mutations for the emergence of the selectable phenotype and the improved fitness could be characterized and then lead us to unveil some new clues to better understand the natural protein evolution.

Directed evolution and rational design have compatible strengths and could reach a larger sequence space when combined. In the previous work of Carole Urbach, a larger omega loop was grafted containing the essential residues of Tem-1 β -lactamase. No improvement was recorded with this method. Considering the omega loop residues of another Class A β -lactamase member (Per1 from *P. aeruginosa*) sharing a better structural alignment with the PBP-A compared to Tem-1 could show different results.

Thanks to the rapid evolution of structural biology and bioinformatic, it is also tempting to try an *in silico* approach in which the main work will be based on finding the potential residues to target, in order to generate a PBP-A variant showing an optimal structural alignment with the two domains of the β -lactamase. Starting a directed evolution experiment using such variant could be successful.

References

1. R. E. Cobb, R. Chao, H. Zhao, Directed evolution : past , present , and future., *AIChE J.* **59**, 1432–1440 (2013).
2. M. T. Reetz, Advances in Catalysis Volume 49, *Adv. Catal.* **49**, 0360–0564 (2006).
3. S. A. LERNER, T. T. WU, E. C. LIN, Evolution of catabolic pathway in bacteria., *Science (80-.).* **146**, 1313–1315 (1964).
4. S. Mills, D.R Peterson, R.L Spiegelman, An extracellular darwinian experiment with a self-duplicating nucleic acid molecule*, *Proc. Natl. Acad. Sci.* **58**, 217–224 (1967).
5. R. Levisohn, S. Spiegelman, Further extracellular Darwinian experiments with replicating RNA molecules: diverse variants isolated under different selective conditions., *Proc. Natl. Acad. Sci. U. S. A.* **63**, 805–811 (1969).
6. D. L. Kacian, D. R. Mills, F. R. Kramer, S. Spiegelman, A replicating RNA molecule suitable for a detailed analysis of extracellular evolution and replication., *Proc. Natl. Acad. Sci. U. S. A.* **69**, 3038–3042 (1972).
7. J. H. Campbell, J. A. Lengyel, J. Langridge, Evolution of a second gene for beta-galactosidase in Escherichia coli., *Proc. Natl. Acad. Sci. U. S. A.* **70**, 1841–1845 (1973).
8. J. O. Deshler, A simple method for randomly mutating cloned DNA fragments by using chemical mutagens and the polymerase chain reaction, *Genet. Anal. Tech. Appl.* **9**, 103–106 (1992).
9. B. W. Glickman, M. Radman, Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction., *Proc. Natl. Acad. Sci. U. S. A.* **77**, 1063–1067 (1980).
10. R. Scheuermann, S. Tam, P. M. Burgers, C. Lu, H. Echols, Identification of the epsilon-subunit of Escherichia coli DNA polymerase III holoenzyme as the dnaQ gene product: a fidelity subunit for DNA replication., *Proc. Natl. Acad. Sci. U. S. A.* **80**, 7085–7089 (1983).
11. K. M. Esvelt, J. C. Carlson, D. R. Liu, A system for the continuous directed evolution of biomolecules., *Nature* **472**, 499–503 (2011).
12. I. J. Fijalkowska, R. M. Schaaper, Mutants in the Exo I motif of Escherichia coli dnaQ: defective proofreading and inviability due to error catastrophe., *Proc. Natl. Acad. Sci. U. S. A.* **93**, 2856–2861 (1996).
13. A. R. Oliphant, A. L. Nussbaum, K. Struhl, Cloning of random-sequence oligodeoxynucleotides., *Gene* **44**, 177–183 (1986).
14. M. T. Reetz, D. Kahakeaw, R. Lohmer, Addressing the numbers problem in directed evolution., *ChemBioChem* **9**, 1797–1804 (2008).
15. R. A. Beckman, A. S. Mildvan, L. A. Loeb, On the fidelity of DNA replication: manganese mutagenesis in vitro., *Biochemistry* **24**, 5810–5817 (1985).
16. J. P. Vartanian, M. Henry, S. Wain-Hobson, Hypermutagenic PCR involving all four transitions and a sizeable proportion of transversions., *Nucleic Acids Res.* **24**, 2627–2631 (1996).
17. R. C. Cadwell, G. F. Joyce, Randomization of genes by PCR mutagenesis., *PCR Methods Appl.* **2**, 28–33 (1992).

18. M. Zaccolo, D. M. Williams, D. M. Brown, E. Gherardi, An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues., *J. Mol. Biol.* **255**, 589–603 (1996).
19. W. P. Stemmer, DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution., *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10747–10751 (1994).
20. W. P. Stemmer, Rapid evolution of a protein in vitro by DNA shuffling., *Nature* **370**, 389–391 (1994).
21. S. W. Santoro, L. Wang, B. Herberich, D. S. King, P. G. Schultz, *An efficient system for the evolution of aminoacyl-tRNA synthetase specificity.* (2002), pp. 1044–1048.
22. C. E. Melançon, P. G. Schultz, One plasmid selection system for the rapid evolution of aminoacyl-tRNA synthetases, *Bioorganic Med. Chem. Lett.* **19**, 3845–3847 (2009).
23. C. C. Liu, P. G. Schultz, Adding new chemistries to the genetic code., *Annu. Rev. Biochem.* **79**, 413–444 (2010).
24. G. P. Smith, Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface., *Science (80-.)*. **228**, 1315–1317 (1985).
25. G. Winter, A. D. Griffiths, R. E. Hawkins, H. R. Hoogenboom, Making antibodies by phage display technology., *Annu. Rev. Immunol.* **12**, 433–455 (1994).
26. R. W. Roberts, J. W. Szostak, RNA-peptide fusions for the in vitro selection of peptides and proteins., *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12297–12302 (1997).
27. G. Yang, S. G. Withers, Ultrahigh-throughput FACS-based screening for directed enzyme evolution., *ChemBioChem* **10**, 2704–2715 (2009).
28. J. J. Agresti *et al.*, Ultrahigh-throughput screening in drop-based microfluidics for directed evolution., *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4004–4009 (2010).
29. E. Mastrobattista *et al.*, High-throughput screening of enzyme libraries: In vitro evolution of a β -galactosidase by fluorescence-activated sorting of double emulsions., *Chem. Biol.* **12**, 1291–1300 (2005).
30. J. A. Francisco, C. F. Earhart, G. Georgiou, Transport and anchoring of beta-lactamase to the external surface of Escherichia coli., *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2713–2717 (1992).
31. P. Fuchs, W. Weichel, S. Dubel, F. Breitling, M. Little, Separation of E. coli expressing functional cell-wall bound antibody fragments by FACS, *Immunotechnology* **2**, 97–102 (1996).
32. M. P. Schreuder, S. Brekelmans, H. van den Ende, F. M. Klis, Targeting of a heterologous protein to the cell wall of Saccharomyces cerevisiae., *Yeast* **9**, 399–409 (1993).
33. E. T. Boder, K. D. Wittrup, Yeast surface display for screening combinatorial polypeptide libraries., *Nat. Biotechnol.* **15**, 553–557 (1997).
34. E. T. Boder, K. D. Wittrup, Yeast surface display for directed evolution of protein expression, affinity, and stability., *Methods Enzymol.* **328**, 430–444 (2000).

35. L. R. Pepper, Y. K. Cho, E. T. Boder, E. V Shusta, A decade of yeast surface display technology: where are we now?, *Comb. Chem. High Throughput Screen.* **11**, 127–134 (2008).
36. D. S. Tawfik, A. D. Griffiths, Man-made cell-like compartments for molecular evolution., *Nat. Biotechnol.* **16**, 652–656 (1998).
37. A. Aharoni, G. Amitai, K. Bernath, S. Magdassi, D. S. Tawfik, High-throughput screening of enzyme libraries: Thiolactonases evolved by fluorescence-activated sorting of single cells in emulsion compartments, *Chem. Biol.* **12**, 1281–1289 (2005).
38. A. D. Griffiths, D. S. Tawfik, Man-made enzymes - From design to in vitro compartmentalisation *Curr. Opin. Biotechnol.* **11**, 338–353 (2000).
39. F. J. Ghadessy, J. L. Ong, P. Holliger, Directed evolution of polymerase function by compartmentalized self-replication., *Proc. Natl. Acad. Sci. U. S. A.* **98**, 4552–4557 (2001).
40. K. Chen, F. H. Arnold, Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide., *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5618–5622 (1993).
41. M. T. Reetz, A. Zonta, K. Schimossek, K.-E. Jaeger, K. Liebeton, Creation of enantioselective biocatalysts for organic chemistry by in vitro evolution., *Angew. Chemie Int. Ed. English* **36**, 2830–2832 (1997).
42. U. T. Bornscheuer, J. Altenbuchner, H. H. Meyer, Directed evolution of an esterase for the stereoselective resolution of a key intermediate in the synthesis of epothilones, *Biotechnol. Bioeng.* **58**, 554–559 (1998).
43. M. Dion *et al.*, Modulation of the regioselectivity of a Bacillus alpha-galactosidase by directed evolution., *Glycoconj. J.* **18**, 215–223 (2001).
44. B. Seelig, J. W. Szostak, Selection and evolution of enzymes from a partially randomized non-catalytic scaffold., *Nature* **448**, 828–831 (2007).
45. F.-A. Chao *et al.*, Structure and dynamics of a primordial catalytic fold generated by in vitro evolution., *Nat. Chem. Biol.* **9**, 81–3 (2013).
46. H. Zhao, Directed evolution of novel protein functions., *Biotechnol. Bioeng.* **98**, 313–317 (2007).
47. H.-S. Park *et al.*, Design and evolution of new catalytic activity with an existing protein scaffold., *Science (80-.)*. **311**, 535–538 (2006).
48. A. Crameri, G. Dawes, E. Rodriguez, S. Silver, W. P. Stemmer, Molecular evolution of an arsenate detoxification pathway by DNA shuffling., *Nat. Biotechnol.* **15**, 436–438 (1997).
49. R. R. Copley, P. Bork, Homology among (beta/alpha)(8) barrels: implications for the evolution of metabolic pathways., *J. Mol. Biol.* **303**, 627–641 (2000).
50. D. M. Z. Schmidt *et al.*, Evolutionary potential of (beta/alpha)8-barrels: functional promiscuity produced by single substitutions in the enolase superfamily., *Biochemistry* **42**, 8387–8393 (2003).
51. J. T. Bridgham, E. A. Ortlund, J. W. Thornton, An epistatic ratchet constrains the direction of glucocorticoid receptor evolution., *Nature* **461**, 515–519 (2009).
52. L. Afriat-Jurnou, C. J. Jackson, D. S. Tawfik, Reconstructing a missing link in the evolution of a recently diverged phosphotriesterase by active-site loop remodeling., *Biochemistry* **51**, 6047–55 (2012).

