The catalytic activity of a DD-peptidase impairs its evolutionary conversion into a beta-lactamase

Gol Mohammad Dorrazehi
Biochemistry and Genetics of Microorganisms (BGM)
Louvain Institute of Biomolecular Science and Technology (LIBST)
Université catholique de Louvain

A thesis submitted for the academic degree of PhD in Science

January 2022

Members of the jury

Prof. François Chaumont (President of the jury)
Université catholique de Louvain

Prof. Waldemar Vollmer
Newcastle University

Prof. Florian Hollfelder
University of Cambridge

Prof. Pascal Hols
Université catholique de Louvain

Prof. Patrice Soumillion (Supervisor)
Université catholique de Louvain

I dedicate this thesis to those children all over the world who are deprived of quality education and never get the opportunity to participate in scientific research, especially the deprived children in my country Iran and my homeland Baluchistan.

Acknowledgments

Although the quest for knowledge has no end, this amazing journey is close to the destination, hence I would like to thank those without them making this journey wouldn't have been possible. For me this unique opportunity was the fulfilment of a dream and I feel deeply fortunate that I have done my PhD dissertation under thorough supervision of Professor Patrice Soumillion. Therefore, my immense gratitude is to him for his enlightening role along this journey. His full support and intellectual guidance was always available for me, and learning from him is the main cause behind this success. Patrice opened my eyes into scientific world of discovering fundamental facts and taught me how to think outside of the box to interpret and question every research observation. Thank you Prof. Patrice!

I would like to sincerely thank all my PhD jury members for their time and the invaluable intellectual discussions and comments that greatly contributed in improving the quality of this thesis.

Thanks to UCLouvain and LIBST for providing the dynamic research culture, international environment and the beautiful campus in Louvain-la-Neuve, which we found a place to learn and to grow. I also thank the Marie Skłodowska-Curie Actions, the European Union's reference programme for doctoral education, for funding this research and providing the great collaborative networks such as ES-Cat for young researchers to further their career stages. A special appreciation for Hollfelder's lab in University of Cambridge for organizing and hosting the ES-Cat and thanks to their efforts and kindness during the secondment and workshops that I spent there. Thank you to Bornberg's lab at University of Münster for hosting my other secontment under ES-Cat.

I will never forget BGM, its friendly and positive environment and my great colleagues for their parts in this journey. Many thanks to Professor Pascal Hols, Professor Bernard Hallet and Dr. Benoît Desguin for their supports and discussions in the BGM group. So much thanks to Véronique Lebrun for all the administrative supports. Thanks a lot to Brigitte Delplace and Damien Evrard for all their excellent technical and logistic supports and management in the lab and thank you Damien for all the contributions in the 'PBP-A' project. Special thanks to you my dear friends Sebastian Worms, Jason Baby Chirakadavil and Steve Alarcia for providing me the biggest helps in the team and during the time in Belgium. Thank you Steve for all the time you gave for the NGS data analysis and python scripting! Thanks to the current and old members of Soumillion's Lab, Alessandro Ratto, Emiliano David Primo, Romain Deschryver, Zia Tariq, Serena Gargiulo, Alexander Stephany, Roland Jaumain, Martin Desmet, Pierre Galka, Julia Wessel, Mohammad Shahneawz Khan and Antoine Godeaux.

Thanks to all the members of BGM lab for making the friendly workplace; thanks to Cyril Magy, Adrien Knoops, Frédéric Toussaint, Johann Mignolet, Claire Stulemeijer, Ophélie d'Udekem, Stefano Di Giacomo, Laura Ledesma, Marie Henry de Frahan, Julien Damoczi, Géraldine Houssa, Maricruz Fernandez, Nicolas Aryanpour, Julian Urdiain-Arraiza, Guillaume Cerckel, Chloé Vanhemmeseel, Florence Vande Capelle, Tamara Pringels and Gérôme Goossens.

I could not get where I am today without the support of my family, especially my father, thank you baba. I thank my parents and all the members of my family and family in-law for their patience that we could not meet for long, and their unconditional wishes and prayers from the far distance. I should mention that I could not make this PhD without my wife (Mahboobeh Dorrazehi), who was beside me throughout every moment of this journey and her unconditional love of family was the source of my encouragement and keeping me enthusiastic for the future of our children, family and society. And a big thank to my two little angels, Haanol and Pahrgol, for making this journey and my life more beautiful. Thanks to all my friends and relatives back home for their encouragements and best wishes. Thanks to all my teachers and mentors from schools to universities back in Iran and Malaysia.

تشکر از تشویق ها و دلگرمی هایتان

شمئ نیکیں واهگانی منت وار
امیت کہ نوکیں پَدریچ گوں أوپار ءُ گوں زانت ءُ زائگءَ رُژنائیں باندانیءِ شوهازءَ بہ رؤت

_ زگریں هون ثُرمپ دلءِ دانکہ نہ بنت مار مُهرَگ
پہ هَلو هالو دَج ایت واهگءِ سیہ مار کےءَ _ *- سید هاشمی*

Thanks to all

Gol Mohammad Dorrazehi

January 2022

Funding acknowledgements

I gratefully acknowledge the funding and trainings provided by Marie Skłodowska-Curie Actions for my doctoral education as Early Stage Researcher (ESR) in the ES-Cat network.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722610 (ES-Cat).
This material reflects only the author's views and the Union is not liable for any use that may be made of the information contained therein.

Abbreviations and Symbols

Amino Acid	3 Letters	1 Letter	Mass	
			Monoisotopic	Average
Alanine	Ala	A	71.03711	71.0788
Arginine	Arg	R	156.10111	156.1875
Aspargine	Asn	N	114.04293	114.1038
Aspartic acid	Asp	D	115.02694	115.0886
Cysteine	Cys	C	103.00919	103.1388
Glutamic acid	Glu	E	129.04259	129.1155
Glutamine	Gln	Q	128.05858	128.1307
Glycine	Gly	G	57.02146	57.0519
Histidine	His	H	137.05891	137.1411
Isoleucine	Ile	I	113.08406	113.1594
Leucine	Leu	L	113.08406	113.1594
Lysine	Lys	K	128.09496	128.1741
Methionine	Met	M	131.04049	131.1926
Phenylalanine	Phe	F	147.06841	147.1766
Proline	Pro	P	97.05276	97.1167
Serine	Ser	S	87.03203	87.0782
Threonine	Thr	T	101.04768	101.1051
Tryptophan	Trp	W	186.07931	186.2132
Tyrosine	Tyr	Y	163.06333	163.1760
Valine	Val	V	99.06841	99.1326

°C	**Degree Celsius**
βME	B-mercaptoethanol or 2-Mercaptoethanol
AA	Amino acid
Amp	Ampicillin
bp	Base pair
Cas9	CRISPR-associated protein 9
CFU	Colony forming unit
CRISPR	clustered regularly interspaced short palindromic repeats
D-Ala	D-Alanine
dATP	DesoxyAdenosine Triphosphate
dCTP	DesoxyCytosine Triphosphate
dGTP	DesoxyGuanosine Triphosphate
DNA	Desoxyribonucleic acid
dNTP	DesoxyNucleotide Triphosphate
dTTP	DesoxyThymidine Triphosphate
Gluc	Glucose
h	Hour
HPLC	High-performance liquid chromatography
IM	Inner membrane
IPTG	Isopropyl-β-thiogalactoside
Kan	Kanamycine
KDa	Kilo Dalton
LB	Luria Bertani
LC	Liquid chromatography–mass spectrometry
L-ara	L-arabinose
L-rha	L-rhamnose
Mg	Milligram
MIC	Minimal inhibitory concentratin

min	Minute
ml	Millilitre
mM	Millimolar
MS	Mass spectrometry
ng	Nanogram
NGS	Next Generation Sequencing
Ni	Nickel
nm	Nanomolar
OM	Outer membrane
PBP	Penicillin Binding Protein
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PenG	Penicillin G
PG	Peptidoglycan
Rmsd	Root Mean square deviation
RNA	Ribonucleic acid
rpm	Rotation Per minute
s	Second
SDS	Sodium dodecyl sulfate
S-S	Disulfide bond
Tet	Tetracyclin
UV	Ultaviolet
WGS	Whole genome sequencing
WT	Wild-type

Thesis Summary

Despite the wealth of knowledge on enzymes, their evolutionary emergence is still poorly understood. The classical view that enzymes are highly precise molecular machines has been replaced with the fact that enzymes, as ensemble of conformers, are able to carry secondary activities. Although these unwanted activities are less intense compared to what the enzymes have been specialized for, they may interfere with the metabolism and impose a cost to the cell. Therefore, in *in vivo* directed evolution of enzymes, restraining effects of fitness costs on the host organism would become beneficial for success. Through successes and failures on directed evolution of a D-alanyl-D-alanine-peptidase (DD-peptidase) from *Thermosynechococcus elongatus*, named PBP-A, into a beta-lactamase, we have accumulated evidences that this conversion follows a counter-intuitive evolutionary trajectory that is hindered by fitness costs on the host bacteria. In this study, we envisioned to further explore the molecular origins of this fitness cost and to evaluate molecular mechanisms that may contribute to fitness improvement. For that, we initially characterized the PBP-A enzyme by *in vitro* assays and found that it carries several activities, beside the previously reported DD-carboxypeptidase (DD-CPase) activity. One of the activities that was found to be effective on peptidoglycan of the host *Escherichia coli* cells, was DD-endopeptidase (DD-EPase) activity. Analysis of the peptidoglycan of the *E. coli* cells under expression of PBP-A showed a less-crosslinked peptidoglycan as the main origin of the fitness cost. To overcome this cost, we examined several parameters including expression level, pH dependence and potential electrostatic interaction between PBP-A and the peptidoglycan. We found that reducing the expression level of PBP-A, lowering its isoelectric point and avoiding acidic pH conditions, significantly improves the fitness cost. We further investigated the introduction of a disulfide bridge, which is conserved in many class A serine β-lactamases, in PBP-A. Fitness improvement of disulfide-bonded PBP-A was significant and analysis of peptidoglycan indicated that the interfering activity on peptidoglycan was significantly reduced. Further *in vitro* characterization of the enzymes shows a severe impact of the

disulfide bridge on the DD-peptidase activity, without compromising the reactivity against penicillin. Based on these results we propose that disulfide bonds played an intragenic (positive) epistatic role, along evolutionary trajectories from PBPs to β-lactamase, to secure the enzyme against deleterious effects of gaining the hydrolytic power, highlighting an overlooked role of disulfide bonds to impact the structure-function relationship of proteins by constraining the conformational dynamics.

To benefit from the single-copy chromosomal expression in the directed evolution approach, we developed a novel scarless method for building chromosomal libraries in genome of *E. coli,* based on a combination of CRISPR-Cas9 cleavage and λ-Red recombination system. This method allows generation of chromosomal libraries up to million independent clones.

We then subjected different variants of PBP-A to directed evolution under the fitness-improving parameters. Although all the attempts to convert PBP-A to a beta-lactamase were not successful, a slightly higher resistance was achievable from the libraries of disulfide-bonded PBP-A, that was strangely not stable upon re-growing of the resistant population.

The results of this thesis clearly demonstrate that fitness cost issues are multi-factorial and that an initial careful attention to every biological parameter may be crucial for the success of a directed evolution campaign, especially when dealing with potentially dangerous enzymatic activity such as a hydrolase.

Table of Contents

Chapter 1

Introduction

1.1 Foreword

Although we have a vast knowledge regarding enzymes and their chemistry, we have still a rather limited understanding about the structure-to-function relationships underlying their exquisite catalytic performances, as well as the mechanisms of their evolutionary emergence. Rational engineering of enzymes to improve or alter their activity in a desired way is not as easy as it seems because a very small change in chemical or physical properties can have a profound impact on the catalytic activity and such impacts are still poorly predictable. Of course, nature does not proceed by rational engineering but through evolutionary trajectories and adaptive mechanisms. Natural evolution is a very slow phenomenon based on a continuous process of random mutation in DNA codes and selection of the most adapted phenotypes in a particular environment. Once an enzyme is positively selected under a certain selection pressure, its beneficial advantage will assure its heritability in upcoming generations. However, despite high precision in their activities, many enzyme perform side-reactions that may prove problematic in crowded cellular pool, which has led to evolution of repair enzymes to cope with such noncanonical activities. Discoveries are increasing on metabolite repair enzymes, which play role in protection of metabolic pathways from undesired interference caused by promiscuous enzymes (Bommer, Van Schaftingen, & Veiga-da-Cunha, 2020; Veiga-da-Cunha, Van Schaftingen, & Bommer, 2020). On the other hand, the emergence of a new activity can also be

associated with some fitness costs due to potential increase in interfering promiscuous activities. So, it is more the difference between benefit and cost that is driving the selection. For non-essential enzymes in particular, this can be, in theory, a major hurdle to evolution since their activity may be beneficial only under specific environmental conditions and there should not be any remaining fitness cost when the activity is not necessary. In this context, compensatory adaptation (positive epistasis) in organism level and/or regulatory rewiring are inevitable to insure beneficial or neutral impact in a variable environment (Domínguez-García, García, Quesada, & Caballero, 2019; McDonald & Ayala, 1978; Millan et al., 2014; Poon & Chao, 2005). However, at the enzyme level itself, little is known about the relative importance of molecular features that are selected for minimizing fitness costs rather than maximizing the beneficial activity in a particular evolutionary trajectory. While enzymes are usually described as optimized by evolution to perform a specific function with very high efficiency, their evolution for not doing what they should not do is probably overlooked.

Directed evolution of enzymes is a modern field of research in biochemistry and molecular biology aiming at generating biocatalysts endowed with new or improved properties by mimicking natural evolution in an accelerated manner. In the broader scope of proteins, it deserved recognition by the Nobel Prize in Chemistry in 2018 (Arnold, 2019). From a fundamental point of view, studying evolved enzymes can contribute to our understandings of structure-function relationships as well as of the evolutionary mechanisms underlying the emergence, optimization and adaptation of enzymatic catalysis. From an applied point of view, engineered enzymes are widely used in a variety of medical, biotechnological and industrial applications and a huge potential is still to be exploited.

A family of cyanobacterial DD-peptidase called PBP-A, which is phylogenetically the closest family to class A β-lactamases, was previously chosen in our team as the starting point of evolutionary study to understand the mechanisms of evolutionary emergence of β-lactamase activities. Through several attempts of

directed evolution, evidences gathered that converting a DD-peptidase into a β-lactamase follows a counter-intuitive evolutionary trajectory. Despite the high similarities in structural scaffolds and conserved active site motifs between PBP-As and class A β-lactamases, directed evolution attempts has failed to convert PBP-A into a β-lactamase. Evidences suggested that fitness cost(s) associated with expression of PBP-A and its variants in *Escherichia coli (E. coli)* was(were) the main obstacle in this conversion (Labarbe, 2006; Trabelsi, 2014). In this thesis, we aim to (i) thoroughly characterize the origins of this(these) fitness cost(s), (ii) evaluate potential strategies towards minimizing them and (iii) finally, to pursue the directed evolution towards a β-lactamase under fitness improving conditions. Despite the established notion that evolvability of an enzyme is positively correlated with enzyme promiscuity, in this study, we highlight that role of promiscuity in evolvability of enzymes could rather be highly context-dependent: firstly, if the promiscuous activity(s) (multifunctionality) is not disturbing the existing cellular networks of biochemical pathways, it would be granted freedom for further developments towards specific functions, otherwise the intrinsic selection of maintaining cellular fitness would restrict the freedom and hinder its further evolution; secondly, more promiscuity, particularly substrate promiscuity, means more competition of molecules for the active site, which can cause competitive inhibition and prevent binding of new substrate(s). We suggest that, in the course of β-lactamase evolution, the fitness cost might have decreased through modification of the enzyme's isoelectric point and reduction of active site dynamics through disulfide bonds. Moreover, the loss of substrate promiscuity was probably facilitated through drastic changes in active site cavity via a longer Ω-loop.

1.2 Molecular Evolution

1.2.1 Random mutations as source of variation

The classical dogma of evolution is known as a slow and repetitive process of random mutations in genes and selection of the advantageous phenotypes. Survival and heritability of selected phenotypes depends on their fitness and

cellular/organismal adaptability (Mousseau & Roff, 1987; Wilson et al., 2006), especially in absence of the selection pressure. Three types of point mutations in nucleotide sequence of a gene are: synonymous, if the mRNA codon codes for the same amino acid; missense, if the mRNA codon codes for a different amino acid; or nonsense, if the mRNA codon becomes a stop codon. Insertion and deletions (InDels) can modify the length of the polypeptide or cause frameshift in open reading frame (ORF) if the number of inserted or deleted base pairs is not divisible by three. Although they were initially considered as silent, synonymous mutations might affect transcription and/or translation and, therefore, impact the evolutionary trajectory of genes (Chamary, Parmley, & Hurst, 2006; Firnberg, Labonte, Gray, & Ostermeier, 2014; Goymer, 2007; Kimchi-Sarfaty et al., 2007; Plotkin & Kudla, 2011). Nonsense mutations lead to protein truncation by introducing earlier nonsense stop codons prior to the natural one.

In neutral theory of molecular evolution, mutations can become beneficial, neutral or deleterious if the corresponding phenotype has a selective advantage, no effect or a disadvantage, respectively. Depending on the level of fitness cost that deleterious mutations introduce to the host cell, they can be lethal or disadvantageous. Lethal mutations would never be seen in a population since they always get purged due to causing high fitness cost, and hence purifying selection, while disadvantageous (not lethal) mutations may become heritable if their fitness effect is tolerable by the cells (see Figure 1-1) (Eyre-Walker & Keightley, 2007; Loewe & Charlesworth, 2006).

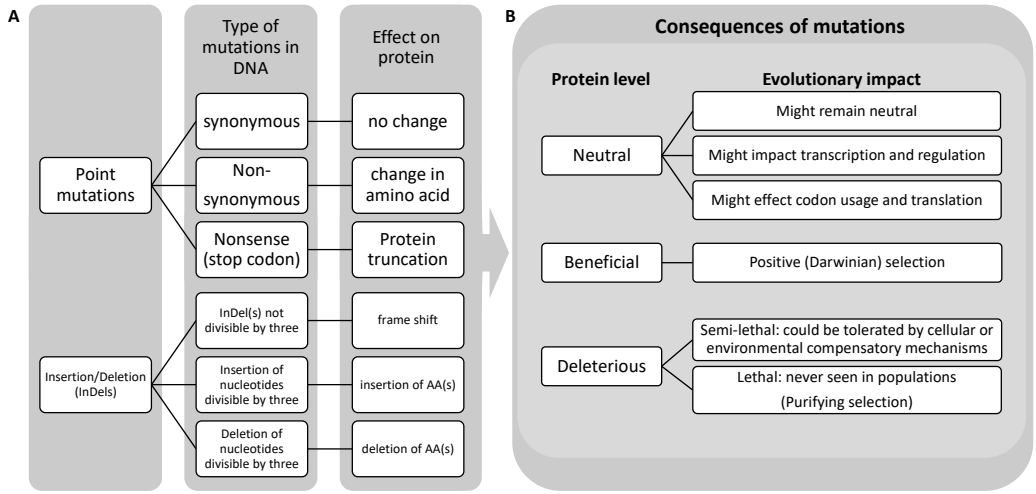


Figure 1-1 : Different mutations in DNA level and prediction of their consequences on protein as well as protein evolution. A) Point mutations in nucleotide sequence of DNA will lead to either synonymous, nonsynonymous or nonsense mutations which can cause no effect, amino acid swapping or truncation in protein level, respectively. Effect of insertion and deletions (InDels) is dependent on number of InDel nucleotide(s): if divisible by three it would insert/delete amino acid(s) and if not, it would cause frameshift mutation, which will lead to disruption of reading frames and incorrect reading translation of the DNA codes. B) Depending on effect of mutation on protein level, it can face different consequences from neutral, beneficial or deleterious. Neutral mutations on protein level such as synonymous mutations might remain neutral or not. Synonymous mutations can be beneficial or harmful to cellular/organismal fitness through their effect on transcription and translation in DNA level (Chamary et al., 2006; Firnberg et al., 2014; Goymer, 2007; Kimchi-Sarfaty et al., 2007; Plotkin & Kudla, 2011). Beneficial mutations, which provide improvement in fitness of cells/organisms in particular environments, become fixed in population as long as selection pressure is positively selecting them. Deleterious mutations may be compensated through compensatory adaptations or face purifying selection and purging from population(s).

1.2.2 Protein's sequence space and fitness landscape

In evolutionary biology, scientists are using the concept of 'sequence space' as a way to represent the range of possible combinations of building blocks in a given entity in central dogma (DNA, RNA or protein). The number of known sequences of proteins is around 10^8 while the extant number of protein sequences has been estimated between 10^{34} and 10^{43} sequences that 4 gigayears of evolution might have explored (Buchholz, Fademrecht, & Pleiss, 2017). Sequence space of proteins represents all possible amino acids in a polypeptide chain and because there is possibility of 20 amino acids to occupy each position, the sequence space is very large. An average-sized protein with 300 amino acid would have a sequence space with 20^{300} possible combination of amino acids (Romero & Arnold, 2009b). Therefore, all the proteins over entire history of life, have explored a tiny fraction of sequence space, which is expanding in a very slow pace by natural evolution (Arnold, 2006; Dryden, Thomson, & White, 2008; Maynard Smith, 1970). It was initially a main question in the field, as asked by J. Maynard Smith (1970), that whether natural selection took a single route in the limitless desert of sequence space (Axe, 2004; Dryden et al., 2008), which was limited towards survival routes and guided by natural selection or not? Or has nature fully explored the sequence space for all possible folded protein structures? However, recent studies on de novo protein design, based on computational modeling and (lowest) free energy levels, has opened an entirely new window of protein structures, functions and assemblies and has thought us that nature has only explored a very small subset of possible routes for protein evolution (Baker, 2019; Boyken et al., 2016; Brunette et al., 2015). Even outside of the chemical domain explored by natural proteins, beyond the 20 naturally occurring amino acids, design of proteins using unnatural amino acids or genetic code expansion is also exploding the explorable sequence/structure space (Hosseinzadeh et al., 2017; Neumann-Staubitz & Neumann, 2016).

Protein's properties are defined by its amino-acid sequence, which is decoded from the DNA sequence of their corresponding genes. Besides functional characteristics, proteins have their unique biochemical and biophysical

properties, such as electrical charge, stability, tendency for aggregation, degradation rate, sub-cellular location and post-translational modification in the crowded cellular environment. Such properties may play as important role as protein's function for cellular and organismal fitness. These properties can be easily perturbed by nonsynonymous (missense) mutations in proteins primary sequence and would become problematic if the mutations have serious consequences against cellular fitness.a Such costly mutations are often compensated by adaptive mutations and/or compensatory mechanisms (DePristo, Weinreich, & Hartl, 2005).

The metaphor of fitness landscape (or adaptive landscapes) is widely used in evolutionary biology to explain the multidimensional dynamics of adaptation and to visualize the relationship between genotypes and their fitness in a genotype space (Figure 1-2). Coordinates in the landscape represent genotypes space while the height of the landscape representing the fitness related to each coordinate. Similar genotypes are located close to each other, while getting more different as they are locating more far from each other (de Visser & Krug, 2014). The environment that a protein is evolving in determines the shape of the fitness landscape meaning that the shape of landscape and evolutionary trajectories of the same protein is different under different environment. A multi peaks/valleys fitness landscape is called 'rugged', where an evolutionary trajectory may become trapped in local peaks while a landscape with a single smooth peak is called 'smooth' fitness landscape, where increasing in fitness through all the trajectories would end the same optimum (global maximum) (Figure 1-2).

In smooth landscape, all beneficial mutations are additive, their acquisition in any order would lead to an uphill trajectory. Contrarily, in a rugged landscape, mutations interact with one another by epistasis, thus the effect on each mutant is dependent on the pre-existing genetic background, and a specific mutation can be beneficial or deleterious depending on this background.