53. R. Huang *et al.*, Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates *Proc. Natl. Acad. Sci.* **109**, 2966–2971 (2012).
54. I. Matsumura, A. D. Ellington, In vitro evolution of beta-glucuronidase into a beta-galactosidase proceeds through non-specific intermediates., *J. Mol. Biol.* **305**, 331–339 (2001).
55. A. Babbie, N. Tokuriki, F. Hollfelder, What makes an enzyme promiscuous? *Curr. Opin. Chem. Biol.* **14**, 200–207 (2010).
56. O. Khersonsky, D. S. Tawfik, Enzyme promiscuity: a mechanistic and evolutionary perspective., *Annu. Rev. Biochem.* **79**, 471–505 (2010).
57. F. A. Kondrashov, E. V. Koonin, A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications *Trends Genet.* **20**, 287–291 (2004).
58. E. V. Koonin, Y. I. Wolf, Constraints and plasticity in genome and molecular-phenome evolution., *Nat. Rev. Genet.* **11**, 487–498 (2010).
59. G. C. Conant, K. H. Wolfe, Turning a hobby into a job: how duplicated genes find new functions., *Nat. Rev. Genet.* **9**, 938–950 (2008).
60. M. Soskine, D. S. Tawfik, Mutational effects and the evolution of new protein functions., *Nat. Rev. Genet.* **11**, 572–582 (2010).
61. M. Soskine, D. S. Tawfik, Mutational effects and the evolution of new protein functions., *Nat. Rev. Genet.* **11**, 572–582 (2010).
62. M. Kaltenbach, N. Tokuriki, Dynamics and constraints of enzyme evolution., *J. Exp. Zool. Part B Mol. Dev. Evol.* **9999B:1-20** (2014), doi:10.1002/jez.b.22562.
63. E. V. Koonin, *The logic of chance :The nature and origin of biological evolution.*
64. M. F. Schenk, I. G. Szendro, M. L. M. Salverda, J. Krug, J. A. G. M. de Visser, Patterns of Epistasis between beneficial mutations in an antibiotic resistance gene., *Mol. Biol. Evol.* **30**, 1779–1787 (2013).
65. D. J. Kvitek, G. Sherlock, Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape., *PLoS Genet* **7**, e1002056 (2011).
66. J. A. G. M. de Visser, T. F. Cooper, S. F. Elena, The causes of epistasis., *Proc. Biol. Sci.* **278**, 3617–3624 (2011).
67. B. C. Dickinson, A. M. Leconte, B. Allen, K. M. Esvelt, D. R. Liu, Experimental interrogation of the path dependence and stochasticity of protein evolution using phage-assisted continuous evolution, *Proc. Natl. Acad. Sci.* **110**, 9007–9012 (2013).
68. D. M. Weinreich, N. F. Delaney, M. a Depristo, D. L. Hartl, Darwinian evolution can follow only very few mutational paths to fitter proteins., *Science (80-.)*. **312**, 111–114 (2006).
69. S. Bershtein, K. Goldin, D. S. Tawfik, Intense neutral drifts yield robust and evolvable consensus proteins, *J. Mol. Biol.* **379**, 1029–1044 (2008).
70. R. D. Gupta, D. S. Tawfik, Directed enzyme evolution via small and effective neutral drift libraries., *Nat. Methods* **5**, 939–942 (2008).
71. G. Amitai, R. D. Gupta, D. S. Tawfik, Latent evolutionary potentials under the neutral mutational drift of an enzyme., *HFSP J.* **1**, 67–78 (2007).
72. J. S. Taylor, J. Raes, Duplication and divergence: the evolution of new genes and old ideas., *Annu. Rev. Genet.* **38**, 615–643 (2004).

73. J. Zhang, Evolution by gene duplication: an update. *Trends Ecol. Evol.* **18**, 292–298 (2003).
74. A. Wagner, Energy constraints on the evolution of gene expression., *Mol. Biol. Evol.* **22**, 1365–1374 (2005).
75. G. H. Perry *et al.*, Diet and the evolution of human amylase gene copy number variation., *Nat. Genet.* **39**, 1256–1260 (2007).
76. H. Innan, F. Kondrashov, The evolution of gene duplications: classifying and distinguishing between models., *Nat. Rev. Genet.* **11**, 97–108 (2010).
77. A. Force *et al.*, Preservation of duplicate genes by complementary, degenerative mutations., *Genetics* **151**, 1531–1545 (1999).
78. A. Stoltzfus, On the possibility of constructive neutral evolution, *J. Mol. Evol.* **49**, 169–181 (1999).
79. B. G. Spratt, L. D. Bowler, Q. Y. Zhang, J. Zhou, J. M. Smith, Role of interspecies transfer of chromosomal genes in the evolution of penicillin resistance in pathogenic and commensal *Neisseria* species., *J. Mol. Evol.* **34**, 115–125 (1992).
80. E. Hilario, J. P. Gogarten, Horizontal transfer of ATPase genes — the tree of life becomes a net of life *Biosystems* **31**, 111–119 (1993).
81. H. Ochman, J. G. Lawrence, E. A. Groisman, Lateral gene transfer and the nature of bacterial innovation., *Nature* **405**, 299–304 (2000).
82. T. J. Treangen, E. P. C. Rocha, Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes., *PLoS Genet.* **7**, e1001284 (2011).
83. R. E. Lenski, S. C. Simpson, T. T. Nguyen, Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness., *J. Bacteriol.* **176**, 3140–3147 (1994).
84. J. E. Bouma, R. E. Lenski, Evolution of a bacteria/plasmid association., *Nature* **335**, 351–352 (1988).
85. D. I. Andersson, B. R. Levin, The biological cost of antibiotic resistance *Curr. Opin. Microbiol.* **2**, 489–493 (1999).
86. X. Wang, G. Minasov, B. K. Shoichet, Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs, *J. Mol. Biol.* **320**, 85–95 (2002).
87. D. C. Marciano, O. Y. Karkouti, T. Palzkill, A fitness cost associated with the antibiotic resistance enzyme SME-1 beta-lactamase., *Genetics* **176**, 2381–2392 (2007).
88. J. Björkman, D. Hughes, D. I. Andersson, Virulence of antibiotic-resistant *Salmonella typhimurium*., *Proc. Natl. Acad. Sci. U. S. A.* **95**, 3949–3953 (1998).
89. A. Fernández *et al.*, Expression of OXA-type and SFO-1 β -lactamases induces changes in peptidoglycan composition and affects bacterial fitness., *Antimicrob. Agents Chemother.* **56**, 1877–1884 (2012).
90. N. Nanninga, Morphogenesis of *Escherichia coli*., *Microbiol. Mol. Biol. Rev.* **62**, 110–29 (1998).
91. L.-T. Sham *et al.*, MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis, *Science (80-.)*. **345**, 220–222 (2014).
92. M. T. Cabeen, C. Jacobs-wagner, Bacterial cell shape, *Nat. Rev. Microbiol.* **3**, 601–610 (2005).

93. Y. Asano, Y. Kato, A. Yamada, K. Kondo, Structural similarity of D-aminopeptidase to carboxypeptidase DD and beta-lactamases., *Biochemistry* **31**, 2316–2328 (1992).
94. C. Goffin, J. Ghuysen, Biochemistry and comparative genomics of SxxK superfamily acyltransferases offer a clue to the mycobacterial paradox : presence of penicillin-susceptible target proteins versus lack of efficiency of penicillin as therapeutic agent., *Microbiol. Mol. Biol. Rev.* **66**, 702–738 (2002).
95. R. Hakenbeck, J. Coyette, Resistant penicillin-binding proteins., *Cell. Mol. Life Sci.* **54**, 332–340 (1998).
96. R. P. Ambler, The structure of beta-lactamases., *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **289**, 321–331 (1980).
97. J. Lamotte-Brasseur *et al.*, Mechanism of acyl transfer by the class A serine beta-lactamase of *Streptomyces albus* G., *Biochem. J.* **279** (Pt 1, 213–221 (1991).
98. N. C. Strynadka *et al.*, Molecular structure of the acyl-enzyme intermediate in beta-lactam hydrolysis at 1.7 Å resolution., *Nature* **359**, 700–705 (1992).
99. H. Adachi, T. Ohta, H. Matsuzawa, Site-directed mutants, at position 166, of RTEM-1 beta-lactamase that form a stable acyl-enzyme intermediate with penicillin., *J. Biol. Chem.* **266**, 3186–3191 (1991).
100. B. F. Gherman, S. D. Goldberg, V. W. Cornish, R. a Friesner, Mixed quantum mechanical/molecular mechanical (QM/MM) study of the deacylation reaction in a penicillin binding protein (PBP) versus in a class C beta-lactamase., *J. Am. Chem. Soc.* **126**, 7652–7664 (2004).
101. S. D. Goldberg, W. Iannuccilli, T. Nguyen, J. Ju, V. W. Cornish, Identification of residues critical for catalysis in a class C beta-lactamase by combinatorial scanning mutagenesis., *Protein Sci.* **12**, 1633–1645 (2003).
102. M. Hata, Y. Tanaka, Y. Fujii, S. Neva, T. Hoshino, A theoretical study on the substrate deacylation mechanism of class C Beta-lactamase, *J. Phys. Chem. B* **109**, 16153–16160 (2005).
103. Z. Wang, W. Fast, A. M. Valentine, S. J. Benkovic, Metallo-beta-lactamase: structure and mechanism., *Curr. Opin. Chem. Biol.* **3**, 614–622 (1999).
104. U. Heinz, H.-W. Adolph, Metallo-beta-lactamases: two binding sites for one catalytic metal ion?, *Cell. Mol. Life Sci.* **61**, 2827–2839 (2004).
105. D. Golemi, L. Maveyraud, S. Vakulenko, J. P. Samama, S. Mobashery, Critical involvement of a carbamylated lysine in catalytic function of class D beta-lactamases., *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14280–14285 (2001).
106. K. Bush, G. A. Jacoby, A. A. Medeiros, A functional classification scheme for beta-lactamases and its correlation with molecular structure., *Antimicrob. Agents Chemother.* **39**, 1211–1233 (1995).
107. J. A. Kelly *et al.*, On the origin of bacterial resistance to penicillin: comparison of a beta-lactamase and a penicillin target., *Science* **231**, 1429–1431 (1986).
108. R. Kirby, Evolutionary origin of the class A and class C β-lactamases, *J. Mol. Evol.* **34**, 345–350 (1992).
109. C. Urbach, J. Fastrez, P. Soumillon, A new family of cyanobacterial penicillin-binding proteins. A missing link in the evolution of class A beta-lactamases., *J. Biol. Chem.* **283**, 32516–32526 (2008).

110. I. Massova, S. Mobashery, Kinship and diversification of bacterial penicillin-binding proteins and beta-lactamases., *Antimicrob. Agents Chemother.* **42**, 1–17 (1998).
111. R. F. Pratt, Functional evolution of the serine [small beta]-lactamase active site, *J. Chem. Soc. Perkin Trans. 2* , 851–861 (2002).
112. L. Chesnel, A. Zapun, N. Mouz, O. Dideberg, T. Vernet, Increase of the deacylation rate of PBP2x from *Streptococcus pneumoniae* by single point mutations mimicking the class A β -lactamases, *Eur. J. Biochem.* **269**, 1678–1683 (2002).
113. M. Peimbert, L. Segovia, Evolutionary engineering of a β -lactamase activity on a D-Ala D-Ala transpeptidase fold., *Protein Eng.* **16**, 27–35 (2003).
114. A. Matagne, J. Lamotte-Brasseur, J. M. Frère, Catalytic properties of class A beta-lactamases: efficiency and diversity., *Biochem. J.* **330 Pt 2**, 581–598 (1998).
115. C. Bourguignon-Bellefroid *et al.*, Importance of the two tryptophan residues in the *Streptomyces* R61 exocellular DD-peptidase., *Biochem. J.* **282 Pt 2**, 361–367 (1992).
116. C. Urbach *et al.*, Structure of PBP-A from *Thermosynechococcus elongatus*, a penicillin-binding protein closely related to class A beta-lactamases., *J. Mol. Biol.* **386**, 109–120 (2009).
117. U. T. Bornscheuer *et al.*, Engineering the third wave of biocatalysis *Nature* **485**, 185–194 (2012).
118. J. D. Bloom, F. H. Arnold, In the light of directed evolution: Pathways of adaptive protein evolution, *Proc. Natl. Acad. Sci.* **106** , 9995–10000 (2009).
119. K. L. Tee, T. S. Wong, Polishing the craft of genetic diversity creation in directed evolution., *Biotechnol. Adv.* **31**, 1707–1721 (2013).
120. S. Lutz, W. M. Patrick, Novel methods for directed evolution of enzymes: quality, not quantity., *Curr. Opin. Biotechnol.* **15**, 291–297 (2004).
121. T. S. Wong, D. Roccatano, M. Zacharias, U. Schwaneberg, A statistical analysis of random mutagenesis methods used for directed protein evolution., *J. Mol. Biol.* **355**, 858–871 (2006).
122. H. Zhao, L. Giver, Z. Shao, J. A. Affholter, F. H. Arnold, Molecular evolution by staggered extension process (StEP) in vitro recombination., *Nat. Biotechnol.* **16**, 258–261 (1998).
123. T. Eggert *et al.*, Multiplex-PCR-based recombination as a novel high-fidelity method for directed evolution., *Chembiochem* **6**, 1062–1067 (2005).
124. N. Hamamatsu *et al.*, Biased mutation-assembling: an efficient method for rapid directed evolution through simultaneous mutation accumulation., *Protein Eng. Des. Sel.* **18**, 265–271 (2005).
125. N. M. Antikainen, S. F. Martin, Altering protein specificity: techniques and applications., *Bioorg. Med. Chem.* **13**, 2701–2716 (2005).
126. A. V Shivange, J. Marienhagen, H. Mundhada, A. Schenk, U. Schwaneberg, Advances in generating functional diversity for directed protein evolution., *Curr. Opin. Chem. Biol.* **13**, 19–25 (2009).
127. P. Hanson-Manful, W. M. Patrick, Construction and analysis of randomized protein-encoding libraries using error-prone PCR., *Methods Mol. Biol.* **996**, 251–67 (2013).
128. P. C. Cirino, K. M. Mayer, D. Umeno, Generating mutant libraries using error-prone PCR., *Methods Mol. Biol.* **231**, 3–9 (2003).

129. C. Bessler, J. Schmitt, K.-H. Maurer, R. D. Schmid, Directed evolution of a bacterial alpha-amylase: toward enhanced pH-performance and higher specific activity., *Protein Sci.* **12**, 2141–2149 (2003).
130. B. Lingen, J. Grötzinger, D. Kolter, M.-R. Kula, M. Pohl, Improving the carboligase activity of benzoylformate decarboxylase from *Pseudomonas putida* by a combination of directed evolution and site-directed mutagenesis., *Protein Eng.* **15**, 585–593 (2002).
131. D. W. Leung, E. Chen, D. V Goeddel, A Method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction., *Technique* **1**, 11–15 (1989).
132. K. R. Tindall, T. A. Kunkel, Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase., *Biochemistry* **27**, 6008–6013 (1988).
133. S. Shafikhani, R. A. Siegel, E. Ferrari, V. Schellenberger, Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization., *Biotechniques* **23**, 304–310 (1997).
134. M. Imhof, C. Schlotterer, Fitness effects of advantageous mutations in evolving *Escherichia coli* populations., *Proc. Natl. Acad. Sci. U. S. A.* **98**, 1113–1117 (2001).
135. M. M. Desai, D. S. Fisher, Beneficial mutation selection balance and the effect of linkage on positive selection., *Genetics* **176**, 1759–1798 (2007).
136. A. E. Firth, W. M. Patrick, GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries., *Nucleic Acids Res.* **36**, 281–285 (2008).
137. W. M. Patrick, A. E. Firth, J. M. Blackburn, User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries., *Protein Eng.* **16**, 451–457 (2003).
138. B.-L. Wang *et al.*, A universal method for directional cloning of PCR products based on asymmetric PCR., *Biotechnol. Appl. Biochem.* **52**, 41–44 (2009).
139. T. S. Rasila, M. I. Pajunen, H. Savilahti, Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, *Escherichia coli* mutator strain, and hydroxylamine treatment., *Anal. Biochem.* **388**, 71–80 (2009).
140. C. B. Glascock, M. J. Weickert, Using chromosomal lacIQ1 to control expression of genes on high-copy-number plasmids in *Escherichia coli*., *Gene* **223**, 221–231 (1998).
141. D. F. Gruber, V. A. Pieribone, B. Porton, H.-T. Kao, Strict regulation of gene expression from a high-copy plasmid utilizing a dual vector system., *Protein Expr. Purif.* **60**, 53–57 (2008).
142. D. A. Drummond, J. D. Bloom, C. Adami, C. O. Wilke, F. H. Arnold, Why highly expressed proteins evolve slowly., *Proc. Natl. Acad. Sci. U. S. A.* **102**, 14338–14343 (2005).
143. C. Park, X. Chen, J. Yang, J. Zhang, Differential requirements for mRNA folding partially explain why highly expressed proteins evolve slowly., *Proc. Natl. Acad. Sci. U. S. A.* **110**, 678–686 (2013).
144. J. L. Cherry, Highly expressed and slowly evolving proteins share compositional properties with thermophilic proteins., *Mol. Biol. Evol.* **27**, 735–741 (2010).

145. P. J. O'Brien, D. Herschlag, Catalytic promiscuity and the evolution of new enzymatic activities, *Chem. Biol.* **6**, R91–R105 (1999).
146. N. Tokuriki, D. S. Tawfik, Chaperonin overexpression promotes genetic variation and enzyme evolution., *Nature* **459**, 668–673 (2009).
147. K. T. Wyganowski, M. Kaltenbach, N. Tokuriki, GroEL/ES Buffering and Compensatory Mutations Promote Protein Evolution by Stabilizing Folding Intermediates, *J. Mol. Biol.* **425**, 3403–3414 (2013).
148. L. Rockah-Shmuel, D. S. Tawfik, Evolutionary transitions to new DNA methyltransferases through target site expansion and shrinkage, *Nucleic Acids Res.* **40**, 11627–11637 (2012).
149. P. Macheboeuf, C. Contreras-Martel, V. Job, O. Dideberg, A. Dessen, Penicillin binding proteins: key players in bacterial cell cycle and drug resistance processes., *FEMS Microbiol. Rev.* **30**, 673–691 (2006).
150. E. Sauvage, F. Kerff, M. Terrak, J. A. Ayala, P. Charlier, The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis., *FEMS Microbiol. Rev.* **32**, 234–258 (2008).
151. J. van Heijenoort, Peptidoglycan hydrolases of Escherichia coli., *Microbiol. Mol. Biol. Rev.* **75**, 636–663 (2011).
152. A. Fernández *et al.*, Expression of OXA-type and SFO-1 β -lactamases induces changes in peptidoglycan composition and affects bacterial fitness., *Antimicrob. Agents Chemother.* **56**, 1877–84 (2012).
153. C. Urbach, J. Fastrez, P. Soumillion, A new family of cyanobacterial penicillin-binding proteins: A missing link in the evolution of class A β -lactamases., *J. Biol. Chem.* **283**, 32516–32526 (2008).
154. R. E. Harkness, D. Mirelman, E. E. Ishiguro, Regulation of D-alanine carboxypeptidase and peptidoglycan cross-linkage in amino acid-deprived Escherichia coli., *J. Bacteriol.* **145**, 845–849 (1981).
155. W. Vollmer, B. Joris, P. Charlier, S. Foster, Bacterial peptidoglycan (murein) hydrolases., *FEMS Microbiol. Rev.* **32**, 259–286 (2008).
156. M. Matsushashi, M. Wachi, F. Ishino, Machinery for cell growth and division: Penicillin-binding proteins and other proteins, *Res. Microbiol.* **141**, 89–103 (1990).
157. D. L. Popham, K. D. Young, Role of penicillin-binding proteins in bacterial cell morphogenesis *Curr. Opin. Microbiol.* **6**, 594–599 (2003).
158. H. Amanuma, J. L. Strominger, Purification and properties of penicillin-binding proteins 5 and 6 from Escherichia coli membranes, *J. Biol. Chem.* **255**, 11173–11180 (1980).
159. B. G. Spratt, Properties of the penicillin-binding proteins of Escherichia coli K12., *Eur. J. Biochem.* **72**, 341–352 (1977).
160. S. R. De León, K. Daniels, A. J. Clarke, Production and purification of the penicillin-binding protein 3 from Pseudomonas aeruginosa, *Protein Expr. Purif.* **73**, 177–183 (2010).
161. J. Basu, R. Chattopadhyay, M. Kundu, P. Chakrabarti, Purification and partial characterization of a penicillin-binding protein from Mycobacterium smegmatis., *J. Bacteriol.* **174**, 4829–4832 (1992).
162. J. Coyette, J. M. Ghuyssen, R. Fontana, Solubilization and isolation of the membrane-bound DD-carboxypeptidase of Streptococcus faecalis ATCC9790. Properties of the purified enzyme., *Eur. J. Biochem.* **88**, 297–305 (1978).

163. A. Gutierrez *et al.*, β -Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity., *Nat. Commun.* **4**, 1610 (2013).
164. G. Zhao, T. I. Meier, S. D. Kahl, K. R. Gee, L. C. Blaszcak, BOCILLIN FL, a sensitive and commercially available reagent for detection of penicillin-binding proteins., *Antimicrob. Agents Chemother.* **43**, 1124–1128 (1999).
165. D. Sirot *et al.*, A complex mutant of TEM-1 beta-lactamase with mutations encountered in both IRT-4 and extended-spectrum TEM-15, produced by an *Escherichia coli* clinical isolate., *Antimicrob. Agents Chemother.* **41**, 1322–1325 (1997).
166. U. K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4., *Nature* **227**, 680–685 (1970).
167. D. A. Drummond, J. D. Bloom, C. Adami, C. O. Wilke, F. H. Arnold, Why highly expressed proteins evolve slowly., *Proc. Natl. Acad. Sci. U. S. A.* **102**, 14338–14343 (2005).
168. A. W. R. Serohijos, Z. Rimas, E. I. Shakhnovich, Protein Biophysics Explains Why Highly Abundant Proteins Evolve Slowly *Cell Rep.* **2**, 249–256 (2012).
169. M. L. M. Salverda *et al.*, Initial mutations direct alternative pathways of protein evolution, *PLoS Genet.* **7** (2011), doi:10.1371/journal.pgen.1001321.

Appendix

Side project contribution

During my PhD, I had the opportunity to be involved in a project (a collaboration between two laboratories in our Institute) dealing with Phage display and Atomic Force Microscopy (AFM).

The main work of this collaboration was published in the following two articles.

- 1- Controlled manipulation of bacteriophages using single-virus force spectroscopy. (*ACS Nano*)

- 2- Multiparametric atomic force microscopy Imaging of single bacteriophages extruding from living bacteria.
(*Nature communication*)

In the first paper, my task was to produce and purify the bacteriophages. Since these phages were displaying a His-tagged protein at their surface, my attempt to prepare high pure phages for the AFM experiments by passing the phage solution through a Nickel column, then to elute the fixed phages using an imidazol gradient, was successful.

In the second paper, my contribution was to prepare and produce strains harboring combined constructions of phagemids and phages.

Controlled Manipulation of Bacteriophages Using Single-Virus Force Spectroscopy

David Alsteens,^{*,§} Emanuele Pesavento,^{*,§} Gilles Cheuvart,[†] Vincent Dupres,[†] Heykel Trabelsi,[‡] Patrice Soumillion,^{*,*†} and Yves F. Dufrène^{*,*†}

[†]Unité de Chimie des Interfaces, Université Catholique de Louvain, Croix du Sud 2/18, B-1348 Louvain-la-Neuve, Belgium, and [‡]Laboratoire d'Ingénierie des Protéines et des Peptides, Institut des Sciences de la Vie, Université Catholique de Louvain, Croix du sud 4/5, B-1348 Louvain-la-Neuve, Belgium. [§]D.A. and E.P. contributed equally to this work.

Filamentous bacteriophages, which include f1, fd, and M13 species, are able to infect a variety of Gram-negative bacteria like *Escherichia coli*.^{1–3} These phages share a common morphology and a circular, single-stranded DNA genome. In contrast to lytic bacteriophages, filamentous phages are assembled in the cytoplasmic membrane and secreted from infected bacteria without cell lysis. Infection begins by the attachment of proteins pIII to the F pilus of a male *E. coli*. The circular single-stranded DNA enters the bacteria where it is converted by the host DNA replication machinery into a double-stranded plasmid replicative form. By rolling circle replication, the replicative form makes single-stranded DNA, and the templates for expression of the coat proteins are formed. These proteins are inserted into the cytoplasmic membrane, waiting for phage extrusion. Finally, morphogenesis and extrusion of phage descendants occur simultaneously by packaging of the single-stranded DNA into coat proteins while the filamentous particle crosses the bacterial wall.^{1,2}

Phage display technology, first introduced by G.P. Smith³ in 1985, permits the presentation of large peptide and protein libraries on the surface of filamentous phages, which affords the *in vitro* selection of peptides and proteins (including antibodies) with high affinity and specificity to almost any target by affinity capture.^{4,5} Phage display involves the fusion of foreign DNA sequences to the phage genome such that the resulting foreign proteins are expressed in fusion with one of the coat proteins. Although all five coat proteins have been used to display proteins or peptides, gene

ABSTRACT A method is described for the site-directed manipulation of single filamentous bacteriophages, by using phage display technology and atomic force microscopy. f1 filamentous bacteriophages were genetically engineered to display His-tags on their pIX tail. Following adsorption on nitrotriacetate-terminated surfaces, force spectroscopy with tips bearing monoclonal anti-pIII antibodies was used to pull on individual phages via their pIII head. Analysis of the force-extension profiles revealed that upon pulling, the phages are progressively straightened into an extended orientation until rupture of the anti-pIII/pIII complex. The single-virus manipulation technique presented here provides new opportunities for understanding the forces driving cell–virus and material–virus interactions, and for characterizing the binding properties of polypeptide sequences or proteins selected by the phage display technology.