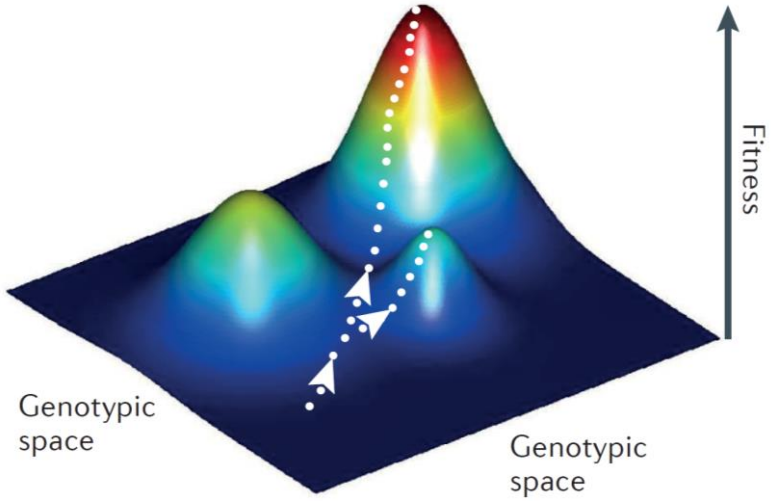


Figure 1-2: Fitness landscapes are often conceptualized as a three-dimensional mountainous landscape. Altitudes in z axis of this landscape are representing the fitness, while different coordinates in the x-y plane are representative of genotypes in the sequence space of a protein. A fitness landscape is called 'rugged', if it contains many local peaks separated by deep valleys, where an evolutionary trajectory may become trapped in local optimum. Average interactions between mutants (epistasis) in a rugged landscape is increased. On the other hand, a 'smooth' fitness landscape contains a single smooth peak see also (Figure 1-8-b,c). Series of small and random genetic changes in the genotype, causes random change in coordinate and depending on the starting coordinate in the landscape, an evolving population climbs uphill in the fitness landscape, by help of a selection pressure for higher fitness (figure adapted from: de Visser & Krug, 2014; see also Romero & Arnold, 2009).

1.2.3 Directed (artificial) evolution

As mentioned above, most of protein sequence space has not been explored by natural evolution, which is a very slow process. Scientists have been harnessing the power of natural evolution in the laboratory (Arnold, 1996; K Chen & Arnold, 1993; Keqin Chen & Arnold, 1991; MacBeath, Kast, & Hilvert, 1998) to accelerate the exploration of the sequence space to finally unlock boundless potentials by engineering proteins and biological systems to address the growing challenges facing humankind. Directed evolution is an accelerated mimic of natural evolution that consists of iterative rounds of mutations in nucleotide sequence of proteins followed by artificial selection or screening for a desired property (Figure 1-3). The field has been awarded with the Nobel Prize in Chemistry in 2018, mainly because of its major impacts in the development of

therapeutic antibodies and enzymes with tailormade biocatalytic properties. In recent decades, applications of directed evolution has broaden not only in engineering and optimization of proteins' properties but also in understanding and development of other biological components such as nucleic acids, biomolecular interactions, regulatory elements, biosynthetic pathways and genetic circuits (Arnold, 1998, 2018; Jäckel, Kast, & Hilvert, 2008; Kuchner & Arnold, 1997; Romero & Arnold, 2009b; Yokobayashi, Weiss, & Arnold, 2002; Zeymer & Hilvert, 2018).

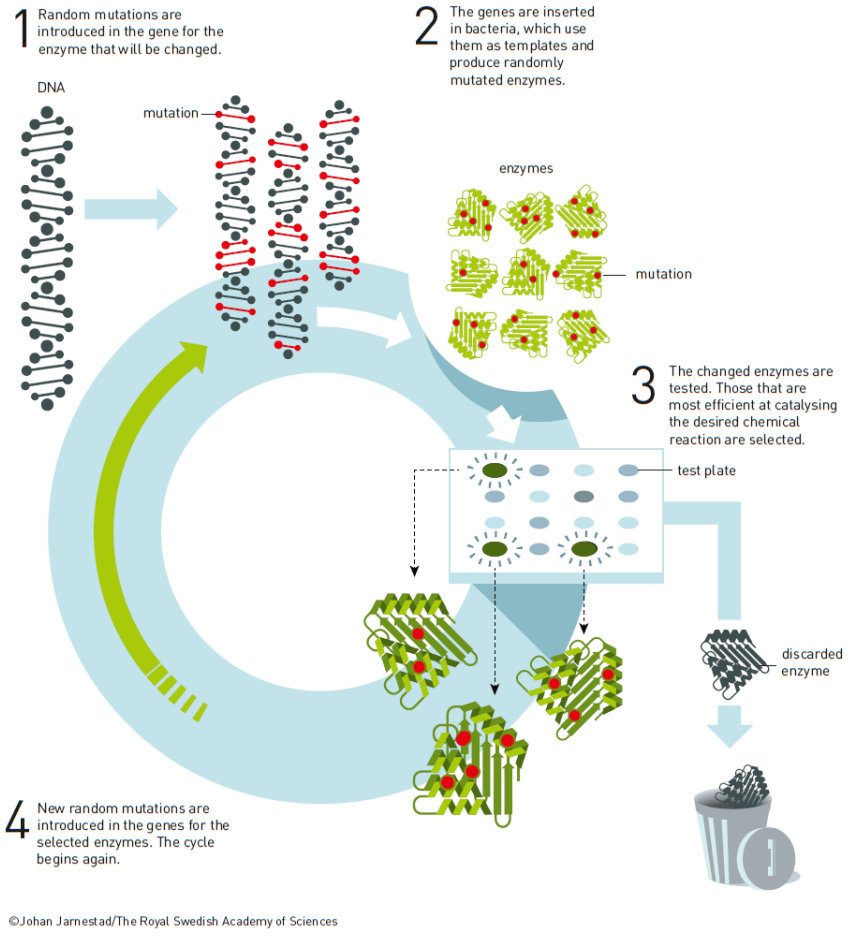


Figure 1-3: Directed evolution of enzymes. **1)** Artificial genetic diversity is introduced randomly in the gene of interest; **2)** expression of genes; **3)** selection of enzymes based on a desired activity; **4)** rounds of directed evolution are repeated to achieve the

maximum values on activity. (image was taken from © The Royal Swedish Academy of Sciences, https://www.nobelprize.org/).

1.2.3.1 Artificial genetic diversification

Although point substitutions are the typical source of generating genetic diversity in directed evolution campaigns, innovative methodologies are increasing to apply InDels, repeats and recombination techniques in order to create gene libraries. Since a very small alteration in chemical or physical properties of enzymes can have a profound impact on their catalysis, controlling the distribution (quality) of mutations across the protein and number (quantity) of mutations at a time and/or the nature of these mutations are very important criteria for designing a starting library (Gargiulo & Soumillion, 2021).

1.2.3.1.1 Random point mutagenesis

The easiest-to-handle and well-established strategies to introduce random mutagenesis on DNA sequence are relying on point mutations. Depending on the length of the target sequence, various methodologies have been developed:

(i) *Error-prone PCR (epPCR):* to cover random mutations on a full-length gene, errors in DNA replication, using low-fidelity DNA polymerase, is the common approach;

(ii) *Degenerate oligos:* for targeting a short window, incorporation of degenerate oligo nucleotides is widely used. Although NNK codon degeneracy (32 codons/20 aa) encoding all 20 canonical amino acids covers the entire sequence space of a protein/peptide, limitation in screening of such astronomical size libraries is the main practical disadvantage. Therefore, controlling codon degeneracy to construct small and smart libraries, which require less screening efforts, has been developed. In later case, reduced amino acid alphabets such as NDT codon (12 codons/12 aa) encoding 12 representative amino acids (Phe, Leu, Ile, Val, Tyr, His, Asn, Asp, Cys, Arg, Ser, Gly) was shown to improve design and screening of focused libraries (Li, Qu, Sun, & Reetz, 2019; Reetz, Kahakeaw, & Lohmer, 2008).

(iii) *DNA shuffling:* genetic libraries constructed by DNA shuffling involve recombination between homologous DNA fragments (Stemmer, 1994). This method combines fragmentation of DNA by DNase I, from 10 to 50 base pairs, and reassembly of fragments via PCR. The initial pool of DNA is either diversified by mixing homologous genes with high sequence identity or using epPCR random mutagenesis.

(iv) *Massive parallel oligo synthesis and assembly:* unlike the very low cost of DNA sequencing, the cost of DNA synthesis is still high for routine application in synthetic biology. However, there is an increasing rate in development of low-cost methods to produce and assemble high-quality synthetic DNA that provide ready for use synthetic DNA constructs. Microarray-based massive oligonucleotide synthesis in multiplexed manner is one of the pioneering technologies in these cases (Hughes & Ellington, 2017; Klein et al., 2016). Nevertheless, limitation in the length of the synthetic constructs and their error-free assembly are still remaining challenges in these approaches.

Furthermore, errors in DNA polymerization can either be applied to create *in vitro* genetic libraries via ep-PCR or to diversify a gene in its *in vivo* context by guiding the low-fidelity DNA polymerase to the target gene in a plasmid or genomic loci.

1.2.3.1.2 InDels libraries

Learnt from natural evolution and aiming further explore the sequence space of a protein beyond the point mutations, there has been a growing trend on development of methods for construction of genetic diversity via InDel mutations (Shortle & Sondek, 1995). However, compare to point mutations, InDels come with a higher cost, which lead to purifying selection against them. The ratio of point mutations to InDels for protein coding sequences in genome of primates is estimated 5:1 and 20:1 for bacteria (Schenkmayerova et al., 2021). Several innovative methods employing different strategies has been introduced

for random InDel library creation: such as insertion/deletion of few nucleotides via synthetic oligonucleotides (Lamminmäki et al., 1999; Qi & Scholthof, 2008), gene-shuffling based approach (Fujii, Kitaoka, & Hayashi, 2006), transposon-based mutagenesis approach (Emond et al., 2020; Skamaki et al., 2020) and cycles of restriction digestion and ligation of building blocks (Tizei, Harris, Withanage, Renders, & Pinheiro, 2021).

1.2.3.2 Library assembly and expression

Incorporation of the diversified oligonucleotides into a gene library is generally carried out via restriction/ligation or homologous recombination methods (overlap PCR or Gibson assembly). Recent developments show advancing on methodologies for assembly of genetic libraries by combining steps of the diversification and assembly. QuickLib and Golden Mutagenesis are among the fast and efficient methods for incorporation and assembly of short degenerated libraries into a single site in a plasmid or multiple site of a DNA fragment, respectively (Galka, Jamez, Joachim, & Soumillion, 2017; Püllmann et al., 2019). Even efficient methods for construction and expression of gene libraries into chromosome of bacteria has been developed (Dorrazehi et al., 2020).

Expression strategy of libraries is mostly dependent on the selection/screening strategy. For *in vivo* selection/screening, an appropriate subcellular localization of either expression in cytoplasm or secretion into periplasm (of Gram-negative hosts) via a signal sequence, is important. On the other hand, for *in vitro* screening of the libraries variants can either be displayed at cell surface, expressed in extracellular environment or expressed in a cell-free system, see Figure 1-6 (Gargiulo & Soumillion, 2021).

1.2.3.3 Library assessment

Selection and screening are the two approaches for protein library analysis. Selection refers to spontaneous elimination of non-functional variants based on autonomous feedback loops, linking phenotype to genotype replication, whereas screening evaluates every protein for a desired function/property.

1.2.3.3.1 Selection

Since negative variants are rejected in selection and only positive variants are subjected to the next round of directed evolution, analysis of large libraries (more than 10^{11}) are possible, which provide an intrinsically high throughput library assessment. Different methods for high throughput selection have been developed for assessment of enzyme libraries that the most widely used are as follows:

Growth complementation and reporter-based selection: in these approaches selection is carried out *in vivo*, where survival of the host cell is linked to the desired property of protein of interest in such a way that only the cells expressing the desired variants are able to survive under the selective pressure. In growth complementation, usually an auxotroph strain is used that only grows when an essential compound is provided or metabolically synthetized. Desired activity of protein of interest is linked to production of the essential compound and the election is carried out in absence of the essential metabolite, where only cells that carry an active variant will grow. Under reporter-based selection, the desired activity of protein of interest confers a reporter/survival phenotype, such as antibiotic resistance. The activity of the enzyme of interest controls the reporter activity and desired variants can be selected based on level of reporter activity (van Rossum, Kengen, & van der Oost, 2013; Xiao, Bao, & Zhao, 2015). Directed evolution of antibiotic resistance or coupling activities of proteins to resistance phenotype are well-established examples of using reporter-based selection in enzyme evolution.

Compartmentalized self-replication (CSR): unlike other high throughput technologies that rely on activity screening and require the detection and sorting of single variants, CSR provides a programmable *in vitro* feedback loop to link a target enzyme activity to its corresponding gene amplification. In a recently reported Programmable External Networkbased CSR (PEN CSR), library of plasmids are transformed and expressed in bacteria followed by compartmentalization in monodisperse water-in-oil droplets, together with a molecular program, DNA polymerase and a lysing buffer. Individual cell will be lysed in the droplets and

the supplemented molecular program assays the level of target enzyme activity and accordingly lead to production of primers for the encoding gene. PCR cycles on the emulsion will lead to amplification of active gene variants that triggered the production of sufficient primers. Finally, active variants will be identified based on their proportion in the pooled gene population (see) (Dramé-Maigné, Espada, McCallum, Sieskind, & Rondelez, 2021; Dramé-Maigné, Zadorin, Golovkova, & Rondelez, 2020; Ghadessy, Ong, & Holliger, 2001).

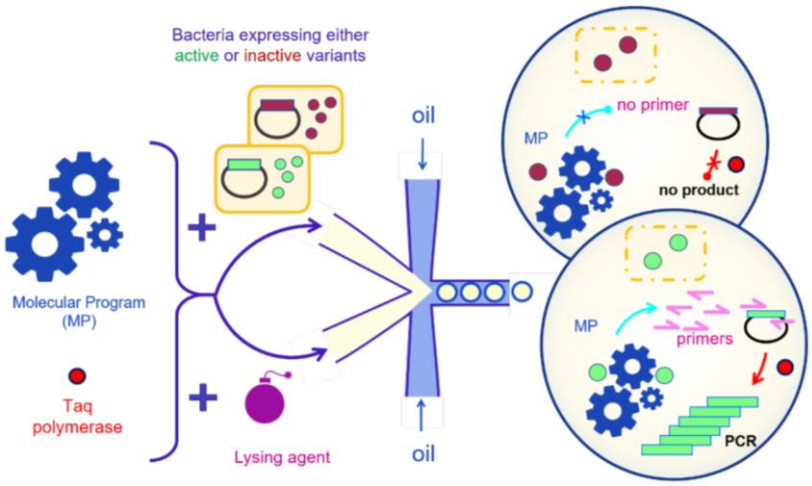


Figure 1-4: Overview of the PEN CSR concept (Dramé-Maigné, Espada, McCallum, Sieskind, & Rondelez, 2021).

Display systems: display technologies that involves direct physical linkage of genotype with phenotype make protein library accessible to the external environment and thus provide a high throughput *in vitro* selection to look for desired properties of proteins. The corresponding genes are linked to the selected protein variants that can easily be isolated and identified.

(i) *mRNA display:* the C-terminal of proteins are covalently linked to their *in vitro* transcribed encoding mRNA via puromycin, a peptidyl acceptor antibiotic that is linked at 3' end of the *in vitro* transcribed mRNA (see Figure 1-5) (Takahashi & Roberts, 2009).

(ii) *Ribosome display:* based on a spacer sequence at 3' end of mRNA, which lacks a stop codon preventing release of translated protein

and keeping associated with the peptidyl tRNA in the ribosome channel, the translated protein is connected to the mRNA and ribosome. The stable complex of mRNA, protein and ribosome provide a cell-free system for *in vitro* selection of proteins and peptides from large libraries (Hanes, Schaffitzel, Knappik, & Plückthun, 2000; He & Taussig, 2002; Plückthun, 2012). Ribosome display is mainly used for selection of binding proteins and antibodies. Similar to mRNA display, the genetic library is transcribed *in-vitro* (Figure 1-5).

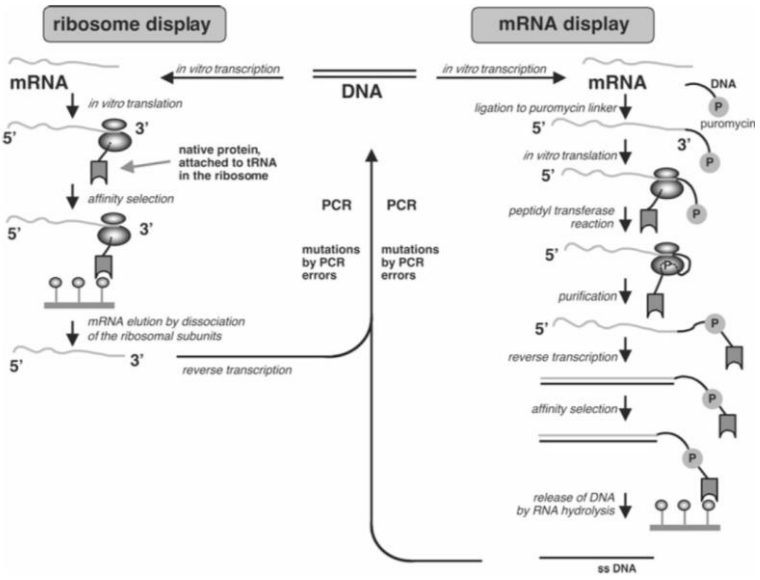


Figure 1-5: Overview of ribosome display and mRNA display systems (figure taken from Plückthun, 2012).

(iii) *SNAP display:* unlike mRNA and ribosome display, SNAP technique is developed to use DNA as the coding nucleic acid to increase stability. In SNAP, protein of interest is expressed in fusion with SNAP-tag (O^6-alkylguanine-DNA-alkyltransferase), which can be attached to template DNA via a thioester bond with the substrate (O^6-benzylguanine) that is covalently bonded to the DNA. SNAP display is widely used for entirely *in vitro* selection of protein binders in emulsion droplets (Houlihan, Gatti-Lafranconi, Kaltenbach,

Lowe, & Hollfelder, 2014; Stein, Sielaff, Johnsson, & Hollfelder, 2007). Alternative to monomers SNAP, SNAP dendrimers are developed for multivalent protein display on a dendrimer-like DNA (Kaltenbach, Stein, & Hollfelder, 2011).

(iv) *Plasmid display:* based on a DNA binding protein fused to the protein of interest, desired variants of proteins, linked to their corresponding plasmids, are selected from the cell lysate of large libraries (Choi, Pack, & Yoo, 2005; Speight, Hart, Sutherland, & Blackburn, 2001).

(v) *Phage Display:* as the most commonly used display techniques for *in vitro* selection of protein binders, phage display rely on ability of phages (such as M13 or T7) to infect bacterial hosts without killing them. The protein, peptide or antibody of interest is fused to one of the coat proteins of the phage and displayed on the surface of the phage particle when the phage is assembled in the host cell. The selection process is entirely conducted *in vitro* via specific interaction of the phage-displayed protein with an immobilized ligand. The corresponding DNA is retained in the phage particle that can be recovered by re-infection and amplification after selection (Dai et al., 2008; G. P. Smith & Petrenko, 1997; Soumillion & Fastrez, 2002; Urvoas, Minard, & Soumillion, 2020). To fulfil the need of eukaryotic proteins for correct folding and post-translational modifications, fully mammalian platform, named retroviral display was also developed (Granieri, Baret, Griffiths, & Merten, 2010; Urban & Merten, 2011; Xiao et al., 2015).

(vi) *Cell surface display:* based on fusing the gene of interest to a gene encoding for cell-surface anchor protein, polypeptides or proteins can be displayed on the surface of living cells, such as the outer membrane of bacterial cells, which can be analyzed and sorted for desired properties (S. Y. Lee, Choi, & Xu, 2003).

1.2.3.3.2 Screening

Assessment of individual protein variants in large libraries is the purpose of screening, which provides the chance of evaluation for all the variants. Coupled with automation, screening methods are well developed for ultrahigh throughput screening processes. Screening methods exist in a range of formats, from solid-phase to micro-, nano- or picoliter aqueous formats depending on type of the expression and detection assay as well as throughput level of the assay (Figure 1-6):

(i) *Microtiter Plates:* microtiter plates miniaturize test tubes to minimize sample volume and increase number of samples by multiple wells. Enzyme activity from either crude cell extracts or purified proteins can be assessed in a microtiter plate. Colorimetric or fluorometric assays, which can measure decrease of substrate or increase of product are the common assays for screening enzyme activities using microtiter plate. Quantification of substrate/product by macroscopic identification and monitoring UV–vis absorbance or fluorescence signal is carried out on a plate reader, which are mostly equipped with agitation and temperature control.

(ii) *Fluorescence-activated cell sorting (FACS):* Individual cells can be sorted based on intensity of their fluorescent signal. Cell surface display (see previous section), can be subjected to sorting by FACS upon the intensity of a fluorescent signal coupled with the activity of protein of interest.

(iii) *In vitro compartmentalization (IVTC):* by mimicking the cellular compartment, which holds together all components of genotype to phenotype dogma, man-made cell-like compartments, such as water-in-oil emulsion droplets or water-in-oil-in-water double emulsion droplets, has been developed to isolate individual genes in an aqueous compartments to form independent reactors for cell-free protein expression and enzyme reaction (Fallah-Araghi, Baret, Ryckelynck, & Griffiths, 2012; Tawfik & Griffiths, 1998). Field of

IVTC is rapidly growing, in particular, by coupling with microfluidics, highly monodisperse droplets can be generated and screened in an fully automated and highly controlled manner. The volume of droplets can be as down as picoliter that can facilitate massive scale-up of screening throughput (Holstein, Gylstorff, & Hollfelder, 2021; Kintses et al., 2012). Depending on the assay and enzyme of interest, screening/sorting of library in microdroplets is carried out as follows:

a. *Fluorescence-activated droplet sorting (FADS):* individual compartments of water-in-oil microdroplets can be sorted based on the intensity of fluorescent signal generated from the encapsulated cell or reaction (Baret et al., 2009). Because of high sensitivity and specificity, FADS is the most widely used method is droplet sorting.

b. *Absorbance-activated droplet sorting (AADS):* unlike FACS and FADS that rely on detection of a fluorescent product, AADS is developed for ultrahigh-throughput microdroplet sorting based on absorbance detection (Gielen et al., 2016). Although it is label-free, AADS is less specific and less sensitive compare to FACS and FADS.

c. *Mass activated droplet sorting (MADS):* based on electrospray ionization mass spectrometry (ESI-MS), MADS has been developed to sort nanoliter droplets with high throughput and high accuracy. A sorting algorithm enables MS-based sample identification and dynamic thresholding and therefore, provides a label-free screening (Holland-Moritz et al., 2020).

(iv) *Digital Imaging (DI):* relying on colorimetric activity assays coupled with digital imaging spectroscopy, DI allows solid-phase high throughput screening of microbial colonies (Delagrave et al., 2001).

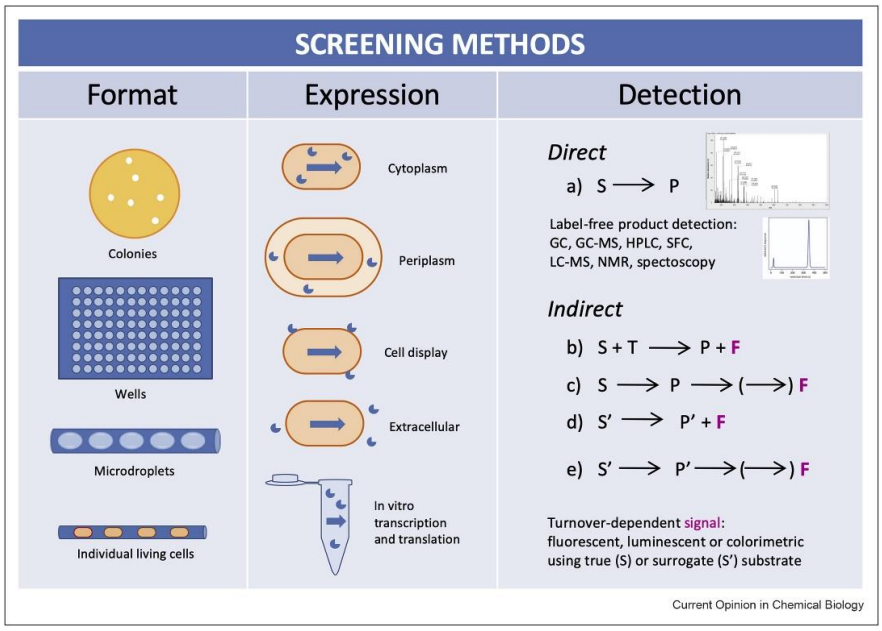


Figure 1-6: Schematic representation of different formats and expression strategies for directed evolution as well as methods for screening assays; a) direct screening assays benefit from direct measurements of substrates or products using spectroscopy or analytical chemistry. On the other hand, indirect assays benefit from: b) detection of co-substrates or co-products; c) coupling the main reaction with a secondary reaction step; d) utilization of a surrogate substrate to produce a detectable side-product or e) initiating the final product of a cascade by a surrogate product (Gargiulo & Soumillion, 2021).

1.2.3.4 In vivo continuous directed evolution

In vivo targeted genetic diversification is mainly relied on applying a low-fidelity DNA polymerase to introduce errors in a target gene located in a plasmid or genomic loci. (Badran & Liu, 2015; Esvelt, Carlson, & Liu, 2011; Halperin et al., 2018; Rix & Liu, 2021; Shinkai & Loeb, 2001; Zhong et al., 2020). *In vivo* continuous directed evolution is aiming to minimize the need for human intervention and remove the need for iterative rounds of mutagenesis and selection/screening cycles. A pioneering system was developed by Esvelt et al. (2011), which exerts a flow of continuous culture of *E. coli* host cells, which are infected by M13 bacteriophage followed by selection based on linking the activity of desired protein to the production of infectious progeny phage carrying the gene of interest (Esvelt et al., 2011). This system was called PACE (phage-assisted continuous evolution), which is illustrated in Figure 1-7.

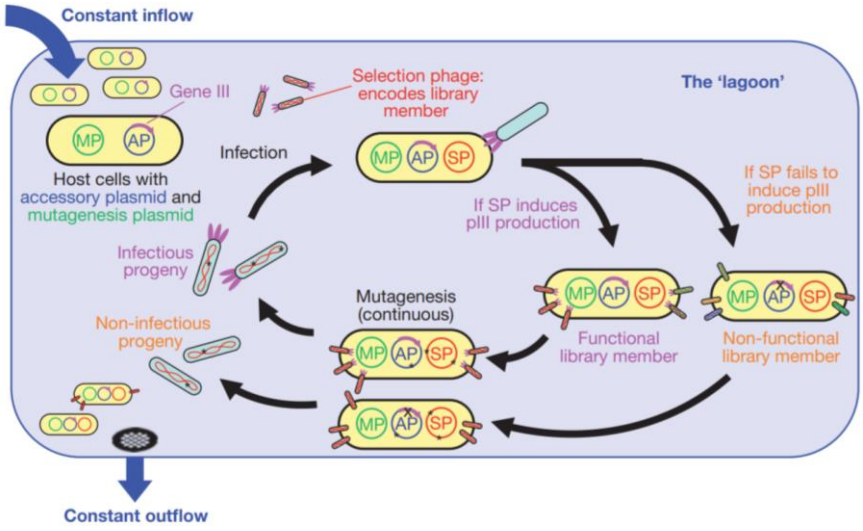


Figure 1-7: Schematic overview of the PACE system in a fixed-volume lagoon. Encoding libraries are continousely evolved by infection of constantly inflowing host cells with selection phage (SP), where only functional variants induce the expression of phage infection requires protein III (pIII) from the accessory plasmid (AP) and release of progeny, thus leading to infection of new host cells. Level of mutations is elevated by inducing the mutagenesis plasmid (MP) that encodes error-prone DNA polymerases (figure taken from Esvelt et al., 2011).

Another recently developed system, called pEvolvR, takes advantage of a RNA-guided nickase variant of Cas9 fused to an error-prone PolI polymerase to specifically diversify a target region in the genome of *E. coli* that can be used for continuous evolution (Halperin et al., 2018). However, this system seems to suffer from off-target effect of the Cas9 system to target unspecific region(s), short coverage of target region (a narrow window up to 350 bp), and bias of mutation rate depending on distance from the nickase cleavage site. Another recent approach for *in vivo* continuous directed evolution in yeast, couples an orthogonal DNA polymerase−plasmid pair, called OrthoRep (Ravikumar, Arrieta, & Liu, 2014; Ravikumar, Arzumanyan, Obadi, Javanpour, & Liu, 2018), with a versatile platform of continuous yeast culture controlled for growth condition as well as functional selection conditions to gradually guide the libraries to follow the desired evolutionary trajectory (Park & Kim, 2021; Rix et al., 2020). Although these *in vivo* approaches offer speed, depth and scalability for directed evolution campaigns, incompatibility of libraries (or individual variants) with host

cell and limit on library size are the main disadvantage of these approaches. However, to bypass this limitation, in the complex context of a living cells that contains networks of metabolic pathways and pressured by homeostasis, such systems must be coupled to a selectable phenotype. In this regard, PACE has an advantage because the cells are sacrificed, and only the phages continue propagation and fresh cells are continuously fed into the lagoon to support the propagation of the phages. Risk of contamination of host cells during self-replication as well as requirement for genetic vectors are among other general limitations of *in vivo* approaches in directed evolution (Tinafar, Jaenes, & Pardee, 2019).

1.2.3.5 In vitro (cell-free) directed evolution

Particularly, if we are aiming to explore sequence space beyond the biological routes that life on earth has been following, *in vivo* diversification and selection would not allow us to reach and scope so. Therefore, cell-free systems have been expanding to circumvent such constraints by developing *in vitro* diversification of a gene of interest (GOI), expression of the respective proteins and, finally, selection or screening of the desired activity/phenotype, which is opening a bright perspective in the field. Cell-free systems, can be assumed as programmable minimal liquids, which have been designed to use minimum complexities of biological machineries, often extracted from cells, to facilitate *in vitro* protein expression from DNA templates.

The *in vitro* approaches for directed evolution are mainly based on compartmentalization of transcription/translation system and gene of interest to be transcribed/expressed, where both of the genotype and phenotype (RNA, peptide or protein of interest) are enclosed in the same compartment (Dodevski, Markou, & Sarkar, 2015; Holstein et al., 2021; Tinafar et al., 2019). However, compartmentalization of a single DNA molecule lead to very low yield of protein expression requiring DNA amplification steps to increases the number of expression templates. Coordination of such *in vitro* DNA amplification, protein expression and enzymatic reaction is tricky due to sensitivity and stability of

droplets against PCR or isothermal amplification. This constraint can be circumvented by utilizing rolling circle amplification (RCA), which is easier to handle in droplets compared to PCR, and stepwise addition of the components into droplets on microfluidic chips (Holstein et al., 2021). Nevertheless, partially *in vitro* approaches such as display technologies, which provide cell-free selection/screening steps, circumvent the need for compartmentalization.

Among many advantages of entirely cell-free approaches, we can refer to construction of large libraries, ability to express toxic proteins, reduced experimental time, utilization of toxic compound for screening/selection and fewer of biases caused by biological system. However, technical limitations such as low yield of protein expression and/or incompatibility of transcription/translation with replication beside the dependency of proteins/enzymes to biological system to properly fold and function, e.g., requiring post-translational modifications such as chaperonin systems for their proper folding, subcellular localization or unknown binding/interactions with other proteins or cellular components and pathways, are among disadvantages of *in vitro* directed evolution (Contreras-Llano & Tan, 2018; Dower & Mattheakis, 2002).

1.2.4 Evolvability

Evolvability refers to the capability of a biological system to generate, adapt and inherit novel phenotypes (Hu & Banzhaf, 2018; Payne & Wagner, 2019). Beside the spontaneous mutations that happen consistently, in response to unfavorable challenges and stress conditions, biological systems generate phenotypic and genetic diversity enhancing their chance of survival (Aertsen & Michiels, 2005; Galhardo, Hastings, & Rosenberg, 2007). Adaptation to a variation and heritability of a mutation, and the phenotypic variation caused by that mutation, is the raw material of biological evolution (Payne & Wagner, 2019). Payne and Wagner (2018) have discussed three major categories for causes of evolvability: (i) phenotypic heterogeneity not only by genetic variation but also without genetic mutation, including stochastic gene expression, error in protein synthesis,

epigenetic modifications and protein promiscuity (Figure 1-8-a); (ii) robustness to DNA mutations caused by a phenotype that is robust to genetic variation; (iii) topography of adaptive landscape (Figure 1-8-b,c) and location of a population (genotype) in this landscape, which is central to evolvability (Payne & Wagner, 2019).

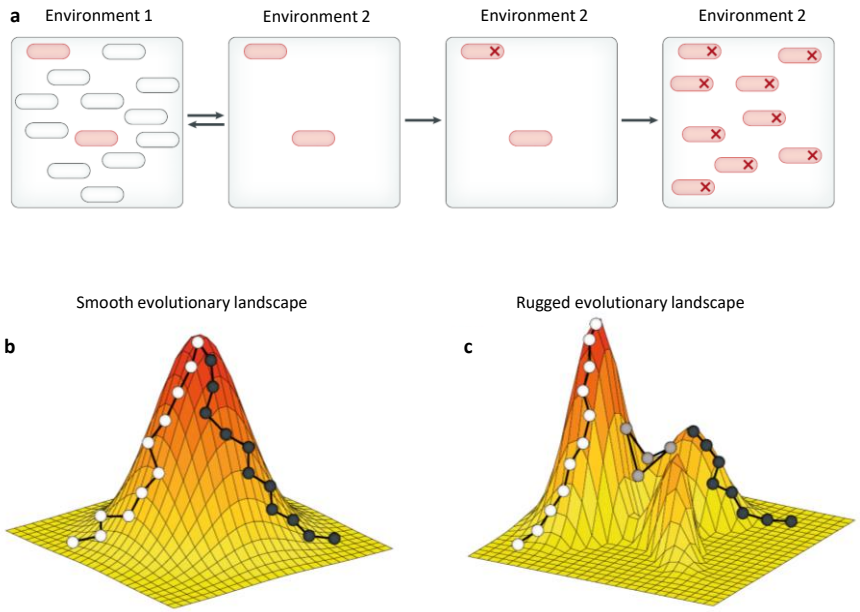


Figure 1-8: Causes of evolvability. **a)** genetic or non-genetic phenotypic heterogeneity can provide subpopulation(s) with novel traits, which might survive in particular conditions. Such heterogeneity is illustrated by the schematic red phenotype as a persister cell in environment 1, which can promote survival of the population against external challenges by environment, like antibiotic disposal (environment 2). This immediate survival can buy time for emergence of a mutation (red cross) that can fixate the phenotype or cause another variant that is adaptive. **b)** A single-peaked or smooth adaptive landscape can facilitate evolvability, because starting from any coordinate in the landscape, selection pressure can guide towards the absolute peak of fitness. On the other hand, in a multi-peaked rugged evolutionary landscape **(c)**, depending on the starting point of a genotype, it might face different trajectories and even may become trapped in a local optimum instead of reaching the global absolute maximum of fitness. (Figure adapted from Payne and Wagner, 2018).

Mutational robustness is defined as the withstanding threshold against deleterious mutations without compromising the phenotype. (Kirschner, Gerhart, & Gerhartt, 2010; Wagner, 2005b, 2005a; Zheng, Guo, & Wagner, 2020). Robustness might have evolved as an intrinsic property to shelter

biological systems against random genetic variations and environmental perturbations (Masel & Siegal, 2009). It is widely accepted that mutational robustness promotes evolvability (Bloom, Labthavikul, Otey, Arnold, & Levitt, 2006; Ferrada & Wagner, 2008; Wagner, 2008). Depending on structure of fitness landscape and condition of new environment, mutational robustness may hinder or promote adaptability (Draghi, Parsons, Wagner, & Plotkin, 2010; Grasso & Chatterjee, 2021; Stern et al., 2014). A very good example of how robustness promotes evolvability is reported by Starr et al., (2017), where they study evolvability of steroid receptors for binding to steroid response elements. They observed significant change between the evolutionary trajectories and mutational neighbourhoods of an ancestral steroid receptor (AncSR1) and its mutant with 11 amino acid substitutions (AncSR1+11p). Although the 11 mutations, distal to the DNA binding domain, did not affect DNA binding specificity, they significantly enhanced the evolutionary capacity of variants towards binding to steroid response elements. Particularly, only 41 out of 160,000 possible mutational variants of AncSR1 (without the 11 substitutions) were able to evolve for binding to steroid response elements, while presence of 11 substitutions enhanced evolvability and 829 variants out of 160,000 were able to evolve to specifically bound steroid response elements (Figure 1-9). These results highlights that how the robustness to mutation can facilitate access to the mutational neighbourhood and enhance evolvability (Starr, Picton, & Thornton, 2017).

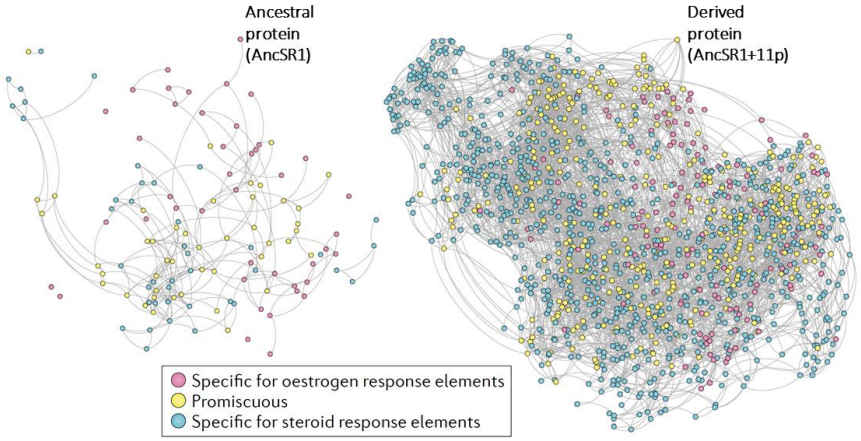


Figure 1-9: Effect of robustness in evolvability of ancestral steroid receptor. Significant difference in the mutational neighbourhoods of the ancestral steroid receptor (AncSR1) and the derived mutant of AncSR1 with 11 amino acid substitutions (AncSR1+11p) is detectable. In absence of 11 substitutions, only 41 out of 160,000 possible variants of the ancestral protein were able to evolve for binding to steroid response elements, whereas in presence of 11 mutations, 829 variants have evolved to specifically bound steroid response elements. (Figure adapted from Payne and Wagner, 2018).

Furthermore, in the field of enzyme engineering, evolvability is correlated with promiscuity of the enzyme and its conformational plasticity (Crean, Gardner, & Kamerlin, 2020; James & Tawfik, 2003b; Tokuriki & Tawfik, 2009a). A plastic conformational space, can act as a malleable clay to be shaped by natural selection. The native activity of enzymes is defined by the major conformation, whereas presence of minor conformations can facilitate promiscuous activities for non-conventional substrates. Through the course of evolution, rounds of random mutation and selection can eventually shift the minor conformations towards a new major conformation, and thus a new native activity (Figure 1-10) (Crean et al., 2020; James & Tawfik, 2003b). However, if the original activity is essential for the fitness, the evolution towards new activity will not take place unless an event such as gene duplication has happened, where among two identical proteins, one can be released from constraint imposed by cellular fitness to keep the original function, and thus, can freely explore the sequence space toward a new function, and the other will remain responsible for the original function (Aharoni et al., 2005; Copley, 2020; Fernández et al., 2005). Eventually,

as mentioned by Keshri et al., two factors define the chance of evolution for a character: probability of gaining a new character and probability of fixing the new character (Keshri et al., 2020).

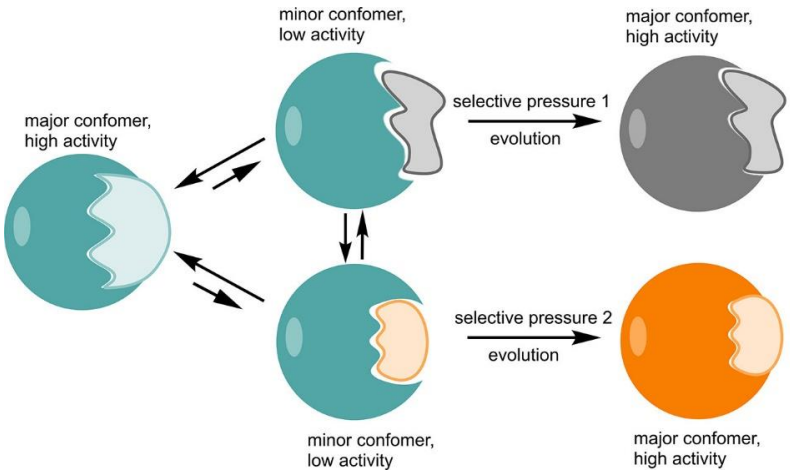


Figure 1-10: Impact of conformational dynamics and promiscuity on evolvability of enzymes. Existence of non-native, secondary conformer(s), shown as minor conformers, in the active center of enzymes can generate promiscuous activities for non-conventional substrates beside the original activity that is due to the major conformer. A drive by course of evolution, which involves rounds of random mutation and beneficial selection, can generate a new major conformer endowing new native catalytic activity. However, newly useful activity is compromising the old activity, and thus the cellular fitness, events such as gene duplication will provide a new platform for divergence towards a new enzyme (Figure from Crean et al., 2020; see also James & Tawfik, 2003 and Fernández et al., 2005).

1.3 Enzyme promiscuity

We learnt in introductory courses and biochemistry textbooks that enzymes are specialist catalysts reacting on specific substrates, but, recently, this view has been profoundly revised as many enzymes has been characterized as promiscuous catalysts. (Babtie, Tokuriki, & Hollfelder, 2010a; Busto, Gotor-Fernández, & Gotor, 2010; Copley, 2003; Khersonsky & Tawfik, 2010).

Enzyme promiscuity can be considered in different types as categorized in three main classes by Hult and Berglund: condition promiscuity (providing favorable conditions towards shifting the original activity), substrate promiscuity (affinity of enzymes for broad substrate spectrum) and catalytic promiscuity (ability of the

same active site to catalyze various reactions) (Hult & Berglund, 2007). It is largely accepted that evolution of neofunctionalization in enzyme prefers to start with promiscuous enzymes and optimize them, than to generating a *de novo* function, therefore, promiscuity is considered as a springboard for functional innovation in enzymes (Fröhlich, Chen, Gholipour, Erdogan, & Tokuriki, 2021; Khersonsky & Tawfik, 2010).

However, promiscuity may prove problematic in crowded cellular environment, which is considered as the evolutionary cause of repair enzymes to circumvent such noncanonical activities. Recent discoveries are emphasizing on role of metabolite repair enzymes in protection of metabolic pathways from promiscuous activities causing undesired interference (Bommer et al., 2020; Veiga-da-Cunha et al., 2020). Such effects of promiscuity as the source of fitness cost is underestimated in the field of directed evolution.

1.4 Protein dynamics

The classical paradigm of 'one sequence, one structure, one function' known as 'simplistic view', has been replace by the 'new view' that proteins are 'dynamic entities' and the extent of dynamic plasticity plays a vital role in their evolvability (Crean et al., 2020; James & Tawfik, 2003b; Tapia-Rojo, Alonso-Caballero, Badilla, & Fernandez, 2021). In 'simplistic view', known for the 'lock and key' metaphor, proteins have a smooth single-well energy diagram with a global minimum, which represents just one structural conformer. Whereas, in the 'new view' corresponds to a rugged energy diagram with many local minima representing an ensemble of conformers with similar but distinct energy levels (Figure 1-11).

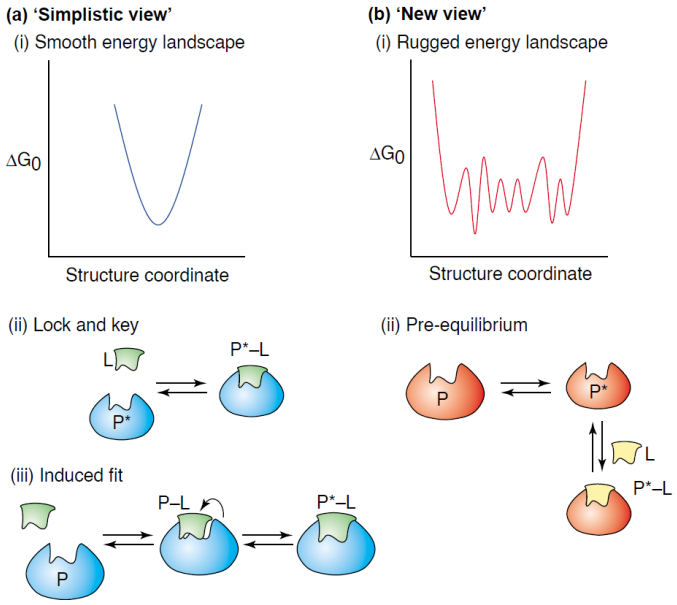


Figure 1-11: Schematic of conformational paradigm in 'simplistic' and the'new view'of proteins. **a)** In simplistic view of proteins, a single-well energy diagram (i) represents only one conformer with a functional state of either lock and key (ii) or induced fit (iii). In other words, in induced fit, the fitted conformation (which is a second well in energy but too high to be populated) is not accessible without ligand (L) but is stabilized with L (reaching opposite energies of the two-well curve when L is present). In fact, induced fit is also compatible with the 'ensemble' view but having wells at higher energy levels that become accessible and populated when L is there. **b)** The new view represents an ensemble of different conformers with similar but distinct energy levels (i), where the mode of function is defined by an equilibrium between pre-existing isomers that only one of the conformer endows the function (ii). (Figure taken from James and Tawfik, 2003).

Protein dynamic is known as the ultimate determinant of enzyme catalysis, since even highly different primary sequences are ending up with highly similar structural folds and much further, even enzymes with highly similar structural folds are known to catalyse disparate reactions. In addition, precision and specificity of enzymes are highly correlated with their conformational dynamics. Therefore, in advanced enzyme studies, in addition to four level of protein structure (primary, secondary, tertiary, and quaternary), the conformational dynamics of proteins is increasingly being applied to predict enzymatic functions (Campbell et al., 2018; Crean et al., 2020; Galdadas et al., 2021; James & Tawfik, 2003b; Petrović, Risso, Kamerlin, & Sanchez-Ruiz, 2018; Schenkmayerova et al.,

2021). Studies has shown that even mutations and dynamic alterations in allosteric sites of enzymes can influence dynamics of catalytic core of enzyme (Galdadas et al., 2021; Simakov, Leonard, Smith, Wymore, & Szarecka, 2017), which shows that the whole fold of protein is playing a role in its final dynamics and conformational space. Mutations and ligands can alter this dynamic, either through direct effect on active core conformation or disturbance of the allosteric communication pathways. For instance, a recent study by Galdadas et al., (2021) shows that 45 of the 90 natural amino acid substitutions identified in clinical isolates of TEM-1, which increase its activity spectrum towards β-lactams and/or enhance stability of the enzyme, are located on allosteric communication pathways and hypothesized to affect the propagation of conformational changes (Figure 1-12). More interestingly, they found that binding of ligands at different allosteric sites in two similar enzymes (TEM-1 and KPC-2 class A β-lactamases), affect the active site very similarly by eventually propagating changes to the same loops (α_3- α_4 loop and the Ω-loop) in both of the enzymes (Galdadas, Qu et al., 2021).

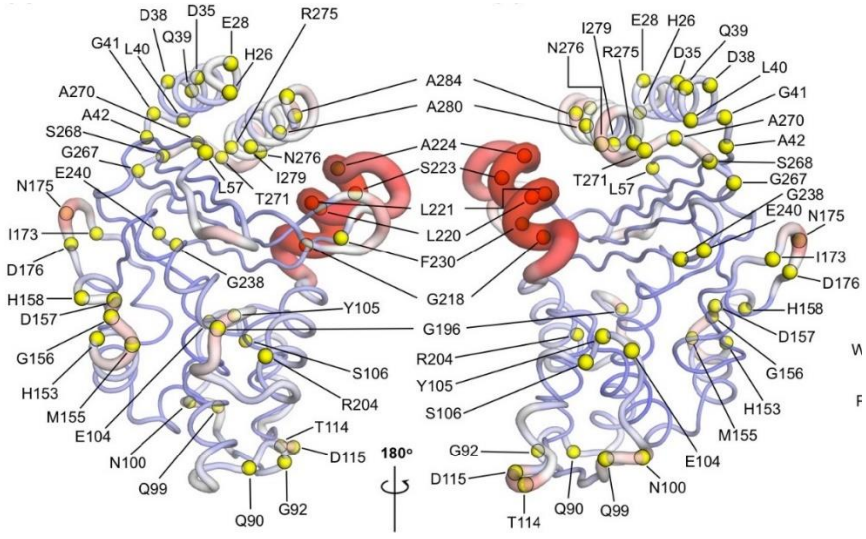


Figure 1-12: Point substitutions that extend activity spectrum and/or stability in natural variants of TEM-1 affecting the dynamic pathway of the proteins. Variant positions (yellow spheres) in TEM-1 β-lactamase, centered at their corresponding Cα, which are located along allosteric communication pathways, suggesting that these mutations made their impact on catalytic behavior of the enzymes by affecting the allosteric dynamics.

The snapshot of average Cα deviation between Apo (no ligand) nonequilibrium (Apo_{NE}) and IB_{EQ} (inhibitor-bound) states, is represented by the color scheme and cartoon thickness of the respective structures (Galdadas, Qu et al., 2021).

1.5 Disulfide bonds

1.5.1 Formation of disulfide bond

Disulfide bonds (S-S bond) in biology are covalent bridges formed upon oxidation and coupling of two thiol groups between pairs of cysteines (Figure 1-13). S-S bond can occur between cysteines of the same or different polypeptides.