KEYWORDS: phage display · bacteriophages · single virus manipulation · atomic force microscopy · force spectroscopy

VII protein (pVII) and gene III protein (pIII) are by far the most commonly used. Notably, filamentous bacteriophages engineered by phage display also offer exciting opportunities in nanotechnology, for example, for the directed synthesis of magnetic and semiconducting nanowires^{6,7} and for the assembly of functional hybrid nanostructures.⁸

While classical biophysical and molecular biology techniques probe the averaged properties of viruses, single molecule assays provide a means to observe, analyze, and manipulate individual virus particles, thereby revealing events and properties that would otherwise be hidden.^{9–12} In this context, recent progress in atomic force microscopy (AFM) methods^{13,14} has enabled researchers to image a variety of viruses in buffer solution,^{15–18} and to measure the forces required to detach them from cell membranes.¹² Yet, the combined use of genetic engineering and AFM for the site-directed manipulation of filamentous phages has never been reported.

*Address correspondence to yves.dufrene@uclouvain.be, patrice.soumillion@uclouvain.be.

Received for review July 10, 2009 and accepted September 10, 2009.

Published online September 21, 2009.
10.1021/nn900778t CCC: \$40.75

© 2009 American Chemical Society

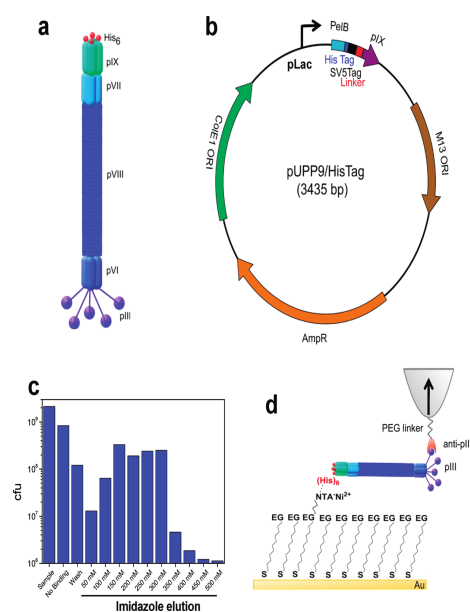


Figure 1. Principle of site-directed single-virus manipulation. (a) Ff filamentous bacteriophages are genetically engineered to display His-tags on their pIX tail. (b) To this end, *E. coli* bacteria are transformed by a phagemid (pUPP9/HisTag) which code for a His₆-pIX protein and infected by a helper phage. (c) The solution of His-tagged phage particles is purified by immobilized ion metal affinity chromatography. In the various input and output fractions, phage particles are detected by their infectivity and titrated as colony forming units (cfu). Approximately 50% of the particles are retained on the column and can be eluted by competition with imidazole. Purified phages are collected by pooling the eluted fractions ranging from 150 to 300 mM imidazole. (d) Following extensive purification, the heterobifunctional filaments are anchored via their His-pIX tail on gold surfaces modified with NTA/EG groups and picked through their pIII head using an AFM tip modified with anti-pIII antibodies.

RESULTS AND DISCUSSION

To address this challenge, we genetically engineered filamentous bacteriophages to display His₆ tags at their pIX, N terminal tail (Figure 1a). We achieved this by using *E. coli* bacteria infected by a R408 helper phage, a f1-type phage,¹⁹ and transformed with a phagemid which code for a His₆-pIX protein (Figure 1b). The obtained tagged filamentous particles are a mixture of R408 helper phage and encapsidated ssDNA of pUPP9/HisTag phagemids. They were precipitated together by polyethylene glycol (PEG) and further purified on a Ni²⁺-NTA column with a ~50% yield (Figure 1c). The successful purification by immobilized metal affinity chromatography confirms the efficient His₆ tag display on the phage particles and also affords a high purity that is required for subsequent AFM analysis. Phages were then attached *via* their His₆ tags on gold surfaces modified with nitrilotriacetate (NTA) groups and picked through their pIII head using an AFM tip modified with anti-pIII antibodies (Figure 1d).

The quality of the phage production and purification procedures was first assessed by AFM imaging (Figure 2). Figure 2 panels a and b show AFM images of mica surfaces on which engineered phages were adsorbed. The surfaces were devoid of large aggregates or debris and clearly showed individual, slightly curved filaments of 7 ± 2 nm diameter (Figure 2c) and 400–1000 nm length. Overall, these values fit with the theoretical dimensions of the engineered phages. Since the phage length is proportional to the size of the encapsidated DNA,⁴ we expect their theoretical contour length, L_c , to be either 500 or 930 nm, depending whether they result from the insertion of the ssDNA from the phagemid or from the helper phage. Last, as we shall see below, it is convenient for further interpretation of the force measurements to take into account the curvature of the filaments and to estimate, from the images, a characteristic distance as the crow flies corresponding to the head-to-tail length of the immobilized phages (L_{ht}). As can be seen in Figure 2d, the distribution of L_{ht} values showed two maxima centered at 480 and 825 nm, corresponding most probably to the phagemid and phage helper particles, respectively. We note that samples that were only purified by PEG precipitation (thus without passage on Ni²⁺-NTA column) were of very poor quality; that is, numerous aggregates were observed, presumably resulting from the co-precipitation of bacterial debris and phage particles. These data demonstrate that thor-

ough purification (PEG precipitation + Ni²⁺-NTA column) of the phages is essential for observing individual viruses and for guaranteeing further reliable single-virus measurements.

Force spectroscopy was then used to pull on the pIII head of individual phages anchored on NTA/EG surfaces (Figure 1d). To this end, AFM tips were functionalized with monoclonal anti-pIII antibodies, using a flexible polyethylene glycol spacer which permits the antibodies to be firmly attached and to be freely oriented. Shown in Figure 3 are the force data obtained between anti-pIII tips and phage surfaces. In 35% of the cases ($n = 360$), the curves showed either single or multiple binding peaks, together with rupture lengths ranging from 50 nm to several hundreds nm. The distribution of the last adhesion forces (F_{adh}), showed a well-defined maximum at 44 ± 14 pN ($n = 360$) that we attribute to specific antibody–antigen interactions for the following rea-

sons: (i) this force value is much smaller than the 150 pN force associated with single His-Ni²⁺-NTA bonds;²⁰ (ii) a marked reduction of adhesion frequency was observed when force curves were recorded in 200 $\mu\text{g/mL}$ solutions of anti-pIII (Figure 3b, inset). In view of the tip surface chemistry that we used (PEG linker) and of the fact that our measured forces are in the range of values reported for single antigen–antibody interactions using similar recording conditions,^{21–25} we conclude that the 44 pN force is likely to reflect the rupture of single pIII/anti-pIII complexes. Adhesion forces in the 50–100 pN range were frequently observed and attributed to the simultaneous detection of two pIII proteins, in agreement with the notion that each phage displays multiple copies of pIII proteins.

Consistent with the above observations, we found that the mean adhesion force ($\langle F \rangle$) increased linearly with the logarithm of the loading rate (r), as expected for specific interactions (Figure 3c).^{23–25} From this relationship, the length scale of the energy barrier was assessed from the slope of the F versus $\ln(r)$ plot, $x_\beta = 0.77$ nm, while extrapolation to zero forces yielded the kinetic off-rate constant of dissociation at zero force, $k_{\text{off}} = r_{F=0} x_\beta / k_B T = 0.74 \text{ s}^{-1}$. Both x_β and k_{off} values are in the range of values reported in previous antigen–antibody studies ($x_\beta = 0.4\text{--}1.0$ nm and $k_{\text{off}} = 10^{-3}\text{--}10^2 \text{ s}^{-1}$).^{22,26,27}

We also note that the 35% binding probability is larger than what we expect considering the surface coverage of the bacteriophages observed in air (Figure 2a). There are several possible reasons for this apparent discrepancy. First, during force measurements some viruses may protrude into the solution with a certain degree of mobility since they are only attached via their pIX tail. We therefore speculate that an anti-pIII tip may detect the same pIII head multiple times. Second, due to dewetting forces associated with drying, the surface density of the bacteriophages may be smaller in air than in liquid. Accordingly, the above data show that AFM with anti-pIII tips can detect and manipulate single bacteriophages via their pIII head up to a maximum force of 150 pN (strength of the His-Ni²⁺-NTA bond), therefore validating the proof-of-concept of “site-directed single-virus force spectroscopy”.

Another pertinent issue is to understand the molecular origin of the force peak profiles (shape, rupture length). Notably, we found that none of the force peaks could be described by the wormlike-chain (WLC) model, which is in contrast with the behavior of modular proteins. Force-extension curves of such proteins generally show nonlinear force peaks that can be fitted

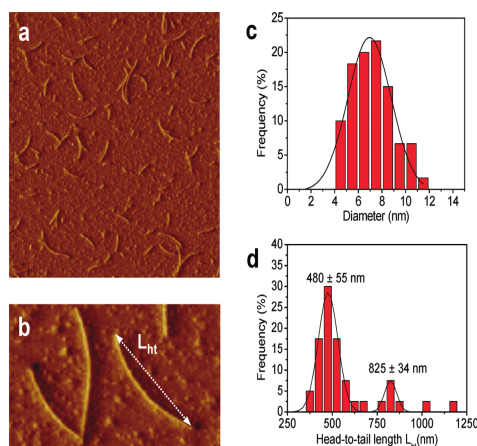


Figure 2. AFM imaging of single viruses. (a,b) Deflection images in air of genetically engineered phages adsorbed on mica (image size: $10 \mu\text{m} \times 10 \mu\text{m}$ (a) and $3 \mu\text{m} \times 1.6 \mu\text{m}$ (b)). Statistical analysis of (c) the diameter and (d) head-to-tail length (L_{ht} , see arrow in Figure 2b) of single filaments (measured on 60 filaments from five images).

by the WLC model, each peak reflecting the force-induced unfolding of secondary structures (α -helices, β -sheets).^{28–32} Thus, rather than reflecting secondary structure deformations, we suggest that the intermediate force jumps correspond to the progressive desorption of the filaments from the surface and/or from neighboring phages (see the image in Figure 2a showing that some phages associate with each other), while the last jump represents the rupture of the pIII–anti-pIII complex.

How about the origin of the observed rupture lengths? Figure 3d shows that the average rupture length was $L_r = 128 \pm 52$ nm, which is much smaller than the phage length, $L_c = 500$ nm. However, the L_r values can be easily interpreted considering the following model (Figure 3e): (i) initially, the engineered phage is a rigid rod¹⁰ that is slightly curved; (ii) upon pulling on its pIII head, it progressively loses its curvature and become more extended since the pIX tail remains attached on the surface. As there is no lateral motion of the tip or the surface, rupture should take place when the phage is in straight oblique position between the His-Ni-NTA and pIII/anti-pIII anchor points. Considering our mean experimental values $L_{\text{ht}} = 480$ nm and the theoretical contour length $L_c = 500$ nm, we find that the expected rupture length should be 140 nm, which matches remarkably well the experimental data, $L_r \approx 130$ nm. The above interpretation is further supported by a controlled experiment in which the phages were attached randomly and via multiple points on surfaces using PEG linkers. In these conditions, the average rupture length dropped to $L_r = 18 \pm 10$ nm (Figure 3d; inset), a value which corresponds to the extension of the

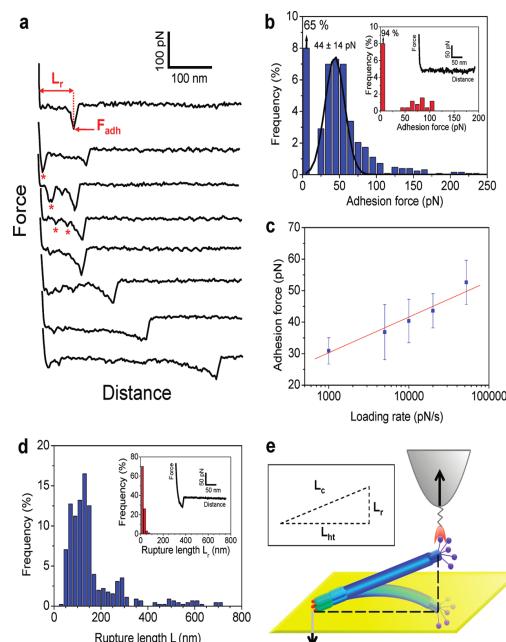


Figure 3. Site-directed force spectroscopy of single viruses. (a) Representative force curves recorded in buffer solution between genetically engineered phages anchored on a NTA/EG surface and AFM tips functionalized with anti-pIII. The last force peak represents the adhesion force of the anti-pIII/pIII pair (see “ F_{adh} ” label), while the occasional intermediate force jumps are attributed to phage-surface or phage-phage desorption events (see asterisk (*) labels). Although most rupture lengths (L_r) were around 100 nm, some were up to several hundreds nm (lower traces). Data obtained using an approach speed of 1000 nm s^{-1} and an interaction time of 500 ms. (b) Histogram showing the distribution of the last adhesion forces (F_{adh}) generated from 1024 curves. The curves were obtained by recording, on $10 \mu\text{m} \times 10 \mu\text{m}$ areas, four force-volume data sets consisting each of 16×16 force curves. Three different tips and three different surfaces prepared independently were used. Shown in the inset is a blocking experiment performed in a $200 \mu\text{g/mL}$ solution of anti-pIII. The histogram was obtained by recording two 16×16 force-volume data sets, using two tips and two surfaces prepared independently. (c) Dependence of the adhesion force on the loading rate applied during retraction. The interaction time (500 ms) and the approach speed (1000 nm s^{-1}) were kept constant. (d) Histogram showing the distribution of rupture length (L_r) from 360 different curves. Shown in the inset is the distribution of L_r values observed when the phages were attached via multiple points using PEG linkers. (e) Schematics of the single-virus pulling experiment: initially, the phages are slightly curved and lying flat on the surface; upon pIII detection and pulling, they are straightened into an extended orientation.