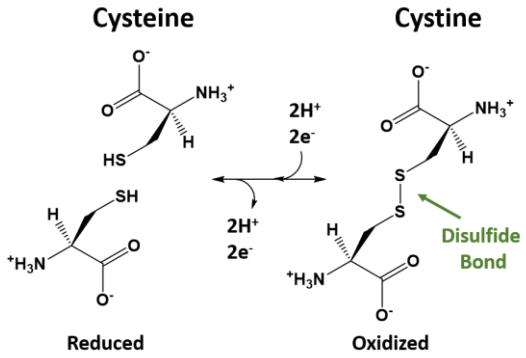


Figure 1-13: Disulfide (S-S) bond formation and cleavage, by oxidation (release of two protons and two electrons) of thiol groups of two cysteine and reduction of the S-S bridged cystine molecule, respectively.

Disulfides are mostly found in secreted proteins and are usually thought to stabilize these proteins. Formation of the disulfide bond is an oxidation reaction, thus the reducing environment of cytoplasm is unfavorable for this reaction. Therefore, cells have developed dedicated subcellular compartments and catalytic pathways for fast and correct formation of disulfide bonds. In prokaryotes, the formation of disulfide bonds is catalyzed in extra cytoplasmic environments. In Gram-negative bacteria, disulfide bond formation is catalyzed in oxidative environments of the periplasm (Denoncin & Collet, 2013; Landeta, Boyd, & Beckwith, 2018), whereas in Gram-positive bacteria, which lack the periplasmic space, disulfide bond catalysis is not fully understood but it is suggested to be catalyzed by thiol-disulfide oxidoreductases located at the cell wall or associated with the cytoplasmic membrane (S. F. Lee & Davey, 2017;

Reardon-Robinson & Ton-That, 2015). In eukaryotes SS bonds are formed in the lumen of the endoplasmic reticulum (ER), catalyzed by the protein disulfide isomerase enzymes (PDI) (Hatahet & Ruddock, 2009; Regeimbal & Bardwell, 2002).

In Gram-negative bacteria, like *E. coli*, oxidative protein folding in the periplasm of the cell envelope demands an electron transfer network provided by specialized catalysts of DsbA/DsbB pathway, for disulfide bond formation, and the DsbC/DsbD pathway, for isomerization of disulfide bond (Figure 1-14). The catalytic mechanism of disulfide bond formation in proteins involves a thiol–disulfide exchange between disulfide donor catalyst and the disulfide acceptor protein (Bushweller, 2020; Denoncin & Collet, 2013; Goemans, Denoncin, & Collet, 2014; Messens & Collet, 2006).

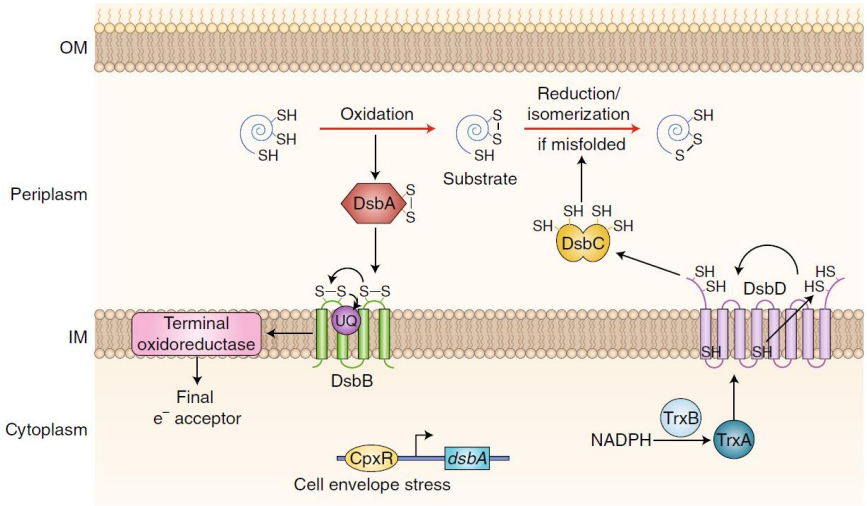


Figure 1-14: Disulfide bond formation and isomerisation in periplasm of *E. coli*. A protein carrying pair(s) of reduced thiol is oxidized by DsbA, by donation of its active site disulfide via a thiol–disulfide exchange reaction. The DsbA will be released in a reduced state, which requires another catalyst, called DsbB, for its re-oxidization and conveying the electrons to ubiquinone (UQ). Disulfide bond isomerization and correction is facilitated by reductive power of DsbC, which receives the electrons of cytoplasmic thioredoxin via DsbD (figure taken from Landeta et al, 2018; see also Messens & Collet, 2006 and Hatahet et al., 2014).

1.5.2 Effect of disulfide bond on biophysical properties of proteins

Unlike the covalent peptide bonds, disulfide bonds are post-translationally introduced into proteins and they are also reversible covalent bonds through

thiol/disulfide exchange reactions. These bonds affect the biophysical characteristics of proteins, such as stability, tendency for aggregation or degradation rate that are important parameters for cellular and organismal fitness (DePristo et al., 2005). Additionally, disulfide bonds will also modulate the level of plasticity and conformational freedom. Protein dynamics have been proposed as the heart of enzyme catalysis, but their identification, analysis and design has proven elusive (Luk et al., 2013; Schenkmayerova et al., 2021). It is expected that dynamics and conformational plasticity has been optimized through evolutionary trajectories of proteins but our understanding of the underlying molecular mechanisms are still blurry.

Although disulfide bonds are well known for their role in stability of proteins, their impact on function and dynamics of proteins, by imposing constraints on conformational space, requires further attention. Emerging evidences are emphasizing that disulfide bonds play an important role as reversible switches in signaling and regulatory mechanisms (Butera et al., 2018; Hogg, 2003; Messens & Collet, 2013; Passam & Chiu, 2019). Intramolecular disulfide bonds are suggested to act by reducing protein dynamics (Ishikawa, Kim, Kwak, Wakasugi, & Fayer, 2007), and even a distal disulfide bridge can influence the dynamics of catalytic core of proteins, see Figure 1-15 (Simakov et al., 2017).

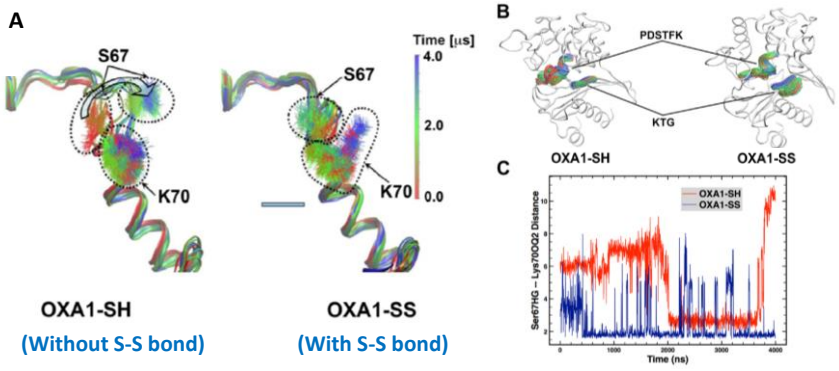


Figure 1-15: A distal disulfide bond in OXA-1 β-Lactamase influences the dynamics of catalytic center by stabilizing the active sites motifs. A) Dynamics of side-chain conformations of catalytic Ser67 and Lys70 by time distribution in presence (OXA1-SS) and absence (OXA1-SH) of disulfide bond. B) Conformational diversity of catalytic motifs $PDS_{67}TFK_{70}$ and KTG. (C) Distance between hydrogen γ (Hγ) of Ser67 and OQ2

of the carboxylated Lys70 in reduced (red) and oxidized (blue) conditions. The disulfide-containing state of the enzyme reveals less freedom of the catalytic center than the non-bridged enzyme (Simakov et al., 2016).

1.5.3 Evolutionary impact of disulfide bonds

There is still much to be learnt about impact of the structure and posttranslational modifications of proteins on rates of their evolution. As reported by Feyertag and Alvarez-Ponce (2017), in a same subcellular location, proteins that contains disulfide bonds show a faster evolutionary rate than proteins with no disulfide bonds. In agreement with their hypothesis that disulfide bonds enhance protein evolvability, by acting as buffers of deleterious mutations through providing stable protein scaffold, they found that disulfide bridge containing membrane proteins evolved 88% faster than proteins without the bridge, and also among extracellular proteins, the rate of evolution for proteins with disulfide bonds is 49% faster compare to proteins lacking disulfide bonds (Figure 1-16). They highlight that disulfide bonds increase designability and evolvability of proteins by (i) making protein structures relatively independent of their primary amino acid sequences, and (ii) thanks to the buffering effect of disulfide bonds to compensate deleterious mutations, nonsynonymous mutations (amino acid substitutions) can accumulate, which endow a precursor for further adaptive evolution (Feyertag & Alvarez-Ponce, 2017).

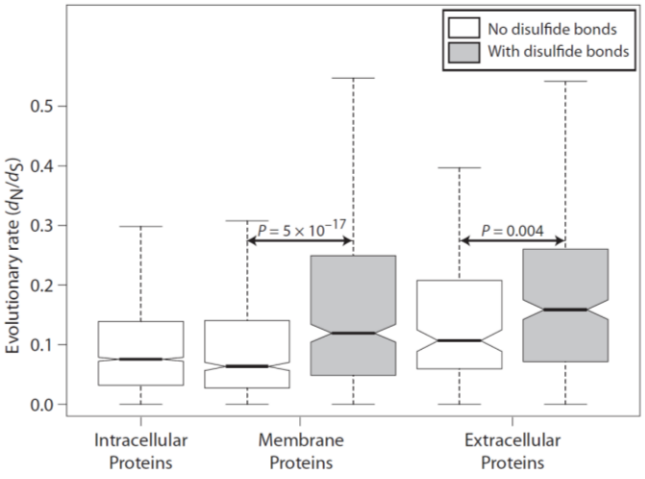


Figure 1-16: Effect of disulfide bonds on evolutionary rates of proteins. Most likely ortholog of human protein-coding genes, were identified in the genome of mouse and

the rate of protein evolution was estimated from the nonsynonymous to synonymous divergence ratio (d_N/d_S). Proteins containing disulfide bonds in their structure are indicated in grey while proteins without any disulfide bond are indicated in white. Outlier proteins are not shown. *P*-values are obtained from the Mann–Whitney's U test (Feyertag & Alvarez-Ponce, 2017).

1.6 Bacterial cell wall

Bacterial cell wall is involved in cell growth and division and it is essential for its morphology and integrity (Rogers et al., 1980). In Gram-negative bacteria, the cell wall appears more complex and composed of two main components of peptidoglycan (PG) layer and outer membrane (OM), where the space between inner membrane (IM) and OM is called the periplasm that also contain the PG sacculus (Figure 1-17-a). Gram-positive bacteria possess a cell wall that is predominantly composed of peptidoglycan as a multilayered structure around the cell membrane (Figure 1-17-b). Due to its uniqueness to bacteria, the bacterial cell envelope serves as a main target for antimicrobial agents such as penicillins. In this section, we will focus on the Gram-negative envelope with emphasis on the protagonists of peptidoglycan biosynthesis and maintenance that are targeted by penicillins as well as phylogenetically related enzymes that are conferring resistance to these antibiotics.

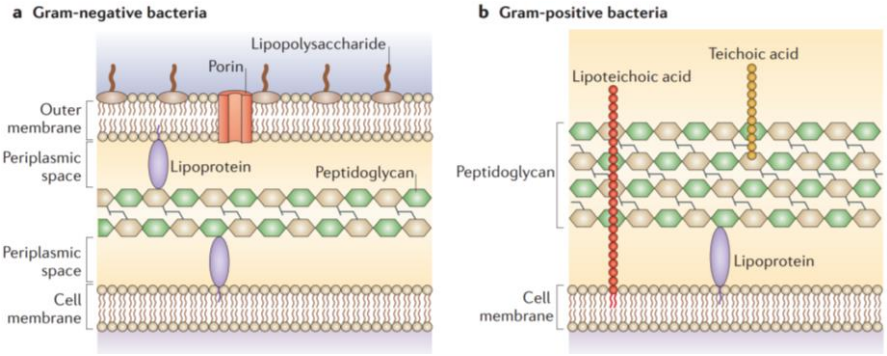


Figure 1-17: Bacterial cell wall. **a)** Gram-negative bacteria have a more complex cell wall containing an additional plasma membrane outside the PG layer and a periplasmic space between the outer membrane and cytoplasmic membrane. The peptidoglycan layer is much thinner than that of Gram-positive bacteria. The passage of small hydrophilic molecules across the outer membrane is facilitated through porins. Outer membrane and peptidoglycan are covalently attached via a lipoprotein. **b)** The cell wall of Gram-positive bacteria is a multilayered thick mesh of peptidoglycan (PG) surrounding the cytoplasmic membrane. PG can represent up to 90% of the Gram-positive cell wall. Teichoic acid glycopolymer is an additional component of Gram-positive cell wall, embedded through

the peptidoglycan layers, and suggested to have important roles such as rigidity of the cell wall, and thus resilience and morphology of the cell as well as providing a net negative charge for the cell (Figure taken from Brown, Wolf, Prados-Rosales, & Casadevall, 2015).

1.6.1.1 Peptidoglycan (PG)

Peptidoglycan (also called murein) is an essential and unique characteristic of bacterial cell wall surrounding the cytoplasmic membrane. The PG has developed to protect the bacterial cell from its high internal osmotic pressure, which pushes the plasma membrane against the cell wall (Cabeen & Jacobs-Wagner, 2005; Pazos & Peters, 2019b; Vollmer & Höltje, 2004). The glycan strands are polysaccharide chains of alternating β-1,4-linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) glycan units cross-linked by unusual peptide bridges containing D-amino acids. These peptides are connected to the lactyl group of MurNAc through an amide bond. The periplasmic precursors of PG are membrane anchored GlcNAc-MurNAc-pentapeptides molecules called lipid II and whose biosynthesis is described in the following sections. The nature and extent of peptide linkages vary among different species, but most of the known cross-links are formed under D-Alanyl-D-Alanine-transpeptidation reactions, which are catalyzed by DD-transpeptidase enzymes (DD-TPases) (Dörr, Moynihan, & Mayer, 2019; Sauvage, Kerff, Terrak, Ayala, & Charlier, 2008b; Verwer, Nanninga, Keck, & Schwarz, 1978; Vollmer & Bertsche, 2008; Vollmer, Blanot, & De Pedro, 2008).

1.6.1.1.1 Peptidoglycan biosynthesis:

Biosynthesis of PG is a multistep process that starts by synthesis of precursors in cytoplasm, followed by assembly of intermediates (lipid I and lipid II) on inner side of plasma membrane and transfer of these intermediates to the external side of the membrane thanks to the flippase proteins. (Barreteau et al., 2008; Bouhss, Trunkfield, Bugg, & Mengin-Lecreulx, 2008; Egan, Errington, & Vollmer, 2020; Sham et al., 2014). The process ends up with polymerization and cross-linking of glycan strands in periplasm to form PG mesh-like macromolecule. The overall process is schematized in Figure 1-18.

Cytoplasmic steps of PG biosynthesis: the cytoplasmic steps of PG biosynthesis, are generally divided into four main sets of reactions: 1) Synthesis of UDP-GlcNAc from fructose 6-phosphate; 2) formation of UDP-MurNAc from UDP-GlcNAc; 3) synthesis of D-glutamic (D-Glu) acid and the dipeptide D-alanyl-D-alanine (D-Ala-D-Ala) building blocks; 4) assembly of UDP-MurNAc-pentapeptide from UDP-MurNAc, amino acids (D-Glu, L-Ala, m-DAP) and dipeptide (D-Ala-D-Ala); where each set of reactions involves specialized enzymatic catalysis (see Figure 1-18). The final product of cytoplasmic steps of PG biosynthesis is the monosaccharide-peptide of UDP-MurNAc-pentapeptide (Barreteau et al., 2008; York et al., 2020).

Membrane steps of PG biosynthesis: the membrane steps involve addition of the GlcNAc to the MurNAc-pentapeptide and translocation of this precursor of PG into periplasm. The Mur*N*Ac-pentapeptide is first transferred by a transferase, called MraY, onto the lipid carrier (undecaprenyl-phopshate) by releasing a uridine monophosphate (UMP) and formation of the product named lipid I. UDP-GlcNAc is then used by MurG enzyme to transfer GlcNAc to lipid I with the release of a uridine diphosphate (UDP) and formation of lipid II. Lipid II is then translocated through a flipping mechanism by MurJ membrane proteins (Yao Liu & Breukink, 2016; Pazos & Peters, 2019a; Typas, Banzhaf, Gross, & Vollmer, 2011). However, recent findings suggests that flippase activity is mostly carried out by MurJ flippase rather than FtsW, which is characterized as a PG polymerase that works in complex with a class B PBP (Sham et al., 2014; Taguchi et al., 2019).

Periplasmic steps of PG biosynthesis: the final steps in PG synthesis to form the sacculus are carried out in periplasmic space, where the final product (lipid II) from previous steps is located. The PG synthesis includes the polymerization of glycan strands catalyzed by glycosyltransferase (GTase) domain of class A high molecular weight penicillin-binding proteins (HMW-PBPs) and/or a complex of SEDS proteins with class B HMW-PBPs , and cross-linking of the strands through their stem peptides that is catalyzed by transpeptidase (TPase) enzymes.

Depending on donor stem peptides, cross-linkage can be catalyzed either by DD-TPase or LD-TPase enzymes (Figure 1-18) (Egan et al., 2020; Pazos & Peters, 2019a; Straume et al., 2020; Taguchi et al., 2019; Typas et al., 2011; Vollmer & Bertsche, 2008). However, most of the cross-linking takes place via DD-TPase reaction between D-Ala at position four of donor pentapeptide with the meso-diaminopimelic acid (mDAP) at third position the acceptor muropeptide, leaving very low amount of pentapeptide behind (Vollmer & Bertsche, 2008). Though less frequent, two stem muropeptides can also be cross-linked via their mDAP residues at position three through a LD-TPase reaction. Much less frequent are the cross-links between D-Ala at position four and the D-iGlu at position two that has been found in Corynebacteria (Schleifer & Kandler, 1972). The attachment of the PG to outer membrane cross-link is formed non-covalently via outer membrane proteins (OMPs) and Pal lipoprotein, and covalently by the lipoprotein Lpp (Braun's lipoprotein) in γ-proteobacteria such as *E. coli*, which is catalyzed by LD-TPase enzymes by acting on a stem tetrapeptide of PG as donor and linking the α-carboxyl group of mDAP to ε-amino group of the C-terminal lysine of Lpp (Godessart et al., 2021; Magnet et al., 2007).

1.6.1.2 Periplasm

The viscous space between IM and OM, in Gram-negative bacteria, is called periplasm where polymeric and elastic mesh-like structure of peptidoglycan is accommodated (Figure 1-17). Proteins secreted to this extracytoplasmic compartment are produced in the cytoplasm and then exported across the IM through secretory pathways thanks to a signal sequence they carry in their N-terminal. Proper folding of these proteins occur in periplasm by the help of molecular chaperones (Driessen, Manting, & van der Does, 2001; Goemans et al., 2014; Miot & Betton, 2004; Wickner, Driessen, & Haril, 1991). Unlike cytoplasmic environment, the periplasm is more directly exposed to the external environment; hence, making its proteins more vulnerable against fluctuations in bacterial growth conditions such as oxidative stress, changes in pH and temperature, osmotic shocks, presence of detergents, etc. (Arts, Gennaris, &

Collet, 2015b; Ezraty, Gennaris, Barras, & Collet, 2017; Merdanovic, Clausen, Kaiser, Huber, & Ehrmann, 2011; Wilks & Slonczewski, 2007).

1.6.1.3 Outer membrane (OM)

OM as the external layer of bacterial cell wall consists of lipopolysaccharides (LPS), proteins and phospholipids (Kamio & Nikaido, 1976; Sperandeo, Martorana, & Polissi, 2019). In *Escherichia coli*, OM is covalently cross-linked to PG through lipoprotein Lpp (also called Braun's lipoprotein), which is bound by its C-terminal lysine to diaminopimelic acid (DAP) of murein and is integrated into the inner layer of the OM by its N-terminal lipid anchor. (V. Braun & Bosch, 1972; V. Braun & Rehn, 1969; Volkmar Braun & Hantke, 2019). Three highly similar L-D-transpeptidases in *E. coli*, namely ErfK, YbiS, and YcfS (also known as LdtA, LdtB, LdtC, respectively), catalyze the replacement of D-Ala linked to meso-DAP in the third position of the muropeptide by the amino group of the C-terminal lysine of Lpp (Magnet et al., 2007). Recently, a similar covalent attachment involving the N-terminal amino group of outer membrane proteins (Omps) has been reported for Gram-negative bacteria that do not express a Lpp ortholog (Godessart et al., 2021; Sandoz et al., 2021).

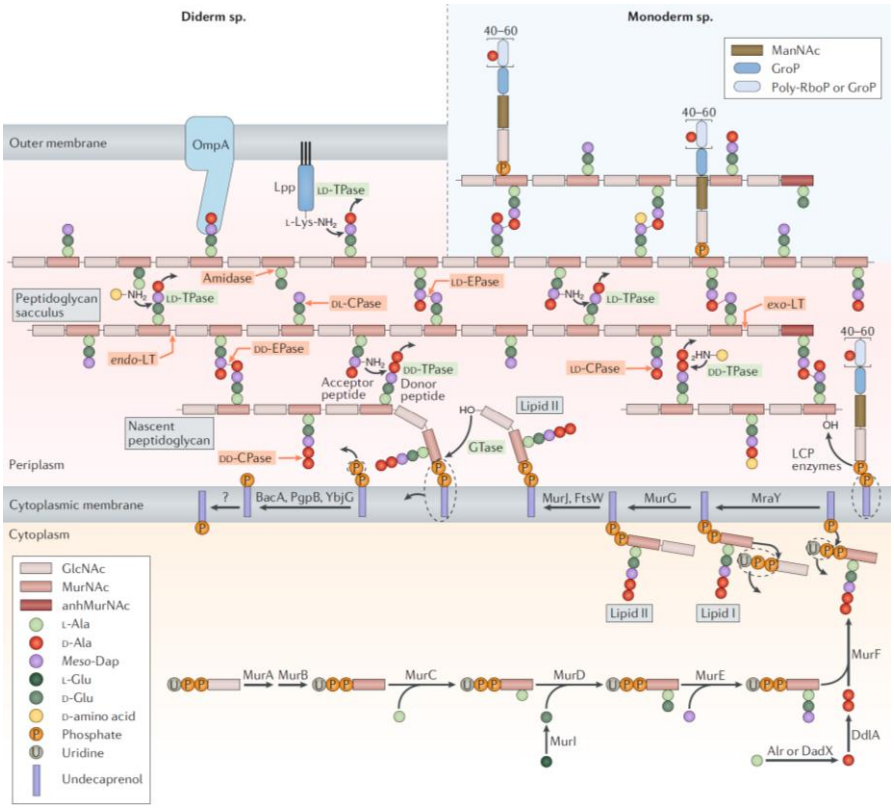


Figure 1-18: Schematic illustration of of peptidoglycan biosynthesis and modification. Essential reactions for the synthesis and hydrolysis of PG in the cell envelopes of monoderm (Gram-positive) and diderm (Gram-negative) bacteria are shown. Cytoplasmic steps involves synthesis of UDP-MurNAc by MurA and MurB followed by stepwise addition of specific amino acids by MurC, MurD, MurE, MurF. Synthesis of D-Ala-D-Ala is carried out by Ddl and racemization of L-amino acids are catalyzed by by Alr or DadX (L-Ala to D-Ala) and MurI (L-Glu to D-Glu). Synthesis of UDP-MurNAc-pentapeptide followed by membrane steps, where MraY and MurG mediate formation of lipid I and then lipid II by attachment to lipid carrier and addition of a GlcNAc, respectively. Lipid II is the translocated (flipped) to the outer leaflet of the cytoplasmic membrane via MurJ and probably FtsW. In periplasm, Glycosyltransferases (GTases) polymerize the nascent peptidoglycan strands from lipid II. Nascent chains are then cross-linked to each other by DD-transpeptidases (DD-TPases). Stem peptides are trimmed by carboxypeptidases (CPases), which can remove the terminal D-Ala from either pentapeptides by DD-carboxypeptidases (DD-CPase) or tetrapeptides by LD-carboxypeptidases (LD-CPase) or remove the mDAP from tripeptides by DL-carboxypeptidases (DL-CPases). Cleavage of crosslinks is carried out by DD-endopeptidases and LD-endopeptidases (EPases). Amidases cleave the bond between the peptide and glycan strands, and lytic transglycosylases (LT) break the glycan chains and release 1,6-anhydro-N-acetylmuramic acid (anhMurNAc). Covalent cross-link of the outer-membrane-associated lipoprotein Lpp is accomplished by LD-transpeptidase (LD-TPases). In diderm species the highly conserved outer membrane protein A (OmpA) is

non-covalently as well as covalently attached to the PG sacculus, which plays role in cell wall integrity. Abbreviations: Alr, Ala racemase, biosynthetic; DadX, Ala racemase, catabolic; DdlA , D-Ala-D-Ala ligase A; GlcNAc, N-acetylglucosamine; GroP, glycerol phosphate; ManNAc, N-acetyl-D-mannosamine; meso-Dap, meso-diaminopimelic acid; MraY, phospho-N-acetylmuramoyl-pentapeptide transferase; MurA , UDP-GlcNAc enolpyruvyl transferase; MurB, UDP-MurNAc dehydrogenase; MurC, UDP-MurNAc L-Ala ligase; MurD, UDP-MurNAc-L-Ala-D-Glu ligase; MurE, UDP-MurNAc-L-Ala-D-Glu-meso-Dap ligase; MurF, UDP-MurNAc-tripeptide-D-alanyl-D-Ala ligase; MurG, undecaprenyldiphosphomuramoylpentapeptide β-N-acetylglucosaminyltransferase; MurI, Glu racemase; MurJ, MOP (multidrug, oligosaccharide-lipid, polysaccharide) transporter (lipid II flippase); MurNAc, N-acetylmuramic acid; Poly-RboP/GroP, poly-ribitol/glycerol phosphate. Different hydrolase activities (CPase, EPase, Amidase, endo- and exo-LT) are highlited in orange and their target is indicated by orange arrows, while transpeptication reactionse (TPase) are highlighted in light green. (Figure taken from Egan et al., 2020).

1.6.2 Penicillin-binding proteins (PBPs) and β-lactamases

1.6.2.1 PBPs

Due to structural analogy between their substrate (D-Ala-D-Ala) and the β-lactam antibiotics such as penicillin, DD-peptidase enzymes bind similarly to β-lactams such as penicillin and because of that, they are also called penicillin-binding proteins (PBPs). PBPs are divided into two groups according to their molecular mass: high molecular weight PBPs (HMW-PBPs, >60 kDa) and low molecular weight (LMW-PBPs, <60 kDa) (Clarke et al., 2009; Sauvage et al., 2008b). HMW-PBPs are known as bifunctional enzymes as they may contain a transglycosylase domain (TGase) beside their DD-transpeptidase (DD-TPase) domain, and based on that they are devided into two classes: class A HMW-PBPs with both TGase and TPase activities and class B HMW-PBPs with only TPase activity (Macheboeuf, Contreras-Martel, Job, Dideberg, & Dessen, 2006). For decades, it was assumed that only class A HMW-PBPs and related enzymes are involved in peptidoglycan polymerization, however, it was shown recently that members of the SEDS (shape, elongation, division and sporulation) protein family catalyze peptidoglycan polymerization. SEDS proteins RodA and FtsW, in complex with class B HMW-PBPs (SEDS–bPBP complex), are found to function as PG polymerases by polymerizing glycan strands, which make them potentially new antibacterial targets (Cho et al., 2016; Emami et al., 2017; Sjodt et al., 2020; Taguchi et al., 2019). On the other hand, class A HMW-PBPs are

suggested to posses activities independent of divisome and elongasome that involved in processing of recently formed peptidoglycan (Straume et al., 2020). LMW-PBPs are known to be acting *in vivo* as DD-carboxypeptidases (DD-CPase) and DD-endopeptidases (DD-EPase) (Clarke et al., 2009; Korat, Mottl, & Keck, 1991). Among three class of LMW-PBPs (A, B and C), class C enzymes possess both DD-CPase and DD-EPase activities and show sequence similarities to class A β-lactamases (Clarke et al., 2009). LMW PBPs are in general compatible with common term of class C PBPs, sometimes subdivided to C1, C2 and C3. Sauvage et al. rather subcategorized class C PBPs into four types by referring to their main representative in *E. coli*: type-4 for PBPs analogous to PBP4, type-5 for analogues of PBP5, type-7 analogues of PBP7 and type-AmpH for PBPs matching to *E. coli* AmpH (Sauvage, Kerff, Terrak, Ayala, & Charlier, 2008a).

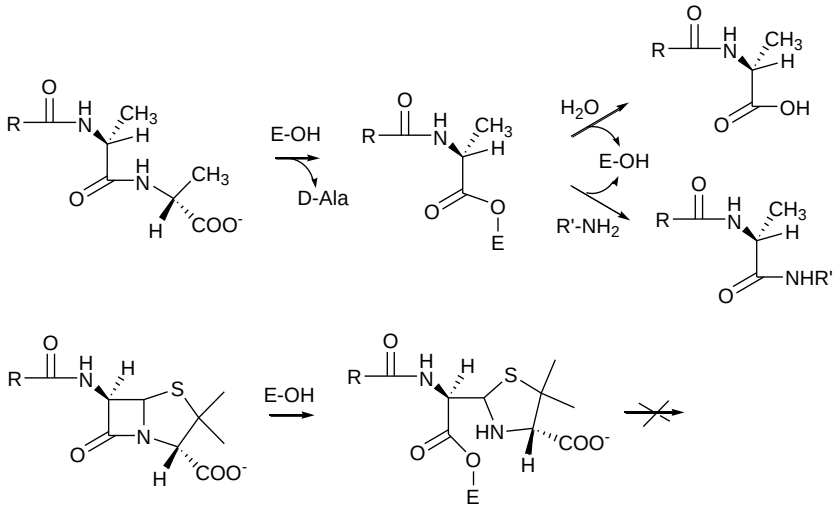


Figure 1-19: Catalysis and inhibition of D-Ala-D-Ala-peptidases. **Up:** DD-peptidases act on the C-terminal D-Alanyl-D-Alanine of the peptidoglycan precursors. The first nucleophilic attack (acylation) by the essential serine generates the acyl-enzyme. The second nucleophilic attack (deacylation) by a water molecule (DD-carboxypeptidases) or an amino group (DD-transpeptidases) regenerates the free enzyme and completes the catalytic cycle. **Down**: DD-peptidases are inhibited by penicillins which mimic the natural substrate. The acyl-enzyme is formed as with the substrate but the active site is then too crowded for the second nucleophile to enter and the enzyme is covalently inhibited.

LMW-PBPs are suggested as ancestors of class A serine β-lactamases (SBLs), such as TEM-, SHV-, and KPC-types. Among known LMW-PBPs, a family of

cyanobacterial PBPs, called PBP-A show the highest sequence similarity (~30%) to class A SBLs and thus, represent the closest phylogenetic relatives to this class of β-lactamases. PBP-A of *Thermosynechococcus elongatus,* contains all the common catalytic features of class A β-lactamases, except for not having a longer Ω-loop, which is re-shaping the active site and carrying an essential Glu166 that is involved in water activation for deacylation process (Fröhlich et al., 2021; Matagne, Lamotte-Brasseur, & Frère, 1998). Overall, sequence and structural comparisons suggested that evolutionary trajectory from PBPs to class A SBLs has involved insertions and deletions as well as extensive optimization through point mutations, which lead to SBL enzymes with generally shorter sequence (<350 residues) than LMW-PBPs (>400 residues) (Fröhlich et al., 2021). However, these changes could just be the result of divergence of protein families and it is still not fully understood how much of these mutations play important role in the functional conversion and/or fitness of class A SBLs. One of the crucial structural feature is thought to be the elongation of the active site Ω-loop in SBLs, which re-shaped the active site cavity and carries a conserved residue (Glu166) that is important in deacylation mechanism. Contrary to this notion, it has been recently discovered that the length of the omega-loop is not a key event in the birth of β-lactamase activity (Gilles Joachim's Thesis, 2016).

1.6.2.2 β-lactamases

Inhibition of PBPs by β-lactams is the origin of these antibiotics' action and, in the course of evolution, some bacteria developed resistance by producing β-lactamase capable of hydrolyzing β-lactams. Although the clinical application of β-lactam antibiotics has a short history of less than a century, it has been estimated by phylogenetic studies that the evolution of β-lactamases has happened more than 2 billion years ago (Bush, 2018; Hall & Barlow, 2004). According to Ambler classification based on homology of amino acid sequences, β-lactamase enzymes can be categorized into four classes: A, B, C, and D (Ambler, 1980; Bush & Jacoby, 2010; Philippon, Slama, Dény, & Labia, 2016). Enzymes of classes A, C, and D β-lactamases are serine hydrolases, while

enzymes of class B are metallo β-lactamase, which require one or more zinc ions for their activity (Galleni et al., 2001; Palzkill, 2013) (Palzkill, 2013). While there is no significant sequence similarities between different classes of SBLs, they share a similar structural architecture (Philippon et al., 2016; Tooke et al., 2019). Generally, emergence and maintenance of antibiotic resistance in bacteria is costly, especially in absence of antibiotics, which is associated with imposing a fitness burden on the cell (Andersson & Hughes, 2010; Andersson & Levin, 1999; Lenski, 1998; Sousa, Magalhães, & Gordo, 2012)). Evolutionary mechanisms of how the resistance has emerged and was fixed in a bacterial population is not fully understood and is of both fundamental and clinical importance (Andersson & Hughes, 2010; Andersson & Levin, 1999; MacLean & Millan, 2019; Millan et al., 2014).

PBPs and β-lactamases belong to the same family of penicillin-recognizing enzymes (PRE) enzyme, and due to common structural fold and conservation of catalytic motifs between them, is largely accepted that β-lactamases has evolved from ancient PBPs (Bush, 2018; Ghuysen et al., 1996; Kelly et al., 1986; Keshri et al., 2020; Koch, 2000; Meroueh, Minasov, Lee, Shoichet, & Mobashery, 2003). Both of the PBPs and β-lactamases employ the very same acylation mechanism to β-lactam molecules, such as penicillin, but it is the inability of PBPs in hydrolysis of acyl-enzyme in deacylation step that makes these enzymes inhibited by penicillins: PBPs as target of β-lactams and β-lactams as substrates of β-lactamases, see Figure 1-19 and Figure 1-20.

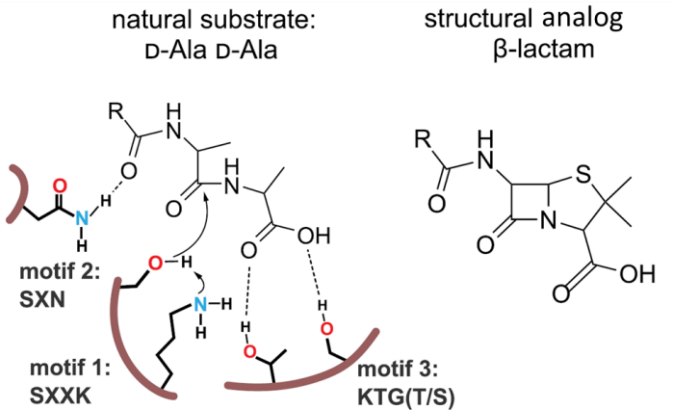


Figure 1-20: Structural analogy between D-Ala-D-Ala substrate of penicillin binding proteins (PBPs) and β-lactams, and schematic of catalytic interactions and mechanisms in the active site of PBPs through the three generally conserved catalytic signatures, SXXK, SXN and KTG(T/S) (figure adapted from Frohlich et al., 2021).

1.6.2.3 Class A β-lactamases

The first discovered, most abundant and well-studied β-lactamases are class A SBLs (Fröhlich et al., 2021; Philippon et al., 2016; Tooke et al., 2019). Because of highly efficient hydrolysis of penicillin, members of this class of β-lactamases are also known as penicillinases. The hydrolysis rate of these enzymes is estimated close to the substrate diffusion limit (k_{cat}/K_M in the range of 10^8 M^{-1} sec^{-1}) (Hardy & Kirsch, 1984). As a member of serine penicillin recognizing enzymes (PREs), class A β-lactamases share two distinct structural domains (see Figure 1-21): an all-α domain made of an array of eight α-helices and an αβ-domain harboring five antiparallel β-sheets and two α-helices. Two hinges are bridging the domains and the active site is located at the interface between the two domains. Three common catalytic motifs are conserved between class A β-lactamases: the S_{70}XXK motif, where the catalytic nucleophile serine is located; the $(S/Y)_{130}$XN motif that is held by a loop from all α-domain and the $(K/R)_{234}$(S/T)G motif forming the floor of the active site by a β-sheet; the numbering is based on Ambler numbering (Ambler et al., 1991; Fröhlich et al., 2021; Joris et al., 1988; Matagne & Frère, 1995; Matagne et al., 1998; Philippon et al., 2016). These three catalytic motifs are also conserved in PBPs and other SBLs (see Figure 1-21), but another structural feature is unique to class A β-

lactamases, a 16 residue loop, called the Ω-loop, which is facing the catalytic core and carries a key catalytic residue acting as a general base for deacylation step, the Glu166 (Adachi, Ohta, & Matsuzawa, 1991; Urbach, Fastrez, & Soumillion, 2008).

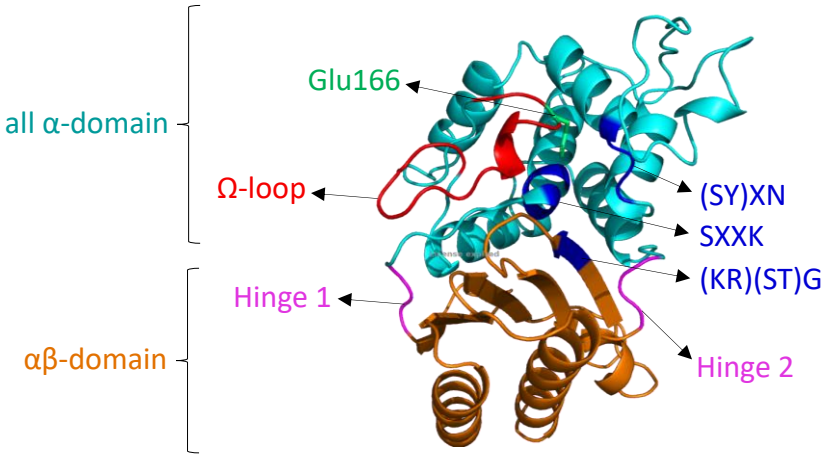


Figure 1-21: General structure of class A β-lactamases. The structure is TEM-1 β-lactamase (PDB: 1BTL). Colors are representing: cyan, the all α-domain; dark orange, the αβ-domain; magenta, the two hinges; blue, the three conserved catalytic motifs; red, the Ω-loop; green the general base Glu166. (Structure is visualized and modified by The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC).

The evolutionary relationships between LMW-PBPs and β-lactamases is more traceable in serine β-lactamases as these enzymes has conserved the similar acylation mechanism to β-lactams, particularly class A SBLs, which show highest sequence similarity to LMW-PBPs. Class C LMW-PBPs (includes Type-5 LMW-PBPs), therefore, are proposed as the evolutionary origin of class A SBLs (see Figure 1-23 and also section 1.6.2.1 for further details). Additionally, existence of promiscuous hydrolysis activities in class C LMW-PBPs strengthen their role as evolutionary ancestor of class A β-lactamases and more strikingly among these promiscuous activities, β-lactamases activity is also reported (Fröhlich et al., 2021; Moon et al., 2018; J. D. Smith et al., 2013; Urbach et al., 2008). Bellow you can see the conservation of same structural fold between class A β-lactamases (TEM-1) and three class C LMW-PBPs, PBP-A from *T. elongatus*, PBP-4 and PBP-5 of *E. coli* (Figure 1-22).

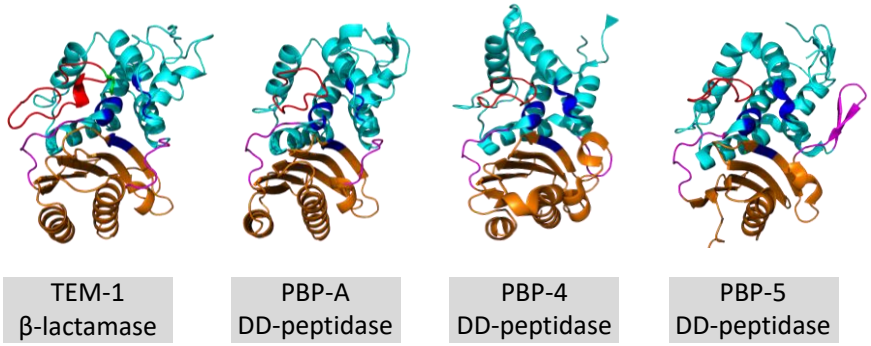


Figure 1-22: Common structural fold of class A β-lactamases such as TEM-1 (PDB: 1BTL) and three class C LMW-PBPs, PBP-A from *T. elongates* (PDB: 2J7V), PBP-4 (PDB: 2EX8) and PBP-5 (PDB: 1NJ4) of *E. coli*. Note that the large extra domains of PBP-4 and PBP-5 are not shown. Same color code of Figure 1-21 is applied. (Structure is visualized and modified by The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC).

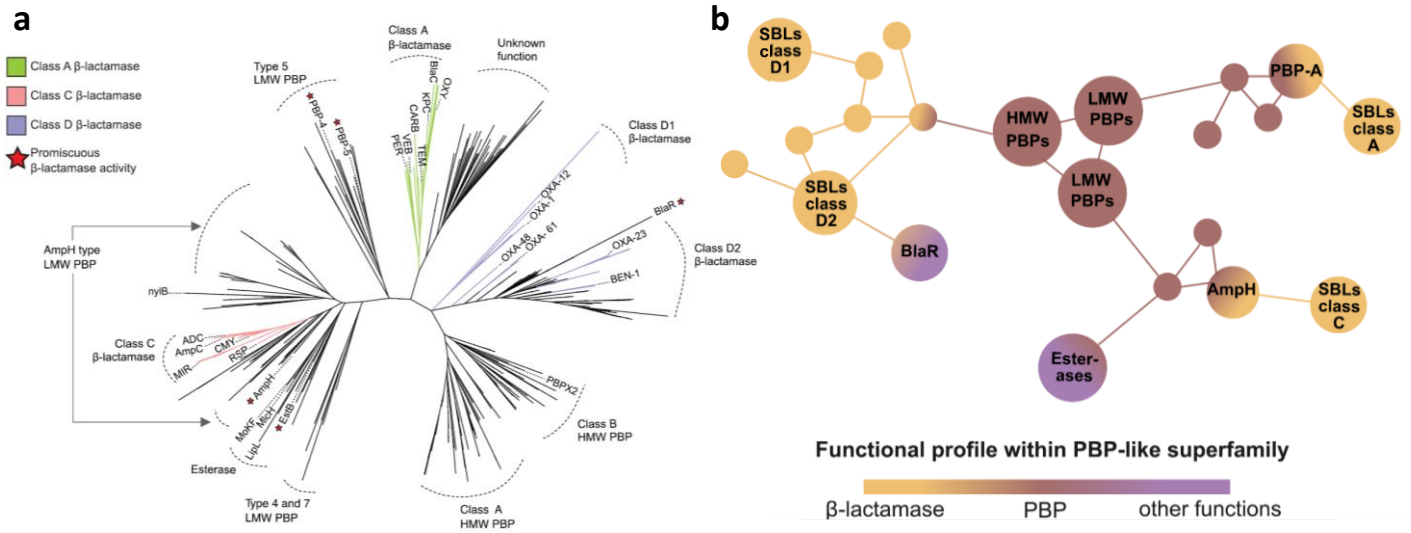


Figure 1-23: Evolutionary relationship within PBP-like superfamily; **a)** phylogenetic tree of PBP-like superfamily shows that, among all PBPs, type-5 LMW-PBPs are the closest relative of class A serine β-lactamases (SBLs). Sequence cluster with unknown functions, which are more close to class A SBLs (up to ~40% sequence identity) and have similar lengths to this class of SBLs, are yet to be characterized and may act evolutionary intermediate between PBP-A (with ~30% sequence identity to class A SBLs) and class A SBLs; **b)** schematic illustration of evolutionary trajectories of the SBLs. Each node represent a cluster of similar sequences that is labeled by their eminent representatives; smaller, unnamed nodes are representing hypothetical or uncharacterized sequences; similarity of sequences is demonstrated by the edges between clusters; color code is representing major functions within the families. PBP in our study belongs to the PBP-A family, which is the closest PBP family, with a known function, to the class A SBLs. (figure adapted from Frohlich et al., 2021).

1.6.2.4 PBP-A

PBP-A from *Thermosynechococcus elongatus* was identified in our lab as a member of cyanobacterial protein family, which was initially annotated as putative class A serine β-lactamases. However, sequence alignments with class A β-lactamases showed a 6-amino acids shorter Ω-loop and lack of the essential Glu166 (Figure 1-24). Further biochemical characterization revealed that PBP-A is rather a DD-carboxypeptidase than a β-lactamase (Urbach et al., 2008). Because PBP-A family has the highest sequence similarity among LMW-PBPs to class A SBLs and also because of its high structural similarity to class A SBLs, and by virtue of its thermostability, PBP-A was chosen as a starting point for directed evolution towards a β-lactamase. The first attempt by Urbach et al. to induce β-lactamase activity in PBP-A by introducing the missing Glu166, at its structurally aligned position of L158 in PBP-A, has only improved by a factor of 50, which failed to generate any resistance to penicillins (Urbach et al., 2009). Many other points mutations were introduced into PBP-A without any improvements of β-lactamase activity (Labarbe (Urbach), 2006).

Several directed evolution experiments were also performed without much success. Error-prone PCR libraries from PBP-A or PBP-A-L158E were built as well as insertional libraires of partially degenerate omega-loop (Labarbe, 2006). These libraries failed to generate penicillin resistance phenotype. Using a codon optimized *pbpA* gene, high expression plasmid and a co-translational signal peptide for periplasmic secretion, variants of PBP-A conferring resistance to 200 mg/l ampicillin were obtained after 5 rounds of epPCR and selection on ampicillin gradient plates (J. Wessel, unpublished). However, these variants completely lost the hydrolytic activity and are degrading penicillin through a slow fragmentation mechanism that is only 10-fold improved compared to PBP-A wt.

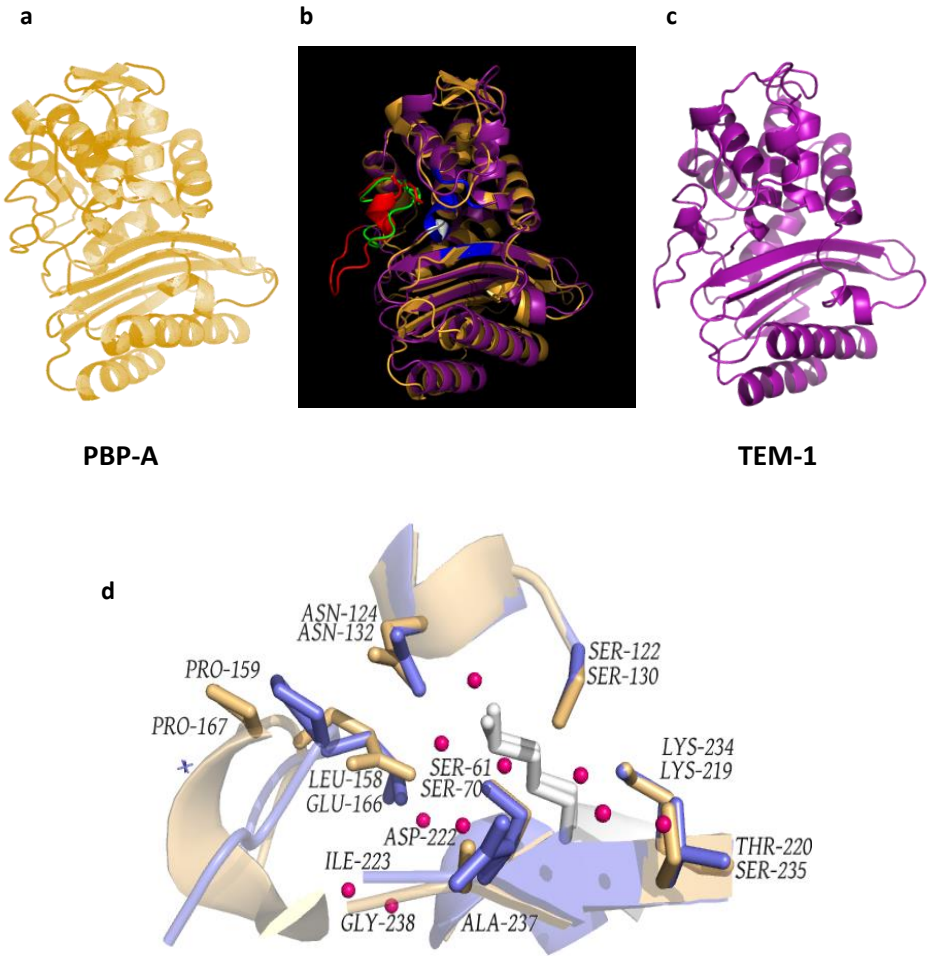


Figure 1-24: High structural similarity between PBP-A and TEM-1; **a)** Structure of PBP-A (PDB: 2J7V); **b)** superimposition of protein structures of PBP-A (colored gold) and TEM-1 (colored purple). The conserved catalytic motifs (SXXK, SXN and KTG(T/S)) are colored blue and the catalytic serine (Ser) is colored white. The major structural difference in the active core, between the two structures, is the Ω-loop, where the shorter (10 residues) Ω-loop of PBP-A is colored green and the longer (16 residues) Ω-loop of TEM-1 is shown by red color; **c)** Structure of TEM-1 (PDB: 1BTL); **d)** Superimposition of the active center of TEM1 and PBP-A.