PEG linkers on the tip and on the surface, and to the stretching of the pIII/anti-pIII complex. Together, these findings show that single filaments were manipulated and extended via their two terminal regions, without deformations of the constituting proteins (unfolding, depolymerization).

CONCLUSIONS

In conclusion, our experiments demonstrate that the combined use of phage display technology and AFM is a powerful platform for the controlled manipulation of individual filamentous bacteriophages. Pulling on the pIII head of bacteriophages anchored by their pIX tail using moderate forces, leads to their progressive desorption from the surface and to their straightening into an extended orientation. We an-

ticipate that the novel method presented here, that is, site-directed single-virus force spectroscopy, will find broad applications in life science and nanotechnology. In biology, the site-directed manipulation of bacteriophages with AFM tips should allow the exploration of the forces that drive the various stages of the bacterial infection process (attachment of pIII proteins to F pili and assembly and secretion of the phages from the bacterial surface). In bio- and nanotechnology, the method should be very useful to characterize the biophysical properties of engineered phages particles and phage displayed peptides or proteins that bind to specific bioligands or to organics/inorganics used for the design of self-assembled hybrid materials.⁸

METHODS

Chemicals and Biologicals. Polyethylene glycol 6000 was purchased from AppliChem (Germany) and other chemicals from Sigma Aldrich (Germany). His Trap FF columns were from GE Healthcare Life Science. *E. coli* strain TG1 is from Invitrogen.

Construction of the Phage Expressing HisTag on pII. The HisTag coding sequence was obtained by annealing the oligos His/F (5'-GCGCGCATGCCGACACCATCACCATGCTAGCGGC-3') and His/R (5'-GCCGTAGCATGTGATGGTATGGTGTGCGCGCATGCGCGC-3'). A fragment encoding SV5Tag-linker-pIX-EcoRI was isolated by PCR from a plasmid expressing a GFP-pIX fusion protein (E. Pesavento, unpublished work) using the primers SV5/F (5'-CACCATGCTAGCGGCAACCAATCCC-3') and pUC/R (5'-CGACGTTGTAAACGACGCG-3'). The HisTag and SV5Tag-linker-pIX-EcoRI fragments were then assembled by overlap PCR using His/F and pUC/R as primers. In this way, an ORF composed of HisTag-SV5Tag-linker-pIX, flanked by *SphI* and *EcoRI* restriction sites, was obtained. This ORF was cloned in a pDANS-sfGFP vector (A. Bradbury, to be published) in place of the sfGFP-pII ORF. The pDANS-sfGFP vector is based on a pUC19 plasmid with an M13 origin. The resulting phagemid, named pUPP9-HisTag (Figure 1b), was electroporated into *E. coli* TG1.

Phage Production and Purification. One liter of LB-Ampicillin (100 µg/mL) medium inoculated by TG1 harboring the phagemid pUPP9HisTag and already infected by the helper phage R408 was grown overnight at 37 °C. After centrifugation, the phages were purified from the supernatant by two successive PEG precipitations. The final pellet was dissolved in 1 mL of Milli-Q water and reprecipitated again for 20 min at 4 °C with 200 µL of PEG/NaCl (20% W/V PEG 6000/2.5 M NaCl). The mixture was centrifuged at 10000 rpm for 20 min at 4 °C and the precipitated phage was finally dissolved in 2 mL of Milli-Q water. This phage sample was loaded onto the His Trap FF column (prepacked with Ni sepharose) beforehand, prepared as recommended by the manufacturer's protocols, and equilibrated at a flow rate of 1 mL/min with the equilibration buffer (50 mM Tris-HCl, 100 mM NaCl, and 20 mM Imidazole pH 7.4). After washing the column with five volumes of equilibration buffer the his tagged phages were eluted with a 50 mM stage imidazole gradient (starting from 50 up to 450 mM) in the same buffer.

Phage Titer Measurements. Serial 10× dilutions of the purified phagemid particles were prepared and 10 µL of each one were mixed with 990 µL of a TG1 culture in exponential phase. After incubation at 37 °C without agitation for 30 min and with agitation for another 30 min, 100 µL of each mixture were spread on LB Petri dishes containing Ampicillin (100 µg/mL) and incubated overnight at 37 °C. The colonies were counted and the phage titer was calculated as colony-forming-units (cfu).

AFM Measurements. AFM measurements were performed at 20 °C in buffered solutions (PBS, pH 7.4), using a Nanoscope IV Multimode AFM (Veeco Metrology Group, Santa Barbara, CA) and oxide sharpened microfabricated Si₃N₄ cantilevers (Microlevers, Veeco Metrology Group). Force measurements were recorded in the force-volume mode consisting of arrays of 16 × 16 force curves on 10 µm × 10 µm areas, using a maximum applied force of 350 pN and a loading rate of 10000 pN/s, calculated by multiplying the tip pulling velocity (nm/s) by the slope of the peaks (pN/nm). A ramp delay of 500 ms was applied between tip approach and tip retraction, while keeping constant the maximum applied force. The spring constants of the cantilevers were measured using the thermal noise method (Picoforce, Veeco Metrology Group), yielding values ranging from 0.01 to 0.015 N/m. Force-volume data were treated using the Nanoscope V 7.30 (Veeco Metrology Group) and Matlab (The Mathworks Inc.) softwares.

For force spectroscopy, AFM tips were functionalized with anti-pII (Mo Bi Tec, Germany) using PEG-benzaldehyde linkers provided by H.J. Gruber (University of Linz), essentially as described elsewhere.³³ Cantilevers were washed with chloroform and ethanol, placed in an UV-ozone cleaner for 30 min, immersed overnight into an ethanolamine solution (3.3 g ethanolamine into 6 mL of DMSO), then washed three times with DMSO and two times with ethanol, and dried with N₂. The ethanolamine-coated cantilevers were immersed for 2 h in a solution prepared by mixing 1 mg of Aceta-PEG-NHS dissolved in

0.5 mL of chloroform with 10 µL of triethylamine, then washed with chloroform and dried with N₂. Cantilevers were further immersed for 5 min in a 0.1% iodine/acetone solution, washed three times in acetone, dried under N₂, and then covered with a 200 µL droplet of a PBS solution containing anti-pII (0.2 mg/mL) to which 2 µL of a 1 M NaCNBH₃ solution were added. After 50 min, cantilevers were incubated with 5 µL of a 1 M ethanolamine solution in order to passivate unreacted aldehyde groups, and then washed with and stored in PBS 10 min later.

For surface immobilization of the phages, silicon wafers (Siltronix) were coated using electron beam thermal evaporation with a 5-nm thick chromium layer followed by a 30-nm thick gold layer. The gold-coated surfaces were cleaned for 15 min by UV and ozone treatment, rinsed with ethanol, dried with a gentle nitrogen flow, and immersed overnight in ethanol containing 0.05 mM of NTA-terminated (10%) and triethylene glycol(EG)-terminated (90%) alkanethiols. After rinsing with ethanol, the tips and supports were immersed in a 40 mM aqueous solution of NiSO₄ (pH 7.2) for 1 h, rinsed with water, incubated in 1 mL of purified phage solution for 2 h, and finally rinsed with PBS. After rinsing, the support was carefully attached to a steel sample puck (Veeco Metrology Group) using a small piece of double face adhesive tape, and the mounted sample was transferred into the AFM liquid cell while avoiding dewetting.

Acknowledgment. The pDANS-sfGFP phagemid was a kind gift of Andrew Bradbury. We thank H. J. Gruber for providing us with the PEG linkers. This work was supported by the National Foundation for Scientific Research (FNRS), the Université catholique de Louvain (Fonds Spéciaux de Recherche), the Région wallonne (Pôle d'excellence NANOTIC), the Federal Office for Scientific, Technical and Cultural Affairs (Interuniversity Poles of Attraction Programme), and the Research Department of the Communauté française de Belgique (Concerted Research Action). Y.F.D. and D.A. are Senior Research Associate and Research Fellow of the FRS-FNRS.

REFERENCES AND NOTES

- Marvin, D. A. Filamentous Phage Structure, Infection, and Assembly. *Curr. Opin. Struct. Biol.* **1998**, *8*, 150–158.
- Russel, M.; Linderth, N. A.; Sali, A. Filamentous Phage Assembly: Variation on a Protein Export Theme. *Gene* **1997**, *192*, 23–32.
- Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors that Display Cloned Antigens on the Virion Surface. *Science* **1985**, *228*, 1315–1317.
- Minard, P.; Urvoas, A.; Soumillion, P., Protein Engineering by Phage Display. In *Protein Engineering Handbook*; Lutz, S., Bornscheuer, U., Eds.; Wiley-VCH Verlag GmbH: Weinheim, Germany, 2008; pp 563–604.
- Kehoe, J. W.; Kay, B. K. Filamentous Phage Display in the New Millennium. *Chem. Rev.* **2005**, *105*, 4056–4072.
- Mao, C. B.; Flynn, C. E.; Hayhurst, A.; Sweeney, R.; Qi, J. F.; Georgiou, G.; Iverson, B.; Belcher, A. M. Viral Assembly of Oriented Quantum Dot Nanowires. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6946–6951.
- Mao, C. B.; Solis, D. J.; Reiss, B. D.; Kottmann, S. T.; Sweeney, R. Y.; Hayhurst, A.; Georgiou, G.; Iverson, B.; Belcher, A. M. Virus-Based Toolkit for the Directed Synthesis of Magnetic and Semiconducting Nanowires. *Science* **2004**, *303*, 213–217.
- Sarikaya, M.; Tamerler, C.; Jen, A. K. Y.; Schulten, K.; Baneyx, F. Molecular Biomimetics: Nanotechnology through Biology. *Nat. Mater.* **2003**, *2*, 577–585.
- Ashkin, A.; Dziedzic, J. M. Optical Trapping and Manipulating of Viruses and Bacteria. *Science* **1987**, *235*, 1517–1520.
- Khalil, A. S.; Ferrer, J. M.; Brau, R. R.; Kottmann, S. T.; Noren, C. J.; Lang, M. J.; Belcher, A. M. Single M13 Bacteriophage Tethering and Stretching. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4892–4897.
- Ivanovska, I. L.; de Pablo, P. J.; Ibarra, B.; Sgalari, G.; MacKintosh, F. C.; Carrascosa, J. L.; Schmidt, C. F.; Wuite, G.

- G. J. L. Bacteriophage Capsids: Tough Nanoshells with Complex Elastic Properties. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7600–7605.
12. Rankl, C.; Kienberger, F.; Wildling, L.; Wruss, J.; Gruber, H. J.; Blaas, D.; Hinterdorfer, P. Multiple Receptors Involved in Human Rhinovirus Attachment to Live Cells. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17778–17783.
 13. Müller, D. J.; Dufrene, Y. F. Atomic Force Microscopy as a Multifunctional Molecular Toolbox in Nanobiotechnology. *Nat. Nanotechnol.* **2008**, *3*, 261–269.
 14. Müller, D. J.; Helenius, J.; Alsteens, D.; Dufrene, Y. F. Force Probing Surfaces of Living Cells to Molecular Resolution. *Nat. Chem. Biol.* **2009**, *5*, 383–390.
 15. Müller, D. J.; Engel, A.; Carrascosa, J. L.; Velez, M. The Bacteriophage Phi 29 Head–Tail Connector Imaged at High Resolution with the Atomic Force Microscope in Buffer Solution. *EMBO J.* **1997**, *16*, 2547–2553.
 16. Kuznetsov, Y. G.; Malkin, A. J.; Lucas, R. W.; Plomp, M.; McPherson, A. Imaging of Viruses by Atomic Force Microscopy. *J. Gen. Virol.* **2001**, *82*, 2025–2034.
 17. Malkin, A. J.; McPherson, A.; Gershon, P. D. Structure of Intracellular Mature Vaccinia Virus Visualized by *in Situ* Atomic Force Microscopy. *J. Virol.* **2003**, *77*, 6332–6340.
 18. Kienberger, F.; Rankl, C.; Pastushenko, V.; Zhu, R.; Blaas, D.; Hinterdorfer, P. Visualization of Single Receptor Molecules Bound to Human Rhinovirus under Physiological Conditions. *Structure* **2005**, *13*, 1247–1253.
 19. Russel, M.; Kidd, S.; Kelley, M. R. An Improved Filamentous Helper Phage for Generating Single-Stranded Plasmid DNA. *Gene* **1986**, *45*, 333–338.
 20. Verbelen, C.; Gruber, H. J.; Dufrene, Y. F. The NTA-His(6) Bond Is Strong Enough for AFM Single-Molecular Recognition Studies. *J. Mol. Recognit.* **2007**, *20*, 490–494.
 21. Berquand, A.; Xia, N.; Castner, D. G.; Clare, B. H.; Abbott, N. L.; Dupres, V.; Adriaensens, Y.; Dufrene, Y. F. Antigen Binding Forces of Single Antisoyzome Fv Fragments Explored by Atomic Force Microscopy. *Langmuir* **2005**, *21*, 5517–5523.
 22. Ros, R.; Schwesinger, F.; Anselmetti, D.; Kubon, M.; Schafer, R.; Pluckthun, A.; Tiefenauer, L. Antigen Binding Forces of Individually Addressed Single-Chain Fv Antibody Molecules. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 7402–7405.
 23. Schwesinger, F.; Ros, R.; Strunz, T.; Anselmetti, D.; Guntherodt, H. J.; Honegger, A.; Jermutus, L.; Tiefenauer, L.; Pluckthun, A. Unbinding Forces of Single Antibody–Antigen Complexes Correlate with Their Thermal Dissociation Rates. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9972–9977.
 24. Verbelen, C.; Christiaens, N.; Alsteens, D.; Dupres, V.; Baulard, A. R.; Dufrene, Y. F. Molecular Mapping of Lipoarabinomannans on Mycobacteria. *Langmuir* **2009**, *25*, 4324–4327.
 25. Morfil, J.; Neumann, J.; Blank, K.; Steinbach, U.; Puchner, E. M.; Gottschalk, K.-E.; Gaub, H. E. Force-Based Analysis of Multidimensional Energy Landscapes: Application of Dynamic Force Spectroscopy and Steered Molecular Dynamics Simulations to an Antibody Fragment–Peptide Complex. *J. Mol. Biol.* **2008**, *381*, 1253–1266.
 26. Merkel, R.; Nassoy, P.; Leung, A.; Ritchie, K.; Evans, E. Energy Landscapes of Receptor–Ligand Bonds Explored with Dynamic Force Spectroscopy. *Nature* **1999**, *397*, 50–53.
 27. Evans, E.; Ritchie, K. Dynamic Strength of Molecular Adhesion Bonds. *Biophys. J.* **1997**, *72*, 1541–1555.
 28. Sotomayor, M.; Schulten, K. Single-Molecule Experiments *in Vitro* and *in Silico*. *Science* **2007**, *316*, 1144–1148.
 29. Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *Science* **1997**, *276*, 1109–1112.
 30. Oberhauser, A. F.; Marszalek, P. E.; Erickson, H. P.; Fernandez, J. M. The Molecular Elasticity of the Extracellular Matrix Protein Tenascin. *Nature* **1998**, *393*, 181–185.
 31. Oesterhelt, F.; Oesterhelt, D.; Pfeiffer, M.; Engel, A.; Gaub, H. E.; Müller, D. J. Unfolding Pathways of Individual Bacteriorhodopsins. *Science* **2000**, *288*, 143–146.
 32. Alsteens, D.; Dupres, V.; Klotz, S. A.; Gaur, N. K.; Lipke, P. N.; Dufrene, Y. F. Unfolding Individual AkSp Adhesion Proteins on Live Cells. *ACS Nano* **2009**, *3*, 1677–1682.
 33. Ebner, A.; Wildling, L.; Kamruzzahan, A. S. M.; Rankl, C.; Wruss, J.; Hahn, C. D.; Hotl, M.; Zhu, R.; Kienberger, F.; Blaas, D.; et al. A New, Simple Method for Linking of Antibodies to Atomic Force Microscopy Tips. *Bioconjugate Chem.* **2007**, *18*, 1176–1184.

ARTICLE

Received 11 Sep 2013 | Accepted 13 Nov 2013 | Published 13 Dec 2013

DOI: 10.1038/ncomms3926

Multiparametric atomic force microscopy imaging of single bacteriophages extruding from living bacteria

David Alsteens¹, Heykel Trabelsi², Patrice Soumillion² & Yves F. Dufrène¹

Force-distance (FD) curve-based atomic force microscopy is a valuable tool to simultaneously image the structure and map the biophysical properties of biological samples at the nano-scale. Traditionally, FD-based atomic force microscopy has been severely limited by its poor temporal and lateral resolutions. Here we report the use of advanced FD-based technology combined with biochemically sensitive tips to image filamentous bacteriophages extruding from living bacteria at unprecedented speed and resolution. Directly correlated multiparametric images of the structure, adhesion and elasticity of infected bacteria demonstrate that the sites of assembly and extrusion localize at the bacterial septum in the form of soft nanodomains surrounded by stiff cell wall material. The quantitative nano-bio-imaging method presented here offers a wealth of opportunities for mapping the physical properties and molecular interactions of complex biosystems, from viruses to tissues.

¹Université catholique de Louvain, Institute of Life Sciences, Croix du Sud, 1, bte L7.04.01, B-1348 Louvain-la-Neuve, Belgium, ²Université catholique de Louvain, Institute of Life Sciences, Croix du Sud, 4-5, bte L7.07.06, B-1348 Louvain-la-Neuve, Belgium. Correspondence and requests for materials should be addressed to Y.F.D. (Yves.Dufrene@uclouvain.be).

Introduced in the 1990s, force-distance (FD) curve-based atomic force microscopy (AFM)^{1–4} enables researchers to correlate structural images of complex biological systems with quantitative maps of their biophysical properties. The principle is to record spatially resolved FD curves across the biological sample^{3,5}. From each curve, a given property (for example, adhesion or elasticity) is extracted, quantified and displayed as grey (or coloured) pixel, the brightness of which reflects the magnitude of the measured property. In doing so, spatially resolved maps of sample properties and interactions are produced. Until recently, the low speed and poor spatial resolution of this method have severely hampered its widespread use in biology. However, recent progress in FD-based AFM has enabled the circumvention of these limitations, and has provided a unique means to record directly correlated, multiparametric images of the structure, adhesion and elasticity of biological systems at (near) molecular resolution and at the speed of conventional topographic imaging (for a recent review, see Dufrene *et al.*⁶). This newer technology has already been used to simultaneously image the structure and the biophysical properties of complex specimens, including human keratinocytes⁷, diatoms⁸, yeasts⁹, lipid phases¹⁰, membrane proteins^{11–13}, virus capsids¹⁴ and pathological neurofibrils¹⁵. In the membrane protein context, the method was applied to map quantitatively the flexibility of individual membrane proteins in their native, folded state at unprecedented submolecular resolution and to correlate protein flexibility with crystal structure¹³. In other studies, the technique enabled structural mapping of the deformation, adhesion, energy dissipation and elasticity of native proteins at a resolution of a few nanometres¹¹, and quantification of the electrostatic field and potential generated by transmembrane protein pores at subnanometer resolution¹². Although valuable, these experiments lack biochemical specificity, that is, they cannot identify nor force probe specific constituents (for example, a given receptor) in complex, heterogeneous systems (for example, living cells). Hence, the quantitative multiparametric imaging of specific biochemical sites using advanced FD-based AFM remains very challenging.

Here, we combine the use of advanced FD-based AFM with specific microscope tips to simultaneously image the structure of heterogeneous biological systems and quantitatively map their adhesion and mechanical properties. As a prototype complex biological sample, we investigate the extrusion of individual filamentous bacteriophages from the surface of living bacteria. Filamentous bacteriophages (Φ), such as fd, f1 and M13, are 880 nm long and 6 nm-wide structures with a circular single-stranded DNA (ssDNA) genome, that are able to infect a variety of Gram-negative bacteria such as *Escherichia coli* (Fig. 1a)^{16–18}. The ssDNA encodes three classes of proteins, that is, replication (pII, pX and pV), morphogenetic (pI, pIV and pXI) and structural (pVIII, pIII, pVI, pVII and pIX) proteins (Fig. 1a). The constitutive proteins are anchored in the inner membrane and incorporated into the virion using a specialized assembly site working as an aqueous channel where the inner and outer membranes are in close contact (Fig. 1b)^{19,20}. Initiation of assembly starts with the addition of pIX and pVII on the packaging signal. Then, pV, which collapses the ssDNA into a flexible rod in the cytosol, is progressively replaced by the major coat protein pVIII as the phage particle is extruded from the outer membrane (Fig. 1b). Once the ssDNA is coated by pVIII, pIII and pVI are added to the late extremity of the virion and the phage is released from the cell, a process that can be blocked by the deletion of pIII (Φ pIIId)^{17,18,21}. Although phage replication and assembly have been extensively studied *in vitro*, the cellular localization of the sites of phage assembly and release are poorly

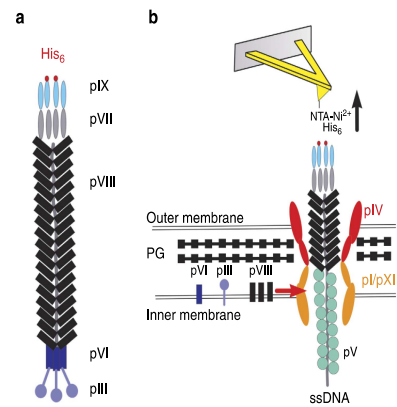


Figure 1 | Probing single bacteriophages using AFM with biochemically sensitive tips. (a) Filamentous bacteriophages were genetically engineered to display His-tag on their pIX tail. (b) Single phages assembled in and escaping from the bacterial cell wall were detected via their His-pIX tail using an AFM tip modified with Ni^{2+} -NTA groups. Shown in the cartoon are the main components of the assembly machinery (see text for details).

understood, owing to the paucity of high-resolution imaging techniques. To address this issue, in this work *E. coli* bacteria infected by His-tagged Φ particles are imaged using advanced FD-based AFM with biochemically sensitive tips, that is, tips functionalized with Ni^{2+} -NTA groups. Directly correlated multiparametric images of the structure, adhesion and elasticity of infected bacteria enable us to demonstrate that the sites of the assembly machinery localize at the bacterial septum in the form of soft nanodomains surrounded by stiff cell wall material. This biospecific multiparametric imaging method offers unprecedented possibilities to quantify and spatially map the multiple biophysical properties of complex biomolecular and cellular systems in physiological conditions.

Results

Fluorescence imaging of infected bacteria. We first analysed the expression of pIX-His-tagged bacteriophages (f1 filamentous bacteriophages²²; Φ w1) in *E. coli* bacteria using fluorescence microscopy, following labelling with Ni^{2+} -NTA fluorescent conjugates (Fig. 2). Phage production occurred predominantly near the septum and the poles of the bacteria (Fig. 2a,d). We compared the behaviour of wt phages with that of mutant phages in which the pIII gene had been inactivated by sequence frameshift (Φ pIIId). pIII proteins are required for infectivity and stability of the particles. Bacteria infected with Φ pIIId produce particles that are not released into the medium but remain associated at the surface in the assembly machinery²³. Bacteria infected with Φ pIIId were homogeneously labelled (Fig. 2b,e), confirming that phages that lack pIII stayed anchored into the cell wall. Labelling specificity was confirmed by showing that non-infected bacteria were not fluorescent (Fig. 2c,f).