1.7 Problem statement

Proteins specificities have been optimized through evolution for a given function but, even though it is often overlooked, high specificity may also prevent proteins to do that they should not, by losing ability of extra

(promiscuous) activities. Although promiscuous activities of proteins are interesting starting points for directed evolution campaigns, they may therefore also be associated with toxicity or fitness costs, which can act as a brake on evolution. This may be particularly acute for enzymes, such as hydrolases, that could chemically damage cellular components. Fitness cost associated with such possible interfering activities may be modulated by gene expression level and/or by other physiological and biophysical properties of the proteins such as size, isoelectric point, cellular localization.

As mentioned earlier, type-5 subcategory of class C LMW-PBPs, which includes cyanobacterial PBP-A, shows the highest sequence similarity to class A β-lactamases (Figure 1-23) and share the common structural fold as well as catalytic motifs (see Figure 1-24) (Fröhlich et al., 2021). As β-lactamases and PBPs are acylated similarly with β-lactam antibiotics, the striking masterwork in the evolutionary trajectory from PBP to beta-lactamase is the attainment of ability to hydrolyze the penicilloyl-enzymes, which PBPs are unable. Previous results in our lab indicated that PBP-A imposes some fitness cost upon expression in *E. coli* cells and this fitness cost was suggested as the main obstacle in failed attempts in converting PBP-A into a β-lactamase. The outcomes from the work of Urbach has shown that many faster growing colonies (big colonies) in a library of PBP-A mutants are harboring mutations in important residues of the active site (Carole Urbach's thesis), which indicates a fitness heterogeneity among colonies of library.

A potential origin of such toxicity(s) would be an interference of PBP mutants with the peptidoglycan metabolism. Based on that, Heykel Trabelsi has hypothesized that such interference may even worsen if hydrolytic power of PBP-A is to be increased without losing its capacity to recognize the original D-Ala-D-Ala substrate. In other word, if the DD-peptidase initially loses its native function towards the D-Ala-D-Ala substrates while maintaining its aptitude to recognize penicillin (exclusive PBP), its potential to evolve into a β-lactamase may increase. Therefore, instead of considering the evolutionary trajectory only as a trade-off between DD-peptidase and β-lactamase activities, he tested a

scenario of evolution in which they included the strict penicillin binding property as an intermediate function. Drift experiments to select strict PBP variants (i.e that have lost DD-peptidase activity) by cultivating the libraries without any selective pressure was conducted. A library of ~10^8 clones PBP-A mutants was created by epPCR and was grown for 10 days under low expression conditions and without any particular selection pressure (just growth). Very interestingly, most mutants providing the best fitness to the bacteria had two mutations in the active site, L158Q and E96V. These two residues are not part of the catalytic machinery of PBP-A but their mutations reshape the active site in such a way that the mutant is poorly active on D-Ala-D-Ala based substrates but still very efficiently acylated by penicillin (Heykel Trabelsi's thesis). However, starting from these mutants, iterative rounds of random mutagenesis (epPCR) and selection with increasing concentrations of ampicillin did not allow the emergence of phenotypically active β-lactamases.

Despite the similarities in structural scaffolds and conserved active site motifs between PBP-As and class A β-lactamases, an important difference is the presence of a six-residues shorter Ω-loop flanking the active site of PBP-A (Figure 1-24). A longer Ω-loop in class A β-lactamases is capping the catalytic pocket, which also holds a general base (Glu166) that is responsible for water activation in the deacylation step. In PBP-A, the structurally aligned position is occupied by a leucine (Leu158). Substitution of this leucine into a glutamate results in a 90-fold increase of the β-lactamase activity with a kcat of 10^{-2} s^{-1}, which is still about 10^5-fold slower compared to a natural β-lactamase and this variant is unable of conferring any penicillin resistance to *E. coli*. The L158E mutant was expressed in cytoplasm of *E. coli* and the crystal structure was determined (Carole Urbach's thesis). All active site elements appear to be superimposable with those of a class A beta-lactamase, highlighting our still poor understanding of the structure-to-function relationship underlying enzyme's efficiency.

Through a retroevolutionary study on TEM-1 β-lactamase, we have also previously shown that the length of the Ω-loop is not essential for β-lactamase

activity (Gilles Joachim's Thesis, 2016). Unpublished results from his work suggests that the length of Ω-loop is not a key event in birth of the β-lactamase activity but an ancestral insertion might have occurred for improving the enzyme efficiency. By analyzing primary sequences of class A β-lactamases, traces of an ancestral InDel in the Ω-loop was evaluated, which provided a plausible scenario that guided the genetic engineering of β-lactamase variants with altered Ω-loop length. Then a directed evolution strategy for evaluating the potential of these shorter Ω-loop scaffolds was used to support β-lactamase activity. By studying the biochemical properties of these variants and analyzing these artificial evolutionary trajectories, the role of the Ω-loop in the structure-function relationships of class A β-lactamases was investigated. Through directed evolution a TEM-1 deletant featuring a 7-residue shorter Ω-loop (named Δ7 variant) and conferring high resistance to ampicillin (2500 μg/ml) was obtained (Gilles Joachim's Thesis) indicating that beta-lactamase activity is not incompatible with a shorter loop. Later, we mutated the evolved Δ7 variant (Δ7c) to create longer Ω-loop through duplication of 21 nucleotides in order to evaluate the compatibility between the deletant sequence and an hypothetical tandem duplication event. We observed that the duplication is not impacting the beta-lactamase activity. While the longer length of Ω-loop is not conferring any advantage on *in vitro* hydrolysis of penicillin compared to the evolved Δ7c, the *in vivo* fitness of the *E. coli* host is significantly improved at low temperature (23C) and under low expression conditions, suggesting that the actual Ω-loop length is securing the enzyme against a deleterious property.

In another attempt, Julia Wessel started by improving the secretion of PBP-A into the periplasm of *E. coli* by using a co-translational signal sequence that is preventing the folding of the protein in the cytoplasm. With this construct, she could evolve PBP-A into poorly active β-lactamases through cycles of epPCR and selection on gradient plates of ampicillin. However, the mutants have completely lost hydrolytic activity and they are degrading the β-lactam by an ineffective cleavage reaction not involving water, which is different from the evolutionary trajectory of β-lactamases through hydrolysis of β-lactams.

Overall, the main lessons from these past studies were: (i) expression of PBP-A in *E. coli* is costly, (2) longer loop in TEM-1β-lactamase is not essential for catalytic activity but may improve fitness of the enzyme upon expression in *E. coli* cells. However, the precise origin of the fitness cost was not known and directed evolution under fitness-improving strategies in *E. coli* remained to be explored.

Moreover, when compared to class A β-lactamases, and in addition to the shorter Ω-loop, there is another molecular characteristic that is absent in PBP-A and in PBPs in general: a disulfide bond between positions 77 and 123 or 69 and 238, in the vicinity of the active site found in 60% of all known class A beta-lactamases. We hypothesized that the disulfide bonds may have an evolutionary role other than only stability, particularly a role on conformational control of the active site. Although such role of the disulfide bond of TEM-1 on conformational dynamic is not reported, a conserved disulfide bridge in OXA-1 class D β-lactamases, which is distal to active site (allosteric site), is shown to impact the dynamics of active site pocket (Simakov et al., 2017). An attractive hypothesis is that introduction of disulfide bond affords a robust scaffold for accepting otherwise deleterious mutations by constraining the conformational dynamics of the enzyme and minimizing deleterious promiscuous activities.

1.8 Objectives of the thesis

In this thesis, we aim at shedding light on restraining effects of fitness costs on *in vivo* directed evolution of enzymes. Our primary goal is to investigate possible molecular origin(s) of the fitness cost associated with expression of the PBP-A protein in *E. coli,* and to evaluate various molecular mechanisms that may improve the fitness of the host bacterium upon expression of the protein. Since the carboxypeptidase activity of PBP-A was only observed on small surrogate thioester substrate mimicking D-Alanyl-D-Alanine, we will initially characterize its *in vitro* activity on large peptidoglycan and small peptidic substrates to better identify potential deleterious activities. Then we will evaluate the implication of PBP-A activity on the fitness cost associated with its expression in the periplasm

of *E. coli*. The peptidoglycan of the expression host will be analyzed for detecting eventual defects and the growth of the bacteria will be characterized in various conditions (various pH, osmotic stress, sensitivity to vancomycin). Other parameters that may impact the fitness will be examined: the expression level, the correct localization of the protein, potential misfolding problems and a possible non specific electrostatic interaction between the enzyme and the peptidoglycan. We will also evaluate if the introduction of a disulfide bridge in the position structurally aligned with the 77-123 bridge found in TEM-1 is improving the fitness of the PBP-A expressing host.

Based on these results, directed evolution experiments will be designed as to minimize fitness cost. Libraries of PBP-A will be prepared by epPCR and the conversion into beta-lactamase by selection on ampicillin gradient plates will be attempted under optimized conditions. In case of failure, we will take advantage of our libraries to perform drift experiments in the absence of selection pressure for any activity with the objective of identifying specific mutations that are associated with gain of fitness. Such mutations could then be fixed in new libraries for subsequent directed evolution campaigns.

Chapter 2

Multiple activities of PBP-A from *Thermosynechococcus elongatus* on peptidoglycan and peptidic substrates

Gol Mohammad Dorrazehi, Hervé Degand, Damien Evrard, Waldemar Vollmer and Patrice Soumillion

Abstract

Although promiscuous enzyme activities are considered as the springboard for new functions, it remained largely overlooked that promiscuity might also hinder evolvability by imposing fitness cost to the organism through interfering activity(s) and/or competitive inhibition between different substrates. Understanding the nature of interfering activities and uncovering origin of such fitness costs is crucial to predict evolutionary emergence of enzyme specificity. Previous observations in our lab indicate that expression of a cyanobacterial penicillin binding protein (PBP-A) in the periplasm of *E. coli* imposes a fitness cost, which is hypothesized to be linked to its activity and may explain why multiple attempts to convert it into a beta-lactamase have failed. In this chapter, based on *in vitro* assays with synthetic peptides as well as purified peptidoglycan (PG) and muropeptides (MPs), we show that PBP-A is an enzyme carrying several activities, including DD-carboxypeptidase and DD-endopeptidase activities with a strong preference fo peptides that are amidated on the gamma-

D-Glutamate in position 2. Furthermore, we have evidences of unusual LD-carboxypeptidase and glucosaminidase activities promoted by PBP-A in *in vitro* assays with PG and MPs. These findings tell us that PBP-A is a potentially dangerous enzyme with multiple activities that may impair its evolvability.

2.1 Introduction

The traditional 'lock and key' model of enzyme function has been empirically replaced by the fact that, beside the natural activity that enzymes have evolved for, many of them are able to catalyze non-conventional reactions (Babtie, Tokuriki, & Hollfelder, 2010b; Busto et al., 2010; Copley, 2003; Hult & Berglund, 2007; Khersonsky & Tawfik, 2010; O'Brien & Herschlag, 1999). Despite the notion that enzyme promiscuity can be considered as a potential starting point for further optimization and innovations towards neofunctionalization in enzymes (Khersonsky & Tawfik, 2010; Pandya, Farelli, Dunaway-Mariano, & Allen, 2014), the toll of a generalist or promiscuous enzyme must be dependent on the capacity of the cell to tolerate such extra activities. If a side activity contributes to organismal fitness or, at least, remains neutral, its maintenance would not be costly for the cell and will be welcomed after positive selection, but maintaining a costly or interfering side activity would need urgent rewiring of the metabolic context or compensatory mechanisms (e.g., gene regulation, epistasis, etc), otherwise it will be subjected to purifying selection (O'Loughlin, Greene, & Matsumura, 2006). A good example of such compensatory adaptation against noncanonical activities of enzymes, is the evolution of metabolite repair enzymes, which are suggested to play a vital role in protection of metabolic pathways from undesired interference caused by deleterious enzymatic side activities (Bommer, Van Schaftingen, & Veiga-da-Cunha, 2020; Veiga-da-Cunha, Van Schaftingen, & Bommer, 2020). Nonetheless, costly intermediates in evolutionary trajectories of enzymes suggest a rugged fitness landscape, where optimization of catalysis may causes fall off into deep low-fitness valleys and dead-end routes (Goldsmith & Tawfik, 2017). In such sharp trajectories, urgent adaptations either in cellular or protein level are required so that the protein can survive against purifying

selection. Cellular adaptation against newly born enzyme activities remains largely unexplored, so does the origin of fitness costs associated with side activities.

D-Alanyl-D-Alanine-peptidases (DD-peptidases) are essential enzymes in biosynthesis and maintenance of peptidoglycan and, due to their covalent inhibition by β-lactam antibiotics such as penicillins, they are often called penicillin binding proteins or PBPs (Blumberg & Strominger, 1974; Georgopapadakou & Liu, 1980; Popham & Young, 2003). This inhibition is the result of promiscuous behavior of PBPs that are acylated similarly by beta-lactams and their natural peptidic substrates. On the other hand, serine beta-lactamases (SBLs) are members of the same superfamily of penicillin recognizing proteins (PRPs) but have evolved as hydrolases of beta-lactam antibiotics as a resistance mechanism. Substrate promiscuities in β-lactamases is thought to be a starting point for the evolution of extended spectrum enzymes that confer further antibiotic resistance to second and third-generation beta-lactams (Martínez, Baquero, & Andersson, 2007; Modi et al., 2021; Salverda, De Visser, & Barlow, 2010). Although molecular mechanisms and structural properties of β-lactamases have been extensively studied, their evolutionary studies highly focused on clinically emerged β-lactamases with extended specificities, which may only represent the tip of the iceberg (Fröhlich et al., 2021). In this large PRP superfamily, our work is aiming to fundamentally address the evolutionary origin of serine β-lactamases from PBPs.

It is largely accepted that during the course of evolution, some PBPs were converted to SBLs for bacterial survival against exposure to β-lactams (Bush, 2018; Keshri et al., 2020; Koch, 2000), indicating that substrate promiscuity of PBPs towards β-lactams not only acted as the origin of antibiotic effect, but at the same time became the springboard for evolution of antibiotic resistance by acquiring β-lactam hydrolytic activity. This evolutionary relationship is supported by a conserved acylation mechanism by β-lactams and, despite a low sequence similarity, a conserved structural fold (Fröhlich et al., 2021; Hall & Barlow, 2004; Keshri et al., 2020; Medeiros, 1997). To date, the highest ~30% sequence identity is found between a family of cyanobacterial low molecular weight PBPs, called

PBP-A and class A SBLs (see Figure 1-23) (Urbach et al. 2009). PBP-A from *Thermosynechococcus elongatus* has been characterized and, based on its activity on surrogate thioester substrate, is suggested as a D-Alanyl-D-Alanine-carboxypeptidase (DD-CPase) enzyme (Urbach et al., 2009, 2008). However, while classical LMW-PBPs feature a C-terminal membrane-anchoring, PBP-A presents an extra N-terminal domain (residues 1-92) predicted as a probable transmembrane anchor (Urbach et al., 2009, 2008). In contrast, serine beta-lactamases are soluble enzymes. Therefore, with the aim to study and evolve PBP-A, its N-terminal domain was removed to produce only the soluble PBP domain. Through success and failures in our directed evolution attempts to convert PBP-A into a beta-lactamase, we have accumulated evidences that fitness cost or toxicity, related to PBP-A expression in *E. coli*, is a main obstacle in this conversion.

In this work, we further characterized the activity of PBP-A with small peptidic substrates and purified muropeptides/peptidoglycan. We detected DD- and LD-carboxypeptidase (DD-CPase and LD-CPase) and DD-endopeptidase (DD-EPase) activities and discuss how these multiple activities may hinder conversion of PBP-A into a β-lactamase in directed evolution experiments.

2.2 Results

2.2.1 PBP-A assay with synthetic peptides

In order to characterize PBP-A activity, we initially incubated the purified wt enzyme and the variant Leu158Glu (L158E) endowed with increased hydrolytic activity on penicillin (Urbach et al., 2008) with two synthetic peptides as small substrate mimicks. The pentapeptides are acetylated at N-terminal L-Ala, and feature an amidated or non amidated gamma-D-Glu in position 2, a mimick of meso-diaminopimelic acid (mDAP*) in position 3 and a C-terminal D-Ala-D-Ala group. The mimick of mDAP contains the correct L- and D-centers and differs from mDAP only by a sulfur (S) instead of a methylene (CH_2) in the gamma position of the side chain (Figure 2-1). The peptidic products in each assay were analyzed by liquid chromatography–mass spectrometry (LC–MS).

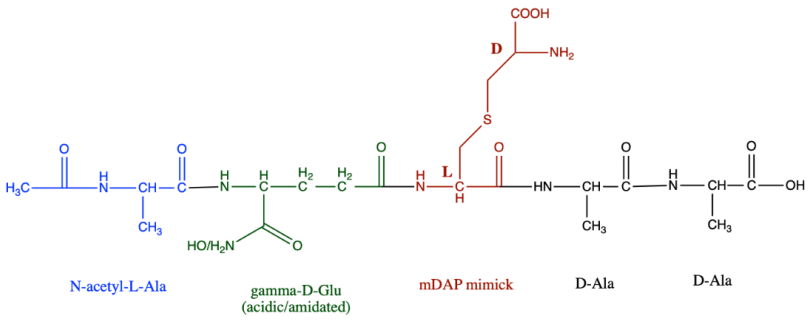


Figure 2-1: Synthetics pentapeptide substrates used in assay with PBP-A enzymes. We designed two pentapeptide (Acetyl–Ala-gamma(D)Glu–mDAP*–(D)Ala–(D)Ala) substrates that contain acetylated N-terminal L-Ala (blue)**,** either having normal (acidic) gamma-D-Glu (isoGlu) or amidated gamma-D-Glu (amidated-isoGlu) (green), a mimick of mDAP (red) and C-terminal D-Ala-D-Ala (black). The mimick of mDAP contains the similar L- and D-centers as mDAP and the only difference with mDAP is a sulfur (S) instead of the second carbon (C) in the side chain.

In assays with synthetic peptides, we observed a strong preference of PBP-A for carboxamide *vs* carboxylate of the gamma-D-Glu (Figure 2-2). The assay was carried out at pH 7.4, where the carboxylic acid is deprotonated and anionic whereas amidated version is neutral. Therefore, we hypothesized that a carboxylic acid may be more similar to amidated isoGlu, and we evaluated the activity of PBP-A with the synthetic substrates at a more acidic pH. Eventhough lowering the pH to 5 result in decrease of activity on amidated peptide, it also allowed the detection of carboxypeptidase activity on the non-amidated one, confirming our hypothesis. and MS spectra of detected peaks in this assays are presented in Figure 2-2.

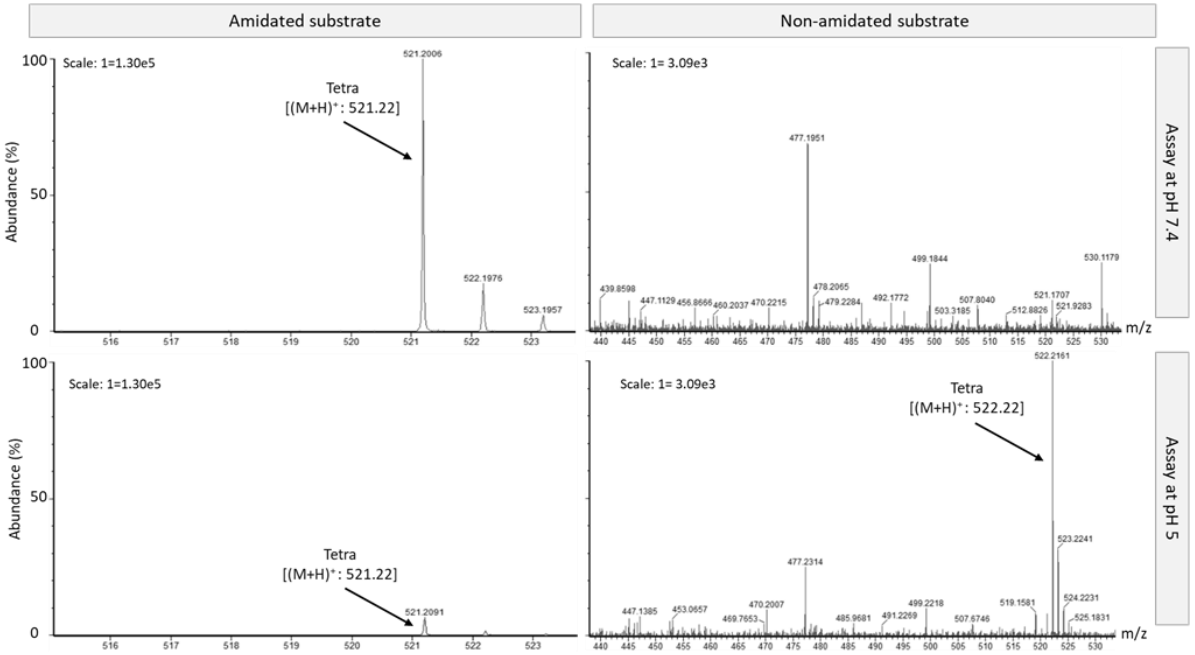


Figure 2-2: Detection amidated tetrtapeptide (Acetyl–Ala-gamma(D)Glu$_{(amidated)}$–mDAP*–(D)Ala [$(M+H)^+$: 521.22]) and non-amidated tetrapeptide (Acetyl–Ala-gamma(D)Glu–mDAP*–(D)Ala [$(M+H)^+$: 522.22]) in assay of PBP-A-wt with non-amidated and amidated pentapeptides, respectively, at two different pH (pH 5 and 7.4). DD-peptidase activity of PBP-A on the amidated peptide is detectable on both acidic and basic pH, while activity on non-amidated pentapeptide is only detectable at acidic pH. However, the lower activity of PBP-A at acidic pH on amidated peptide is probably due to effect of the pH on enzyme efficiency.

To further confirm the DD-peptidase activity of the PBP-A on the amidated peptide, we performed this assay with negative controls. For that, we incubated the peptide with (i) PBP-A in presence of penicillin G (PenG) as the inhibitor, (ii) inactive mutant of PBP-A (PBP-A-S61A), and (iii) no enzyme. Abundance of substrate (Penta) and product (Tetra) detected in MS spectra from these assays are presented in Figure 2-3.

Significant conversion of pentapeptide to tetrapeptide detected in the assay of PBP-A with synthetic pentapeptide, while no conversion upon inhibition of PBP-A by penicillin G (PenG) as well as inactivation of the catalytic serine (Ser61), confirms the previously reported DD-CPase activity of this enzyme (Urbach et al., 2008). Tripeptides and peptidic crosslinks indicative of respective LD-carboxypeptidase and DD-transpeptidase activities were not observed in these assays.