FD-based AFM visualizes single bacteriophages. We then imaged *E. coli* bacteria infected by Φ w1 using FD-based AFM with Ni^{2+} -NTA-terminated tips (Fig. 3). While structural images revealed a smooth, featureless surface (Fig. 3a), directly correlated adhesion maps were highly contrasted, adhesion events being

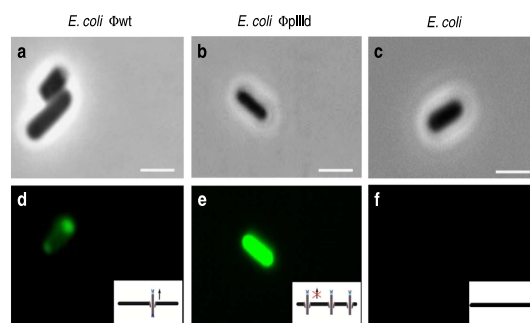


Figure 2 | Fluorescence imaging of bacteria infected by bacteriophages. Phase (a–c) and fluorescence (d–f) images of *E. coli* cells infected by Φ wt (a,d) or by Φ pIII Δ particles lacking pIII (b,e) and of non-infected *E. coli* cells (c,f), following labelling with Ni^{2+} -NTA fluorescent conjugates. Scale bar 2 μm .

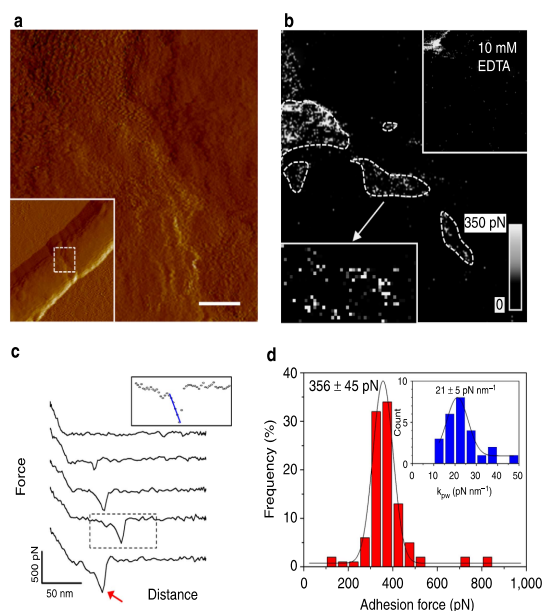


Figure 3 | Imaging the extrusion of single bacteriophages from living bacteria using FD curve-based AFM. (a) FD-based AFM structural image (scale bar, 200 nm; error signal; inset: low-resolution image; the high-resolution image was recorded in the dashed square) and (b) directly correlated adhesion map (upper inset: same experiment in the presence of EDTA; lower inset: enlarged view of an adhesive area) of an *E. coli* cell infected with Φ wt particles recorded in PBS with a Ni^{2+} -NTA tip. (c) Representative FD curves. The red arrow on the bottom curve indicates the maximum adhesion force corresponding to the rupture of the His- Ni^{2+} -NTA bond. The linear shape of the adhesion force peaks (inset: linear fit) indicates that the phage-cell wall complex behaves as a Hookean spring. (d) Histogram of adhesion forces ($n = 500$ adhesion events from three independent experiments, that is, using three different tip preparations and three independent cultures). The inset shows the distribution of the spring constant values associated with the phage-cell wall complex ($n = 500$).

primarily detected in the septum or polar regions (Fig. 3b). In these regions, the adhesion events corresponded to an apparent surface density of $617 \pm 294 \text{ sites } \mu\text{m}^{-2}$ ($n = 1,700$ FD curves from 17 different maps from three independent experiments). As FD curves are recorded at frequencies three orders of magnitudes higher

than in classical FD-based AFM, the imaging speed and resolution are greatly increased (~ 8 min for a 256×256 pixel image versus ~ 80 min for a 64×64 pixel image). From high magnification views (Fig. 3b, lower inset), we estimate that the resolution of our adhesion images is a few nanometres. Increasing the resolution further,

although technically feasible, would have no sense in this study as this would lead to detecting the same sites multiple times. We note that structural and adhesion maps were obtained at a maximum applied force of 1 nN, yielding an indentation of ~ 50 nm. The contrast of our structural images (Fig. 3) was similar to that obtained in conventional contact mode, and was not altered by the applied force (within the 0.1–1 nN range).

We believe that the measured adhesion events correspond to the specific detection of single phage particles for several reasons. First, non-infected bacteria imaged with the same method showed essentially no adhesion events (Fig. 4i); second, treatment of infected bacteria with 10 mM EDTA led to a substantial decrease of adhesion frequency, indicating that NTA tips lacking Ni^{2+} are no longer able to detect His-tagged particles (Fig. 3b, upper inset); third, bacteria infected with ΦpIIIId were massively labelled (Fig. 4g), confirming that phages lacking pIII remain surface-associated; and fourth, adhesive FD curves showed well-defined adhesion force peaks (Fig. 3c, see red arrow on lower curve) of 356 ± 45 pN (mean $\pm$ s.d., $n=500$ adhesive events from three independent experiments) (Fig. 3d). These adhesion events correspond to the rupture of the bond between the Ni^{2+} -NTA tip and a single His-tagged bacteriophage since they are in the range of the force expected for single Ni^{2+} -NTA–His interactions at the loading rate that we used ($\sim 10^7$ pN s $^{-1}$) (refs 24–26).

Interestingly, most adhesion events were not randomly distributed but organized into nanodomains, 190 ± 258 nm 2 in size ($n=105$ domains from 17 different maps from five independent experiments) (Fig. 3b; clusters are highlighted by dashed lines). The nanoscale organization of Φwt is reminiscent of the behaviour of lytic bacteriophages shown to form discrete 50–200 nm lytic holes at the septum and occasionally at the poles 27 . We suggest that the assembly machinery can only be located near the septum (and the poles), where peptidoglycan is synthesized. Indeed, pI, pIV and pXI are known to form a channel across the whole bacterial cell wall breaking the peptidoglycan layer 28 . As no enzymes encoded by the phage are able to disrupt peptidoglycan, we expect that the assembly machinery should be integrated at the sites of peptidoglycan synthesis, that is, where cell division takes place.

Nanomechanics of the phage–cell wall complex. How do single phages respond to mechanical force? FD curves recorded for Φwt were well-described by a linear fit, rather than by a worm-like chain model, which is characteristic of a Hookean behaviour (Fig. 3c, inset). Considering the spring constant of the AFM cantilever (k_c) and the slope (s) of the linear portion of the raw deflection versus piezo displacement curves, the apparent spring constant of the phage–cell wall complex (k_{pw}) is given by $k_{\text{pw}} = (k_c \times s)/(1 - s) = 21 \pm 5$ pN nm $^{-1}$ (mean $\pm$ s.d., $n=500$ adhesive curves from five independent experiments) (Fig. 3d, inset). As the bacteriophage is a non-deformable filament 29,30 , k_{pw} primarily reflects the bacterial cell wall stiffness. Supporting this view, our measured values are close to the bacterial spring constant measured for *E. coli* 31 . We note that FD curves recorded for bacteria infected with ΦpIIIId particles also showed linear force behaviours with similar spring constant values.

Notably, the linear force peaks that we observed here are very different from those reported earlier for Φwt particles that were simply adsorbed on solid surfaces 30 . Pulling on adsorbed phages yielded non-linear force peaks that were preceded by extended elongation events, up to several hundreds nanometres, reflecting the progressive desorption and straightening of the phages into extended orientations. Such force signatures were never observed here, leading us to conclude that the detected Φwt particles were

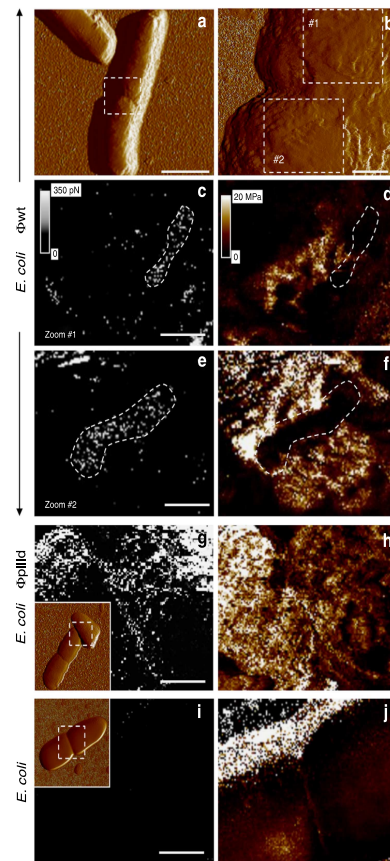


Figure 4 | Bacteriophage extrusion localizes into soft nanodomains. (a) Low- and (b) high-resolution FD-based AFM structural images (error signal) of *E. coli* cells infected by Φwt particles recorded in PBS with Ni^{2+} -NTA AFM tips. Scale bar in a, 2 μm . Scale bar in b, 300 nm. (c,e,g,i) Directly correlated adhesion maps and (d,f,h,j) elasticity maps (z-scale shows Young's modulus values) of *E. coli* cells infected by Φwt (c–f) or by ΦpIIIId particles lacking pIII (g,h) and of non-infected *E. coli* cells (i,j). Scale bars are: (c) 150 nm, (e) 150 nm, (g) 500 nm and (i) 300 nm. Images c, d, e and f correspond to the zoomed areas shown in panel b. The insets in panels g and i are the low-resolution topography images on which images g–j were recorded (see dashed squares). The dashed lines in c–f emphasize the organization of the bacteriophages into soft nanodomains surrounded by stiff material. Similar results were obtained in seven independent experiments (that is, experiments using seven different tip preparations and seven independent cultures) for *E. coli* cells infected by Φwt particles (see also Fig. 5), in five independent experiments for *E. coli* cells infected by ΦpIIIId particles and in three independent experiments for non-infected *E. coli* cells.

actively extruded from the bacteria and not simply adsorbed on the cell surface in a non-specific manner. Consistent with this, similar linear force peaks were observed for bacteria infected with ΦpIIIId particles.

Bacteriophage extrusion localizes into soft nanodomains. Next, we used multiparametric imaging to correlate the structural, adhesion and mechanical properties of infected bacteria (Figs 4 and 5). Correlated adhesion maps (Figs 4c,e and 5a,c,e) and elasticity (Young's modulus) maps (Figs 4d,f and 5b,d,f) obtained for bacteria infected by Φ wt revealed that bacteriophages were preferentially detected in soft regions (0.6 ± 0.2 MPa; $n=100$ curves from three independent experiments), somewhat similar to the softness of native bacteria (1.0 ± 0.3 MPa; $n=100$ curves from three independent experiments), and generally surrounded by much stiffer cell wall properties (11.4 ± 5.9 MPa; $n=100$ curves from three independent experiments). Such quantitative correlations were found in multiple independent experiments (Figs 4 and 5). Cells infected

by Φ pIIIId (Fig. 4g,h) showed a dense distribution of phages in the adhesion image, that was correlated with heterogeneous, complex cell wall properties, that is, soft spots surrounded by a stiff matrix. By contrast, non-infected cells (Fig. 4i,j) showed no phage detection and homogeneously soft properties (1.0 ± 0.3 MPa; $n=100$ curves from three independent experiments) that are typical of those reported for native *E. coli* bacteria³². The differences in local mechanical properties that we unravelled here may be biologically relevant. Because the assembly machinery of the phages is localized near the septum, that is, where peptidoglycan synthesis takes place, we expected that these nanodomains may interfere with peptidoglycan assembly, leading to the accumulation of newly formed cell wall material around these sites.

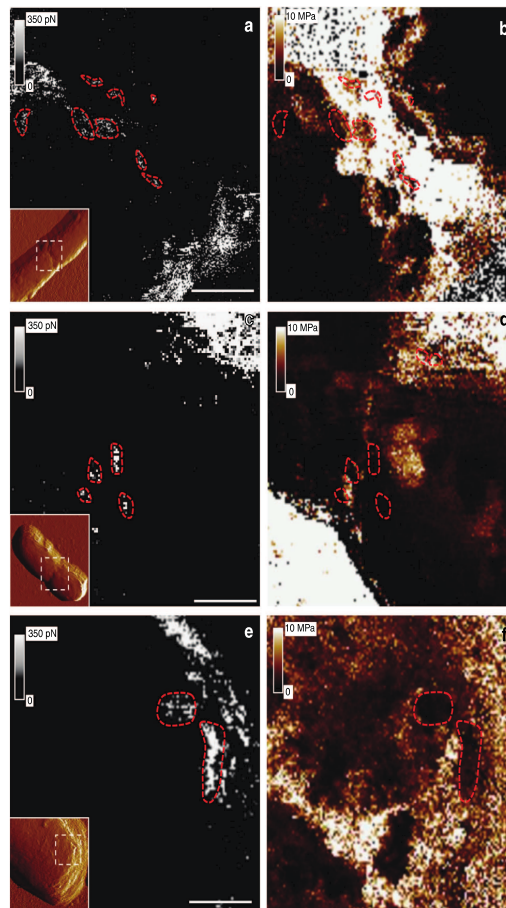


Figure 5 | Simultaneous imaging of the adhesion and elasticity of infected bacteria. (a,c,e) Adhesion maps and (b,d,f) directly correlated elasticity maps recorded by FD-based AFM with Ni^{2+} -NTA AFM tips on *E. coli* cells infected by Φ wt. The insets in panels a,c,e are the low-resolution structural images on which adhesion and elasticity images were recorded. The dashed lines emphasize the organization of the bacteriophages into soft nanodomains surrounded by stiff material. Scale bars are: (a) 250 nm, (c) 250 nm and (e) 125 nm.

Discussion

A crucial challenge in current biological sciences is to characterize the structure, physical properties and interactions of cellular and biomolecular systems under physiological conditions. Our experiments demonstrate that advanced FD-based AFM combined with biochemically sensitive tips represents a powerful multiparametric characterization platform for the simultaneous imaging of the structural and biophysical properties of complex biosystems. To the best of our knowledge, this is the first time that correlated structural, adhesion and elasticity images of complex biological samples are recorded at high temporal and spatial resolutions, and with biochemical specificity. The speed (few minutes) and resolution (few nanometres) are similar to those of classical topographic imaging, thus exceeding markedly the performances of conventional FD-based AFM. In addition to mapping single-receptor sites and sample elasticity, individual FD curves can also provide quantitative information on the nanomechanics of the stretched receptors. Using this method, we provide the first direct visualization of the assembly machinery of bacteriophages on live cells, revealing that it localizes near the septum in the form of soft nanodomains surrounded by stiffer cell wall material. Applicable to a wide variety of complex biological systems, from viruses to tissues, this quantitative, biospecific imaging method offers new opportunities for understanding crucial biological processes and functions.

Methods

Bacterial strains. *E. coli* TGI: F' [traD36 proAB⁺ lacP⁺ lacZAM15]supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSMJ5, (r_K m_C⁻)) was from Invitrogen. His-tag pIX phages were produced by the strain TGI harbouring the phagemid pUPP9His/Tag⁴⁰ and already infected by the helper phage R408, a derivative of the filamentous phage f1 (ref. 22). His-tagged pIX pIII phages (pIII deficient) were produced by the strain TGI, transformed with the pUPP9His/Tag phagemid and infected by a filamentous phage in which the pIII gene is inactivated by a frameshift in an open reading frame inserted between the sequences encoding the signal peptide and the mature pIII³³.

Fluorescence microscopy. *E. coli* bacteria were washed three times by centrifugation (2,000 g for 5 min) in phosphate-buffered saline (PBS), incubated in 50 μg ml⁻¹ of Ni²⁺-NTA-Atto conjugates for 1 h in the dark at 20 °C and washed two times with PBS. Images were acquired with a Zeiss Axio Observer Z1 equipped with a Hamamatsu camera C10600.

AFM measurements. AFM measurements were performed at 20 °C in buffered solutions (PBS, pH 7.4) with a Bioscope Catalyst AFM (Bruker AXS Corporation, Santa Barbara, CA), using the newly developed peak force tapping mode (Peak-Force-Quantitative Nano-Mechanics). Cells were immobilized onto high-molecular weight polyethylenimine-coated glass petri dishes, as described elsewhere²⁴. Bacterial cultures (200 μl) were centrifuged three times at 2,000g for 5 min and resuspended in PBS. The cell suspension was deposited onto polyethylenimine-coated glass for 1 h and then extensively rinsed with PBS. V-shaped Scan Asyst fluid cantilevers (Bruker AXS Corporation) with a spring constant of 0.45–0.77 N m⁻¹ were coated by electron beam thermal evaporation with a 5-nm-thick Cr layer followed by a 30-nm-thick Au layer. Gold-coated cantilevers were immersed overnight in ethanol containing 0.1 mM of NTA-terminated (10%) and triethylene glycol-terminated (90%) alkanethiols (ProChimia, Poland), rinsed with ethanol and then further incubated with a 40-mM solution of Ni²⁺. Multiparametric images were recorded at 256 pixels × 256 pixels with a driving frequency of 0.25 kHz, and an applied force of 1 nN. Cantilever spring constant and sensitivity were calibrated before and after each experiment. Data processing was performed using the commercial Nanoscope Analysis software (Bruker AXS Corporation). Young's moduli were calculated using the Sneddon model for a conical tip.

References

- Ludwig, M., Dettmann, W. & Gaub, H. E. Atomic force microscope imaging contrast based on molecular recognition. *Biophys. J.* **72**, 445–448 (1997).
- Gad, M., Itoh, A. & Ika, A. Mapping cell wall polysaccharides of living microbial cells using atomic force microscopy. *Cell Biol. Int.* **21**, 697–706 (1997).
- Heinz, W. F. & Hoh, J. H. Spatially resolved force spectroscopy of biological surfaces using the atomic force microscope. *Trends Biotechnol.* **17**, 143–150 (1999).
- Grandbois, M., Dettmann, W., Benoit, M. & Gaub, H. E. Affinity imaging of red blood cells using an atomic force microscope. *J. Histochem. Cytochem.* **48**, 719–724 (2000).
- Hinterdorfer, P. & Dufrene, Y. F. Detection and localization of single molecular recognition events using atomic force microscopy. *Nat. Methods* **3**, 347–355 (2006).
- Dufrene, Y. F., Martinez-Martin, D., Medalsy, I., Alsteens, D. & Müller, D. J. Multiparametric imaging of biological systems by force-distance curve-based AFM. *Nat. Methods* **10**, 847–854 (2013).
- Heu, C., Berquand, A., Elie-Caille, C. & Nicod, L. Glyphosate-induced stiffening of HaCaT keratinocytes, a peak force tapping study on living cells. *J. Struct. Biol.* **178**, 1–7 (2012).
- Pletkapić, G., Berquand, A., Radić, T. M. & Svetličić, V. Quantitative nanomechanical mapping of marine diatom in seawater using peak force tapping atomic force microscopy. *J. Physiol.* **48**, 174–185 (2012).
- Chopinet, L., Formosa, C., Rols, M. P., Duval, R. E. & Dague, E. Imaging living cells surface and quantifying its properties at high resolution using AFM in QITM mode. *Micron* **48**, 26–33 (2013).
- Picas, L., Rico, F. & Scheuring, S. Direct measurement of the mechanical properties of lipid phases in supported bilayers. *Biophys. J.* **102**, L01–L03 (2012).
- Medalsy, I., Hensen, U. & Müller, D. J. Imaging and quantifying chemical and physical properties of native proteins at molecular resolution by force-volume AFM. *Angew. Chem. Int. Ed. Engl.* **50**, 12103–12108 (2011).
- Pfreundschuh, M., Hensen, U. & Müller, D. J. Quantitative imaging of the electrostatic field and potential generated by a transmembrane protein pore at subnanometer resolution. *Nano Lett.* **13**, 5585–5593 (2013).
- Rico, F., Su, C. & Scheuring, S. Mechanical mapping of single membrane proteins at submolecular resolution. *Nano Lett.* **11**, 3983–3986 (2011).
- Carrasco, C. *et al.* DNA-mediated anisotropic mechanical reinforcement of a virus. *Proc. Natl Acad. Sci. USA* **103**, 13706–13711 (2006).
- Wegmann, S., Medalsy, I. D., Mandelkow, E. & Müller, D. J. The fuzzy coat of pathological human Tau fibrils is a two-layered polyelectrolyte brush. *Proc. Natl. Acad. Sci. USA* **110**, E313–E321 (2013).
- Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317 (1985).
- Russel, M., Linderth, N. A. & Sali, A. Filamentous phage assembly: variation on a protein export theme. *Gene* **192**, 23–32 (1997).
- Marvin, D. A. Filamentous phage structure, infection and assembly. *Curr. Opin. Struct. Biol.* **8**, 150–158 (1998).
- Marciano, D. K., Russel, M. & Simon, S. M. Assembling filamentous phage occlude pIV channels. *Proc. Natl Acad. Sci. USA* **98**, 9359–9364 (2001).
- Marciano, D. K., Russel, M. & Simon, S. M. An aqueous channel for filamentous phage export. *Science* **284**, 1516–1519 (1999).
- Rakonjac, J., Feng, J. N. & Model, P. Filamentous phage are released from the bacterial membrane by a two-step mechanism involving a short C-terminal fragment of pIII. *J. Mol. Biol.* **289**, 1253–1265 (1999).
- Russel, M., Kidd, S. & Kelley, M. R. An improved filamentous helper phage for generating single-stranded plasmid DNA. *Gene* **45**, 333–338 (1986).
- Rakonjac, J. & Model, P. Roles of pIII in filamentous phage assembly. *J. Mol. Biol.* **282**, 25–41 (1998).
- Verbelen, C., Gruber, H. J. & Dufrene, Y. F. The NTA-His₆ bond is strong enough for AFM single-molecular recognition studies. *J. Mol. Recognit.* **20**, 490–494 (2007).
- Friddle, R. W., Noy, A. & De Yoreo, J. J. Interpreting the widespread nonlinear force spectra of intermolecular bonds. *Proc. Natl Acad. Sci. USA* **109**, 13573–13578 (2012).
- Alsteens, D. *et al.* High-resolution imaging of chemical and biological sites on living cells using peak force tapping atomic force microscopy. *Langmuir* **28**, 16738–16744 (2012).
- Witte, A. *et al.* Endogenous transmembrane tunnel formation mediated by ΦX174 lysis protein E. *J. Bacteriol.* **172**, 4109–4114 (1990).
- Russel, M. Filamentous phage assembly. *Mol. Microbiol.* **5**, 1607–1613 (1991).
- Khalil, A. S. *et al.* Single M13 bacteriophage tethering and stretching. *Proc. Natl Acad. Sci. USA* **104**, 4892–4897 (2007).
- Alsteens, D. *et al.* Controlled manipulation of bacteriophages using single-virus force spectroscopy. *ACS Nano* **3**, 3063–3068 (2009).
- Vadillo-Rodriguez, V., Schooling, S. R. & Dutcher, J. R. In situ characterization of differences in the viscoelastic response of individual gram-negative and gram-positive bacterial cells. *J. Bacteriol.* **191**, 5518–5525 (2009).

32. Francius, G. *et al.* Bacterial surface appendages strongly impact nanomechanical and electrokinetic properties of *Escherichia coli* cells subjected to osmotic stress. *PLoS One* **6**, e20066 (2011).
33. Legendre, D., Soumillion, P. & Fastrez, J. Engineering a regulatable enzyme for homogeneous immunoassays. *Nat Biotechnol.* **17**, 67–72 (1999).
34. Vadiño-Rodríguez, V. *et al.* Comparison of atomic force microscopy interaction forces between bacteria and silicon nitride substrata for the three commonly used immobilization methods. *Appl. Environ. Microbiol.* **70**, 5441–5446 (2004).

Acknowledgements

This work was supported by the National Fund for Scientific Research (FRS-FNRS), the Université catholique de Louvain (Fonds Spéciaux de Recherche), the Federal Office for Scientific, Technical and Cultural Affairs (Interuniversity Poles of Attraction Programme) and the Research Department of the Communauté française de Belgique

(Concerted Research Action). Y.F.D. and D.A. are, respectively, Research Director and Postdoctoral Researcher of the FRS-FNRS.

Author contributions

D.A., H.T., P.S. and Y.F.D. designed research; D.A. and H.T. performed research; D.A., H.T., P.S. and Y.F.D. analysed the data; and D.A., H.T., P.S. and Y.F.D. wrote the paper.

Additional information

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Alsteens, D. Multiparametric atomic force microscopy imaging of single bacteriophages extruding from living bacteria. *Nat. Commun.* **4**:2926 doi: 10.1038/ncomms3926 (2013).