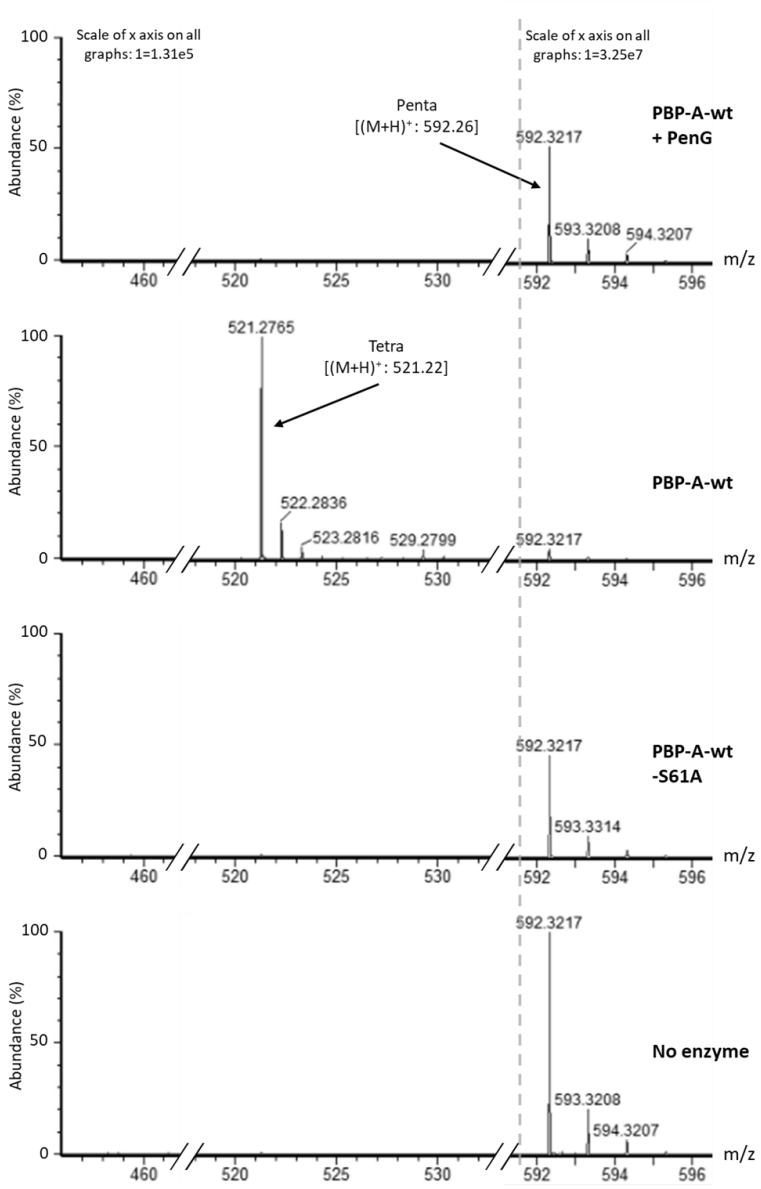


Figure 2-3: Hydrolysis of amidated pentapeptide to tetrapeptide by PBP-A. The spectra represent detection of substrate (amidated pentapeptide, Acetyl–Ala-gamma(D)Glu$_{(amidated)}$–mDAP*–(D)Ala–(D)Ala) [$(M+H)^+$: 592.26]) and product (amidated tetrapeptide, Acetyl–Ala-gamma(D)Glu$_{(amidated)}$–mDAP*–(D)Ala [$(M+H)^+$: 521.22]) in assays of PBP-A-wt in prersense of PenG inhibitor, PBP-A-wt, catalytic mutant (S61A) of PBP-A and a control reaction without enzyme, with amidated pentapeptide. Note that the scale of the spectra after dashed line (for Penta peaks) is 3.25×10^7, while for detection of Tetra is 1.31×10^5.

2.2.2 *In vitro* assay of PBP-A with PG extracts from *E. coli*

PBP-A wt and PBP-A L158E variants were incubated (at pH 7.5) with purified peptidoglycan (PG) extracts from two *E. coli* strains: BW25113Δ6LDT strain, which lacks all Ldt transpeptidases (Kuru et al., 2017) and, therefore, is rich in TetraTetra crosslinks, and CS703-1 strain, which lacks PBPs (PBP 1A, 4, 5, 6 and 70) (Denome, Elf, Henderson, Nelson, & Young, 1999) and thus is rich in penta muropeptides. Samples were digested with cellosyl lysosyme (muramidase) and the released muropeptides were analyzed by HPLC. Qualitative results, representing MS peaks as well as proposed activities and quantitative results for the assays are presented in Figure 2-4 and Figure 2-5, respectively. Detailed quantitative analysis can be found in Appendix 9-1 and Appendix 9-2.

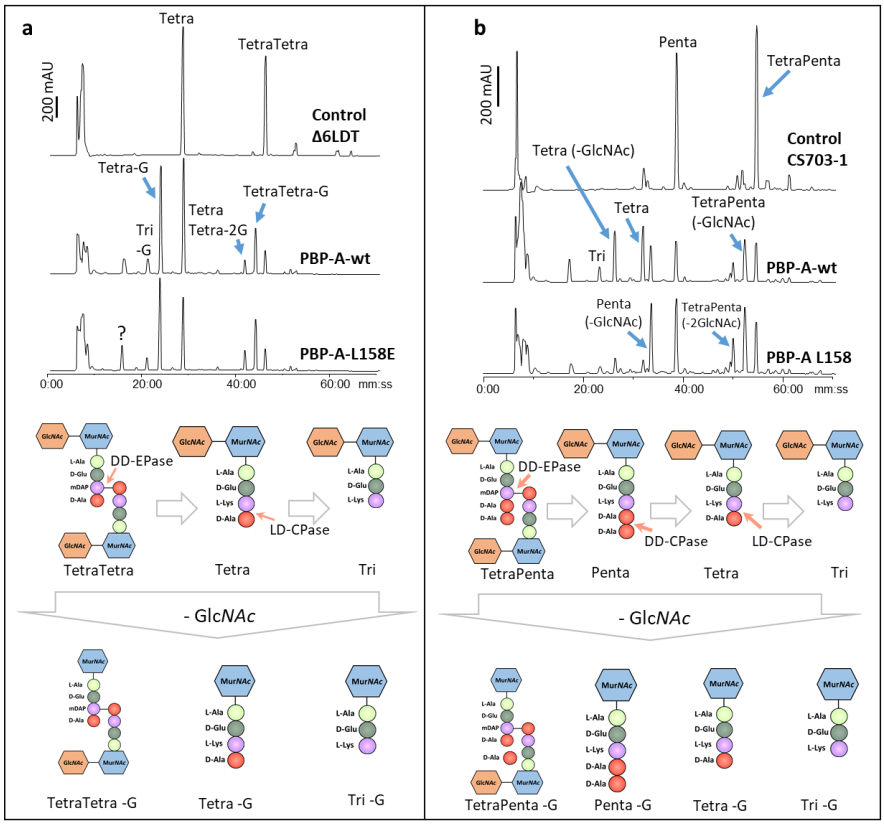


Figure 2-4: MS peaks and presumed activities detected in assays of PBP-A with purified PG of *E. coli*. **a)** Assay with PG from *E. coli* Δ6LDT (rich in TetraTetra). **b)** Assay with PG of *E. coli* CS703-1 (rich in Penta). In both assays, detection of Tri in samples treated

with PBP-A, while no detection in control suggests LD-CPase activity. A Glc*NAc* removal (glucosaminidase) activity is observed in both assays, resulting in cleavage of the bond between Glc*NAc* and Mur*NAc* in a disaccharide unit. -G and -2G represent removal of one or two Glc*NAc* from the muropeptides, respectively. Abundance of all molecules observed in each samples are presented in Appendix 9-1 and Appendix 9-2 and MS analysis data for each peak are available in Appendices 9-4 to 9-17. Lower parts illustrate schematic of the activities observed in these assays.

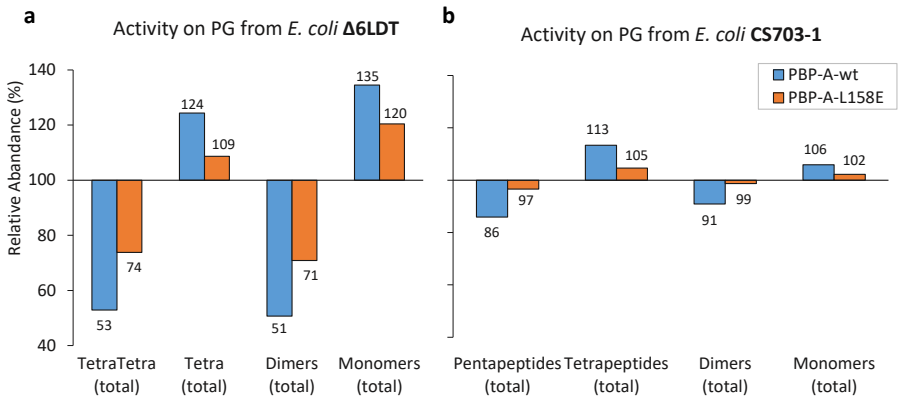


Figure 2-5: Activities of PBP-As on purified peptidoglycan (PG) from *E. coli*. Graphs are representing abundance of peptides detected in the assays of PBP-A-wt and PBP-A-L158E with PG from *E. coli* Δ6LDT and CS703-1 strains relative to the control (no PBP-A) samples in each assay. The assays were carried out at pH 7.5. **a)** Activity of both PBP-As on PG of *E. coli* Δ6LDT (rich in TetraTetra crosslinks) shows a significant decrease in total TetraTetra crosslinks as well as total dimers, while increase in sum of Tetra monomers plus converted Tetra monomers due to LD-CPase activity (Tetra+Tri), which confirms DD-EPase activity. This is further detectable as decrease in total dimers, while increase in total monomers. , **b)** Activity of PBP-As on PG of CS703-1 (rich in Penta) shows significant increase in amount of total tetrapeptides, while decrease in pentapeptides, confirming DD-CPase activity. Additionally, decrease in total dimer crosslinks, while increase in total monomers, confirms DD-EPase activity. Abundance of all molecules in the MS spectra of this assay are available in Appendix 9-1 and Appendix 9-2. HPLC spectra for detected peptides in these assays are available in Appendix 9-3 and MS analysis for each peak are available in Appendix 9-4 to Appendix 9-17.

In these assays, we further confirm the DD-CPase activity but also observed a new DD-EPase activity of PBP-A. DD-CPase activity is confirmed by significant increase in amount of Tetra, while decrease in Penta (Figure 2-4-b). DD-EPase activity is revealed by the decrease in TetraPenta, TetraTetra and in total dimers, and by the increase in total monomers. Such activity will directly reduce PG crosslinking. Both activities are lower for PBP-A-L158E, compare to PBP-A-wt (Figure 2-4).

Unexpectedly, Figure 2-4 also show the detection of tripeptides that may result from LD-CPase activity on tetrapeptides or LD-dipeptidase activity on pentapeptides, as well as the detection of Glc*NAc* removal from Glc*NAc*-Mur*NAc*-peptides. However, when the experiments were performed with shorter incubation time (2 h instead of 6), removal of Glc*NAc* was not detected. Additionnally, Glc*NAc* removal was not observed in assay with TriLysArg muropeptides, on which PBP-A had no activity (Appendix 9-18).

Higher activity of PBP-A on PG of E. coli at acidic pH

Since the PG of *E. coli* is not amidated, we also performed assays under acidic conditions. PBP-A-wt and PBP-A-L158E variants were incubated with purified peptidoglycan (PG) at pH 7.5 and 5.0. A sample without PBP-A treatment was used as control in each condition. Released muropeptides were analysed by HPLC. Results in Figure 2-6 represent abundance of muropeptides in assays with PBP-As relative to the control. Detailed quantitative data are available in Appendix 9-19 and Appendix 9-20.

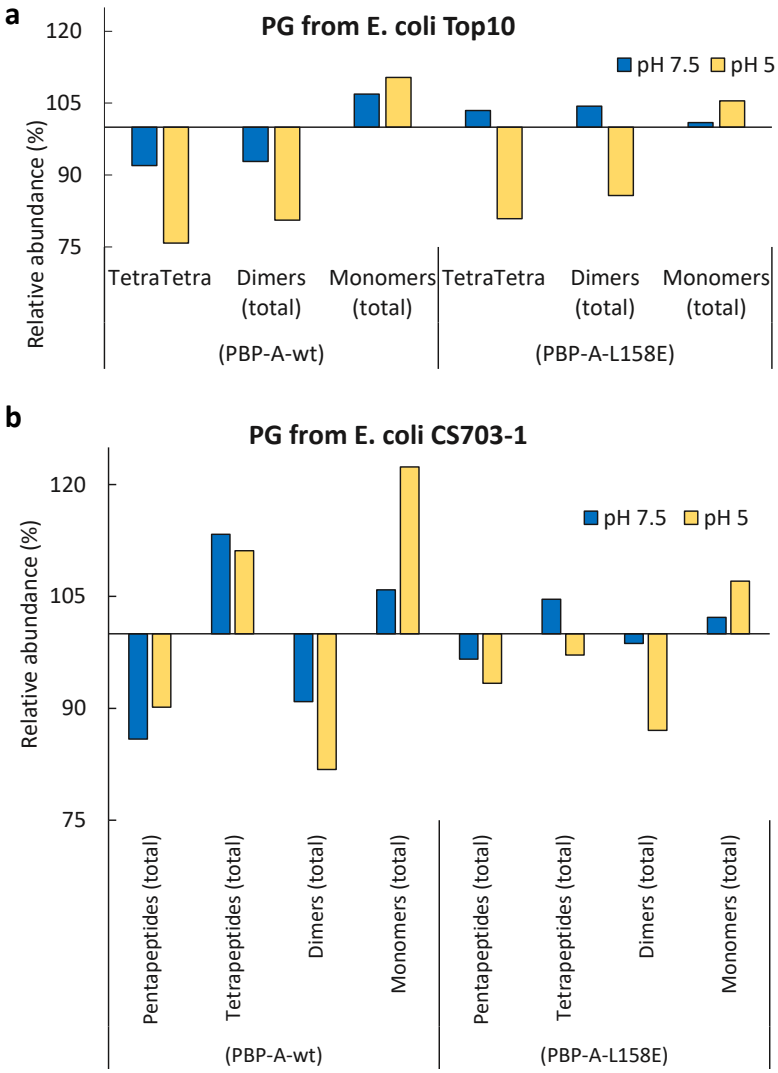


Figure 2-6: *In vitro* activity of PBP-As on PG of *E. coli*. PG from *E. coli* Top10 and CS703-1 strains were incubated with PBP-A-wt and L158E followed by cellosyl digestion and analysis of muropeptides by HPLC. The graphs are representing abundance of different peptides detected in assay of PBP-As with PG, relative to the control (no PBP-A treatment) samples in each assay. **a)** Activity of PBP-As on PG of *E. coli* Top10 at pH 7.5 and 5.0. **b)** Activity of PBP-As on PG of *E. coli* CS703-1 at pH 7.5 and 5.0. In both assays, PBP-As show higher activity on PG at acidic (pH 5) than basic pH (pH 7.5). Abundance of all muropeptides detected at different pH is both assays are available in Appendix 9-19 and Appendix 9-20.

As shown in Figure 2-6, activity of PBP-As on PG of *E. coli* is significantly higher at acidic pH. This effect is particularly traceable on both carpxypeptidase activity

(DD-ECpase) on Penta muropeptides and EPase activity by cleavage of dimer crosslinks.

2.3 Discussion

In a quest to find the origin of fitness cost associated with PBP-A expression in *E. coli*, we hypothesized that PBP-A may act on the peptidoglycan of the *E. coli.* With this aim, we performed *in vitro* assays of PBP-As (wt and L158E) with synthetic peptides as well as purified PG and muropeptides of *E. coli* as substrates. However, due to lack of a convenient kinetic assay, kinetic parameters on efficiency of PBP-A towards different peptidic substrate remains to be determined.

We conducted assays with two synthetic peptides that both are acetylated at N-terminal L-Ala, containing a good mimic of mDAP in position 3 and two D-Ala at C-terminal with a difference in amidated/acidic isoGlu at position 2. While results from these assays confirmed the predicted DD-CPase activity on pentapeptide, we observed a preference of PBP-A for amidated substrate. It is widely known that amidation can occur at α-carboxyl group of D-Glu and the ε-carboxyl group of mDAP (Vollmer, Blanot, et al., 2008). Interestingly, it has been reported by Weckesser and colleagues (Jürgens, Drews, & Weckesser, 1983) that they observed amidation of carboxylic groups in the peptide subunits of cyanobacterial peptidoglycan. However, they did not determine if amidation was on the gamma-D-Glu and/or mesoDAP residue. Since PBP-A in our study is also from a cyanobacterial origin, we hypothesize that the PG of cyanobacteria is amidated on the D-Glu residue and PBP-A is mainly evolved for activity on this substrate. Nonetheless, it remains to be tested whether amidation of ε-carboxyl group of mDAP is affecting PBP-A activity or not.

Amidation of the D-glutamate at position 2 has been reported in Gram-positive species such as *Streptococcus pneumoniae*, *Staphylococcus aureus*, or *Mycobacterium tuberculosis* (DeJesus et al., 2017; Figueiredo et al., 2012; Zapun et al., 2013). The amidation mechanism has been studied in details and reported to be carried out by the essential amidotransferase MurT/GatD complex, where MurT catalyzes

the amidation (via a ligation mechanism similar to Mur ligases) and GatD, which is a glutaminase, provides the ammonia (Morlot et al., 2018; Nöldeke & Stehle, 2019). MurT and GatD are expressed from the same operon and it has been reported to have homologues in cyanobacterium *Nostoc punctiforme* (Münch et al., 2012). However, BLAST analysis of nucleotide and amino acid sequences of both ORFs of SACOL1951 (MurT) and SACOL1950 (GatD) from *Staphylococcus aureus* (subsp. aureus COL, accession number: CP000046.1) on the genome of *T. elongatus* (accession number: BA000039.2), did not reveal presence of homologous proteins.

More interestingly, in assays with synthetic peptides, we observed that PBP-A becomes active on non-amidated substrate at acidic pH. At pH 5, about 10% of the D-Glu should be in its protonated form which is much more similar to the amide in terms of charge and hydrogen bonding pattern. Additionally, the higher activity of PBP-A at acidic pH was also observed in assays with purified PG of *E. coli* (Figure 2-6), which is in agreement with activity of PBP-A on small peptides, because the PG of *E. coli* is not amidated (Jürgens et al., 1983).

Additional *in vitro* assays, performed on PG extracts of *E. coli*, confirmed both DD-CPase activity and further indicate presence of DD-endopeptidase (DD-EPase) activity. The DD-EPase activity is confirmed by cleavage of isopeptidic bridges in tetratetra and tetrapenta dimers between the D-alanyl residue in position 4 and the D-stereospecific side chain of mDAP residue in position 3. Promiscuous activities LMW-PBPs were previousely reported, e.g., among eight endopeptidase paralogues in *E. coli*, PBP4, PBP7 and AmpH serine hydrolase enzymes confer DD-CPase and DD-EPase activities and are involved in maturation, remodeling and recycling of PG (Sauvage et al., 2008b; Voedts et al., 2021). Particularly, similar to PBP-A in our study, PBP4 of *E. coli*, which is reported to possess DD-CPase and DD-EPase activities (Clarke et al., 2009), imposes a fitness cost upon overexpression in *E. coli* (D. E. Nelson & Young, 2001). PBP4 of *Staphylococcus aureus* is also reported to have DD-TPase activity in addition of DD-CPase and DD-EPase activities (Wyke, Ward, Hayes, & Curtis, 1981).

To our knowledge, it is the first time that a LD-CPase activity is observed for a PBP. Based on PBP-A structure (PDB: 2j8y) superimposed with the structure of PBP6 of *E. coli* in complex with a substrate fragment (PDB: 3itb), it is quite surprising since positioning the 3-4 peptide bond in place of the 4-5 peptide bond would generate an important steric clash between the enzyme and the L-Lys side chain in position 3. Such activity would require an alternative positioning of the substrate and probably also a local unfolding of the omega-loop for opening up the pocket of the active site. Moreover, the LD-CPase activity was only detected with PG substrates and not with the small pentapeptides. Thus, further verification of this activity are required (e.g., assay in presence of inhibitor (PenG)). Since our enzymes are expressed in the periplasm before purification, LD-CPase activity may originate from a contaminating enzyme.

Surprisingly, glucosaminidase activities are observed under assays of PBP-As with cellosyl-treated PG from delta6-LDT and CS703-1 strains (Figure 2-4). However, the Glc*NAc* removal activity was not observed in analysis of PG from PBP-A-expressing *E. coli* cells (*in vivo* assays in next chapter) and not detected either with TriLysArg muropeptides (Appendix 9-18) which is not recognized by PBP-A as substrate. This suggests that cellosyl may be involved but the activity should be somehow promoted by PBP-As since this activity was never observed with cellosyl-digested muropeptides alone. However, a previous study by Shoichet et al., has reported that in the structure of preacylation complex of PBP6 from *E. coli* (PDB: 3ITB) with a synthetic peptidoglycan surrogate, NAM-(L-Ala-D-isoGlu-L-Lys-D-Ala-D-Ala), the substrate is recognized by a dimer of PBP6, with one active site occupied as expected by D-Ala-D-Ala but the active site of the other subunit is occupied by the NAM residue (Y. Chen et al., 2009). This report and our results might point towards a possible *in vitro* glucosaminidase activity by PBPs on small muropeptides featuring disaccharides groups but may be less relevant on larger glycan chains *in vivo*.

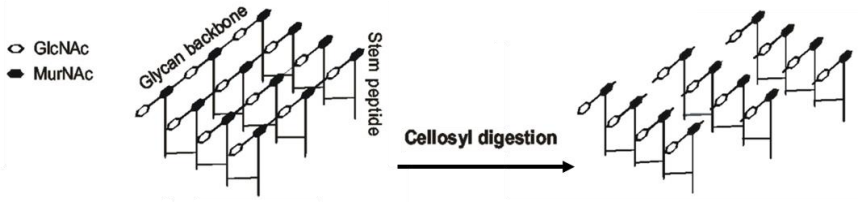


Figure 2-7: Schematic illustration of PG digestion by cellosyl muramidase. Similar to activity of lysozymes, cellosyl cleaves the beta-1,4-glycosidic bond between MurNAc and Glc*NAc*, producing Glc*NAc*-Mur*NAc* disaccharide from the polymeric PG backbone (Myhre et al., 2004). Glc*NAc* removal activities in assay of PBP-A was only detected in *in vitro* assay of with cellosyl-digested PG (Figure 2-4) and not in analysis of PG profile from PBP-A-expressing cells (see section 3.2.4, chapter 3).

2.4 Conclusion

In this chapter, we discovered that PBP-A has multiple *in vitro* activities against peptide/muropeptide substrates and shows a strong preference for muropeptides that are amidated on D-Glu in position 2. The expected DD-CPase of PBP-A was observed in two different assays with synthetic peptides and purified muropeptides/PG. Furthermore, DD-EPase activity was also detected and may reduce the crosslinking of PG *in vivo* and at least partially explain the fitness cost problems. Since activity on non amidated PG is increased at acidic pH, pH-dependent fitness problem may also be expected when PBP-A is expressed in the periplasm of *E. coli*.

2.5 Contribution of authors:

H.D kindly carried out the MS analysis of PBP-A assays with synthetic peptides. D.E helped in enzyme purifications. W.V generously collaborated on the assays of PBP-A on peptidoglycan and muropeptides as well as HPLC and MS analysis and discussions on results of these assays. The rest of the works including development of theory, preparation of mutants of PBP-A, assays, discussions on data and preparation of manuscript are carried out by G.M.D and supervised by P.S.

2.6 Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722610 (ES-Cat).

2.7 Material and methods

2.7.1 Strains and plasmids

Genes of *pbp-a-wt* and *pbp-a-L158E* were cloned into low-copy pBAD43 plasmid and expressed under araBAD promoter in *E. coli* Top10 strain (see 7.1.2 for details). Sequence of *pbp-a* gene and map of pBAD43 plasmid are available in general methodology chapter.

Catalytic mutant of PBP-A (mutated at Ser61) was prepared by site directed mutagenesis via QuickLib protocol (Galka et al. 2017). Following primers were used for the site directed mutagenesis of serine at position 61 of PBP-A into alanine. The substitution (T→G) and, overlapping sequences between two primers are indicated as underlined and bold, respectively:

PBP-S61A_fw:

AATGTCGGCG**GTGACCAAGTGTTTCCGG**CGGCCGCTACCATTAAA TTCCCGATCCTGG

PBP-S61A_rv: **CCGGAAACACTTGGTCAC**

2.7.2 Purification of PBP-A proteins

PBP-A proteins were purified from total soluble lysate (see purification in general methodology) by affinity chromatography thanks to the 6-histidine (6xHis) tag fused to the C-terminal of the PBP-A proteins.

2.7.3 PBP-A activity assay using synthetic peptides as substrates

A reaction containing 5 µL of enzymes (0.15-0.2 mg/mL), 10 µL of 5 mM synthetic pentapeptides (Acetyl–Ala-gamma(D)Glu–mDAP*–(D)Ala–(D)Ala) or amidated pentapeptide (Acetyl–Ala-gamma(D)Glu[amidated]–mDAP*–(D)Ala–(D)Ala) and 10 µL HEPES buffer (20 mM, pH 7.4), incubated for 1 hour at room temperature, followed by analysis by nanoUPLC-MS. For inhibition of PBP-A, 10 µL of 100 mM penicillin G (PenG), prepared in HEPES buffer (20 mM, pH 7.4), was used instead of the buffer in the above-mentioned reaction.

2.7.4 Analysis of peptides by nanoUPLC-MS

In-solution samples are prepared by adding 1 % Trifluoroacetic acid (TFA) to the samples to reach pH below 7.

2.7.4.1 Peptide separation using nanoUPLC

Peptide mixture was separated by reverse phase chromatography on a NanoACQUITY UPLC MClass system (Waters) working with MassLynx V4.1 (Waters) software. 5 µL of each were injected on a trap C18, 100Å 5 µm, 180 µm x 20 mm column (Waters) and desalted using isocratic conditions with at a flow rate of 15 µL/min using a 99% formic acid and 1% (v/v) ACN buffer for 3 min. Peptide mixture was subjected to reverse phase chromatography on a C18, 100Å 1.8 µm, 75 µm x 150 mm column (Waters) PepMap for 35 min at 35°C at a flow rate of 300 nL/min using a two part linear gradient from 1% (v/v) ACN, 0.1 % formic acid to 40 % (v/v)) ACN, 0.1 % formic acid for 15 min and from 40% (v/v) ACN, 0.1 % formic acid to 85 % (v/v)) ACN, 0.1 % formic acid for 10 min. The column was re-equilibrated at initial conditions after washing 30 min at 85% (v/v)) ACN, 0.1 % formic acid at a flow rate of 300 nL/min. For online LC-MS analysis, the nanoUPLC was coupled to the mass spectrometer through a nano-electrospray ionization (nanoESI) source emitter.

2.7.4.2 LC-QTOF-MS/MS Analysis (DDA)

DDA(Data Dependent Analysis) analysis were performed on an SYNAPT G2-Si high definition mass spectrometer (Waters) equipped with a NanoLockSpray dual electrospray ion source (Waters).Precut fused silica PicoTip^{R} Emitters for nanoelectrospray, outer diameters: 360 µm; inner diameter: 20 µm; 10 µm tip; 2.5" length (Waters) were used for samples and Precut fused silica TicoTip^{R} Emitters for nanoelectrospray, outer diameters : 360 µm; inner diameter: 20 µm; 2.5" length (Waters) were used for the lock mass.solution. The eluent was sprayed at a spray voltage of 2.8 kV with a sampling cone voltage of 25 V and a source offset of 30 V. The source temperature was set to 80°C. The cone gas flow was 20 liters/h with a nano flow gas pressure of 0.4 bar and the purge gas was turned off. The SYNAPT G2Si instrument was operated in DDA (data-dependent mode), automatically switching between MS and MS2. Full scan MS and MS2 spectra (m/z 50- 2000) were acquired directly after injection to 35 min in resolution mode (20,000 resolution FWHM at m/z 400) with a scan time of 0.1

sec. Tandem mass spectra of up to 10 precursors were generated in the trapping region of the ion mobility cell by using a collision energy ramp from 17/19 V (low mass, start/end) to up to 65/75 V (high mass,start/end). Charged ions (+1,+2, +3, +4) are selected to be submitted to the MSMS fragmentation over the m/z range from 50 to 2000 with a scan time of 0.25 sec. For the post-acquisition lock mass correction of the data in the MS method, the doubly charged monoisotopic ion of [Glu[1]]-fibrinopeptide B was used at 100 fmol/µl using the reference sprayer of the nanoESI source with a frequency of 30 s at 0.5 µl/min into the mass spectrometer.

2.7.4.3 ESI-QTOF data processing

Data were processed with MassLynx V4.1(Waters) . Spectrum for each sample were extracted from the TIC (Total Ion Chromatogram) and submitted to a search of masses of interest.

2.7.5 Activity assay on purified PG or muropeptides of *E. coli*

HPLC based activity assays were carried out in a final volume of 50 µL and in buffer containing 20 mM buffer agent (HEPES or NaAcetetate), 100 mM NaCl, 0.05% Trition-X 100 (reduced), 10 µL of the substrate and 5 µM of protein(s). HEPES/NaOH was used for reactions at pH of 7.5, NaAcetate was used for reactions at pH 5.0. The reaction mixture was incubated for 2 hours in a thermoshaker at 37°C, 900 rpm for the indicated time. The reaction was stopped by boiling the samples for 10 min at 100°C. For standard activity assays on PG, purified PG sacculi was added to the reaction mixture and after the incubation the cellosyl digestion followed overnight. Activity assays on muropeptides were performed on predigested PG sacculi with cellosyl. The pH value of the muropeptides was adjusted before they were added to the reaction mixture. Purified, cellosyl-digested, non-reduced and desalted muropeptide TriLysArg was a gift from J.-V. Höltje, MPI Tübingen, Germany. The reaction products were reduced with sodium borohydride and analysed by reversed-phase HPLC (see below).

2.7.5.1 Assays of PBP-As with PG of E. coli

Isolated PG from the *E. coli* Top10 and CS703-1 strains were used in these assays. Assay condition: a mixture containing 20 mM NaAcetate pH 5, 100 mM NaCl, 0.05% Trition X-100 and 5 μM of each PBP-A protein, followed by 2 or 6 hours incubation at 37°C. After boiling, the samples were digested overnight with cellosyl, reduced, and the released muropeptides were analyzed by reversed phase HPLC using buffer A (50 mM NaP (pH 4.31), + 10 ul NaN_3) and buffer B (75 mM NaP (pH 4.95), 15 % Methanol, + 10 ul NaN_3 and monitoring absorbance at 205 nm.

2.7.5.2 Assay of PBP-As with muropeptides

Here we used muropeptides (products of cellosyl digestion on PG) or anhydromuropeptides (products of SLT digestion) as substrates. A 1:1 mixture of *E. coli* CS703-1 and *E. coli* BW25113Δ6LDT, which was either pre-digested with cellosyl or soluble lytic transglycosylase (SLT), was used with 5 μM of each PBP-A protein under assay condition of 20 mM HEPES/NaOH pH 7.5, 100 mM NaCl, 0.05% Trition X-100 and incubated for 6 hour at 37°C. The muropeptides were reduced while the anhydromuropeptides were not reduced, and the pH was adjusted to pH 4 with phosphoric acid.

2.7.6 Preparation of PG and muropeptides (cellosyl products)

PG sacculi were isolated from *E. coli* cells as described (Glauner, 1988). PG sacculi were isolated from *E. coli* cells as described (Glauner, 1988). Purified PG from *E. coli* CS703-1 and BW25113Δ6LDT strains were mixed with cellosyl (0.5 μg/mL) in 80 mM sodium phosphate, pH 4.8 and incubated overnight in a thermal shaker at 37°C and 900 rpm. The reaction was stopped by boiling and the sample was centrifuged for 10 min at 10.000 × *g* (ambient temperature) to separate the soluble muropeptides (supernatant) from insoluble material. The muropeptides were either used as substrate in enzymatic activity assays or analysed by reversed-phase HPLC.

2.7.7 Reduction of muropeptides with sodium borohydride

Prior to HPLC analysis, muropeptides were reduced with sodium borohydride (Glauner, 1988). Muropeptides were mixed with the same volume of 0.5 M sodium borate pH 9.0 in a 2 mL Eppendorf tube with a punctured lid. The reduction process was started by adding ~ 1 mg of sodium borohydride powder with a spatula. The samples were centrifuged for 20 min at 2000 × *g* at ambient temperature during the reduction. The pH value of the samples was adjusted to 4-5 with phosphoric acid, the sample was centrifuged (5 min, 12000 × *g*, ambient temperature) and the supernatant was analysed by HPLC.

2.7.8 Reversed-phase HPLC analysis of muropeptides

Reduced muropeptides were separated on a Prontosil 120-3-C18-AQ 3 μm reversed-phase HPLC column (Bischoff) as described (Glauner, 1988). The column was pre-equilibrated with solvent A (50 mM sodium phosphate, pH 4.31, 0.0001% NaN_3). Muropeptides were separated at 55°C with a 90 min linear gradient to solvent B (75 mM sodium phosphate, pH 4.95, 15% (or 30% methanol to facilitate elution of all anhydromuropeptides)). Muropeptides were detected by UV absorbance at 205 nm and analysed by the Laura V4.2.11.129 software (LabLogic System Ltd.).

Chapter 3

Fitness costs associated with PBP-A expression in *E. coli*

Gol Mohammad Dorrazehi, Waldemar Vollmer and Patrice Soumillion

Abstract

Deciphering the origin of fitness costs related to protein expression is crucial to understand their evolutionary behavior and evolvability. However, in directed evolution experiments, such fitness costs are mostly overlooked by relating them to biophysical properties of proteins, e.g., protein misfolding and/or aggregation, while less is known about potential fitness costs due to activity or interfering interaction in physiological conditions. Based on previous observations that expression, even at low level, of a cyanobacterial DD-peptidase (PBP-A) is associated with growth defect of the *E. coli* host, we performed phenotypic characterization and investigated possible origin(s) of the fitness costs. We initially found that the fitness cost is (i) correlated with expression level and activity of PBP-A, (ii) more pronounced at acidic pH, and (iii) associated with an increased sensitivity to osmotic stress and vancomycin. Additionally, decreasing the isoelectric point (pI) of PBP-A by mutating charged residues at the surface, showed a beneficial effect on the fitness. Beyond that, various parameters that

may impact the fitness such as periplasmic addressing and possible interfering activity or interaction were examined. Mass spectrometry (MS) analyses revealed reduction in peptidoglycan (PG) crosslinking due to the carboxypeptidase and endopeptidase activities of PBP-A that were previously observed *in vitro* on muropeptide substrates. These results are casting light on importance of biochemical properties as well as physiological behaviors of proteins in directed evolution campaigns.

3.1 Introduction

In evolutionary fitness landscape metaphor, a negative trade-off between an original phenotype evolving towards a new phenotype will be represented as a sharp trajectory connecting two peaks in the adaptive landscape (see Figure 1-2), with the low fitness valleys highlighting a cost to be paid to cross the trajectory. In such 'sharp' evolutionary trajectories valleys have to be surrounded to avoid death-causing environment that leeds to longer trajectory but without downhill moves, in order to reach another fitness/function peak (Figure 3 1). Such detour could be provided by compensatory/adaptive mechanisms such as intra- or inter-genic epistatic mutations, rewiring the affected metabolism, and/or regulation of transcription/translation to bypass the cost. Mutations causing downhill moves could be tolerated as long as the fitness cost is not too deleterious and can be maintained for the time necessary to acquire the next uphill mutation. The fitness loss in sharp trajectories may be associated with metabolic burden, misfolding or interfering activities of proteins. Natural evolution takes advantage of various strategies, such as regulatory mechanism, mutational robustness, induction of molecular chaperones, intra- or intergenic epistasis and/or positive mutations, to minimize fitness costs in sharp trajectories and to enhance the probability of selection for low-fitness variants.

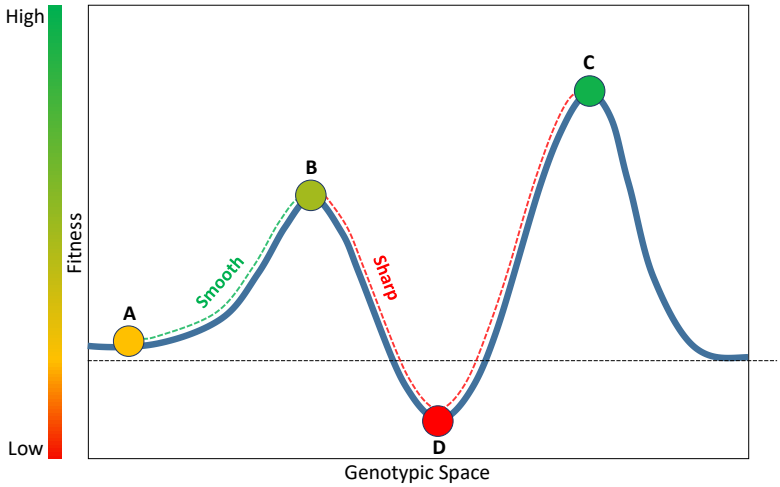


Figure 3-1: 'Sharp' and 'smooth' evolutionary trajectories. In conceptualization of fitness landscapes, where altitudes are representing the fitness and coordinates in the floor are representative of genotypic space, a trajectory would be called 'sharp', if it faces a fitness barrier (a deep fitness valley e.g., from variant 'B' to 'C', shown in red dashed line), or 'smooth' if the trajectory contains uphill climbing (e.g., from variant 'A' to 'B', shown in green dashed line). The color code in left side can be assumed as fitness spectrum from low (red) to high (green). The black horizontal dashed line, represent a conceptual tolerance threshold below which the fitness cost becomes deleterious (see 1.2.2 in Chapter 1 for more information about fitness landscape).

Several strategies, e.g., co-expression/induction of molecular chaperones, fusion to solubilization tags, and reduction in cellular protein expression level have been used by molecular biologists to enhance the fitness of hosts expressing evolving proteins (Costa, Almeida, Castro, & Domingues, 2014; de Marco, Deuerling, Mogk, Tomoyasu, & Bukau, 2007; Knappik, Krebber, & Plückthun, 1993; LaVallie & McCoy, 1995; Schlapschy & Skerra, 2011; Yasukawa et al., 1995). On the other hand, optimization of codon usage, temperature, time and duration of induction and concentration of inducer as well as changing promoters, are among the most common strategies to optimize expression yield (Packiam, Ramanan, Ooi, Krishnaswamy, & Tey, 2020; Trösemeier et al., 2019). Although these parameters are effective in the total rate of protein production. In such case, an alternative approach is to control the initial copy number of the genetic constructs (M. E. Lee, DeLoache, Cervantes, & Dueber, 2015; Nadler, Bracharz, & Kabisch, 2019), which correlates with the expression level of protein(s) of

interest, without compromising total cellular protein expression. Nonetheless, effects associated to the biochemical and physiological properties of a given protein, such as its activity, specific interaction, electrostatic behavior and/or subcellular localization, mostly remain unexplored. Here, we performed phenotypic characterization of *E. coli* expressing a periplasmic cyanobacterial murein hydrolase, called PBP-A, under different pH, osmotic and antibiotic stress.

Evolutionary rate of proteins is also influenced by their expression level as highly expressed proteins are known to evolve slower that their low expressed counterparts (Drummond, Bloom, Adami, Wilke, & Arnold, 2005; Pál, Papp, & Hurst, 2001), which is thought to be mostly related to the protein biophysics and global selection against the deleterious fitness impact of misfolded proteins (Serohijos, Rimas, & Shakhnovich, 2012a). Recently another factor referred as gene copy number has come under attention in regards to their evolution. Studies have highlighted the bias caused by variability in copy number of a gene on corresponding phenotype(s) and their evolution (Dorant et al., 2020; Rodríguez-Beltrán et al., 2020; Ryan et al., 2014). These biases have roots in either expression noise, genetic dominance effect or genotypic/phenotypic variation. This becomes more crucial for *in vivo* directed evolution campaigns since the genetic libraries are mostly prepared in plasmids and selection is carried out in living cells. In such system variability in plasmid copy number between cells would generate heterogeneity in expression and/or activity level of proteins. Moreover, although plasmid libraries are relatively easier to prepare, they endow constraints such as the need for antibiotic selection markers, fitness effects related to plasmid carriage (Alonso-del Valle et al., 2021), genetic dominance effect of variants in the same cell (Rodríguez-Beltrán et al., 2020). An alternative approach to minimize these biases, is to express the variants from chromosomal-integrated single copy libraries. A study by Ryan et al. (2014) has compared the expression between plasmid and chromosomal expression of two fluorescent proteins in yeast and shown that variability in total expression is highly dependent on copy number. As the copynumber of plasmid varies between cells in an isogenic

population, so does the expression level, whereas, on the other hand, the variability in expression was sgnificatly reduced by expressing both of the fluorescent proteins from the genome. The effect of copy number on potential cytotoxicity due to the recombinant protein activity may add up and is generally not considered. In this chapter, we also investigate the effect of copy number on a fitness cost associated with expression of PBP-A. For that, we express the *pbp-A* gene in periplasm of *E. coli* under medium-copy (15-20 copies/cell), low-copy (~5 copies/cell) and single-copy under induced and basal expression.

Bacterial cell wall is an essential scaffold surrounding the plasma membrane that maintains the cell shape and integrity. In Gram-negative bacteria, it consists of an outer membrane (OM) lipid bilayer barrier and the periplasmic space, which contains the peptidoglycan (PG) mesh-like structure (Chattaway, 1982; Dörr et al., 2019; May & Grabowicz, 2018; Nikaido, 2003a). The external bilayer of OM is an asymmetric bilayer that harbors lipopolysaccharides (LPS) in the outer leaflet and phospholipids in its inner side. OM of Gram-negative bacteria plays an important role in isolation and protection of the cells from external environment by providing an impermeable barrier surrounding the periplasm. However, OM facilitates the trafficking of essential molecules through β-barrel porins embedded in the bilayer (Kamio & Nikaido, 1976; Muheim et al., 2017; Sperandeo et al., 2019). Small-scaffold molecules (molecular weight less than 600 Da), such as β-lactam antibiotics, presumably permeate through the non-specific porins by passive diffusion and enter to the periplasmic space, while large scaffold (larger than 600 Da) antibiotics such as vancomycin that exceed the size exclusion of the β-barrel porins, would not pass through the outer membrane unless it is somehow disturbed (Figure 3-2-a) (Muheim et al., 2017).

In *Escherichia coli*, OM is covalently cross-linked to PG through lipoprotein Lpp (also called Braun's protein), which is bound by it's C-terminal lysine side chain to diaminopimelic acid (DAP) of murein and is integrated into the inner layer of the OM through it's N-terminal lipid anchor (Braun & Bosch, 1972; Braun & Rehn, 1969; Braun & Hantke, 2019). Three highly similar L-D-transpeptidases in

E. coli, namely ErfK, YbiS, and YcfS (also known as LdtA, LdtB, LdtC, respectively), are in charge of this anchoring (Figure 3-2-b) (Magnet et al., 2007).

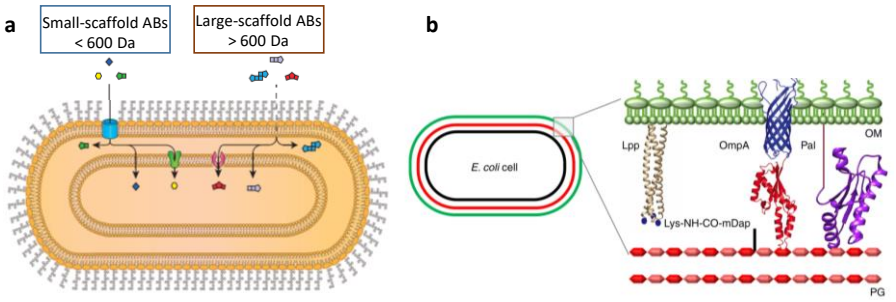


Figure 3-2: Outer membrane (OM) of *E. coli* as a permeability barrier against external environment. **a)** diffusion of antibiotics (ABs) across the cell envelope of Gram-negative bacteria. Generally, antibiotics with molecular weight smaller than 600 Da (small-scaffold ABs) can easily pass the size exclusion of OM porins, while ABs with larger molecular weight 600 Da (small-scaffold ABs) inefficiently diffuse through the OM (Muheim et al. 2017). **b)** Schematic illustration of Lpp-PG cross-link in *E. coli* cell wall. The left part depicts the three layers of cell wall in a general scheme of *E. coli* cell. The cell wall consists of the inner membrane (IM, black) surrounding the cytoplasm, periplasmic space where the peptidoglycan mesh is presented (PG, red), and the outer membrane (OM, green). In the right part, magnified schematic view of the outer layer of cell wall is depicted. The only covalently bond lipoprotein to the PG is Lpp, which cross-links to PG through its C-terminus and while integrated into the OM by a lipid part linked to a cysteine in its N-terminus. OmpA is an integral membrane protein composed of a β-barrel domain embedded in the OM bilayer and a periplasmic domain that facilitate a noncovalent interaction with the PG. Pal is a lipoprotein involved in noncovalent interaction of OM with the PG. (Figure adapted from Mathelié-Guinlet et al. 2020).

The unique PG structure is composed of glycan strands made of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc), which are cross-linked via short peptide stems attached to the MurNAc residues (Figure 3-2) (Vollmer & Bertsche, 2008; Vollmer, Blanot, et al., 2008). In *E. coli*, the stem peptide is initially synthesized as L-Ala–D-iGlu–mDAP–D-Ala–D-Ala pentapeptide, which carries a meso-diaminopimelic acid (mDAP or m-A_2pm) at position 3 (Figure 3-2-a) with an alpha carbon in the L-configuration and a side chain α-carbon in the D-configuration (Pazos & Peters, 2019b; Vollmer & Bertsche, 2008). Glycan strands are generally cross-linked via the carboxyl group of D-Ala at position 4 and the ε-amino group of the mDAP at position 3 (Figure 3-2). This reaction is carried out by D-Alanyl-D-alanine transpeptidase enzymes

(DD-TPases) recognizing D-Ala-D-Ala motif as substrate and cleaving the terminal D-Ala during the transpeptidation. Less frequent, bridging between the main and side chains of two mDAP residues is catalyzed by LD-TPases that will release the D-Ala in position 4. Besides crosslinking, muropeptides are also trimmed by various carboxypeptidases (DD-, LD- and DL-CPases)and cleaved by endopeptidases (DD- and LD-EPases) and amidases (Figure 3-3), while glycan chains are cleaved by lytic transglycosylases (LT) (Egan et al., 2020; Vollmer, Joris, Charlier, & Foster, 2008). β-lactam antibiotics share structural similarities with D-Alanyl-D-Alanine substrate and react with DD-peptidases to form stable acyl-enzymes, leading to their inhibition and bacterial lysis (André Zapun, Carlos Contreras-Martel, Thierry Vernet, 2008). Because of that, DD-peptidases are also called penicillin binding proteins or PBPs. On the other hand, serine beta-lactamases belong to the same superfamily of penicillin recognizing enzymes but have acquired the capacity to hydrolyze penicillins very efficiently, therefore providing antibiotic resistance to their expression hosts. PBP-A, from *Thermosynechococcus elongatus*, is a PBP that is closely related to class A beta-lactamases. Its recombinant expression in the periplasm of *E. coli* has been previously shown to impose a fitness cost to the bacteria and it is thought to be the main obstacle in its directed evolution towards a β-lactamase.

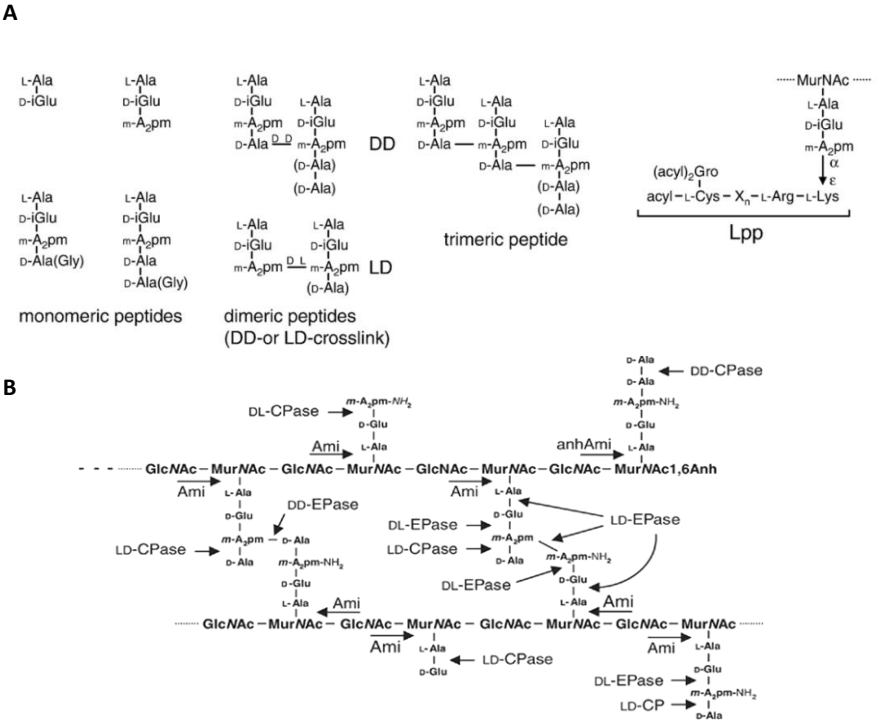


Figure 3-3: **A)** Different monomeric and crosslinked peptides found in PG structure. On the right side the cross-link of Braun's lipoprotein, Lpp, with PG is presented. **B)** Target hydrolytic sites of the different murein hydrolases in peptidoglycan of *E. coli* or *B. subtilis*. Amide bonds crosslinking MurNAc and the L-alanine of the stem peptide are cleaved by N-acetylmuramyl-L-alanine amidases (Ami); specific amidases (anhAmi) target the terminal 1,6-anhydroMurNAc residues. Amide bonds between and within peptides are cleaved by endopeptidases (DD-EPase, LD-EPase or DL-EPase). Both cross-linked and monomeric peptides are targeted by these enzymes. Peptide bonds between the C-terminal D- or L-amino acids are cleaved by carboxypeptidases (DD-CPase, LD-CPase, DL-CPase). (Figure adapted from Vollmer and Bertsche, 2007; Vollmer et al. 2007; See also: Egan et al. 2020).

In this chapter, we validated the implication of PBP-A activity to the fitness cost associated with its expression in *E. coli* and characterized the phenotypes of strains producing the wild-type PBP-A (PBP-A-wt) and PBP-A-L158E mutant that features higher hydrolytic activity on penicillins (Urbach et al., 2008). We also investigated the effect of gene copy number of the fitness cost and, on a quest to precisely identify the origin of fitness cost, we analyzed the PG profile of the *E. coli* cells under expression of PBP-A.

3.2 Results

3.2.1 Fitness cost of PBP-A expression in *E. coli*

In this section, we evaluate the effect of PBP-A expression on fitness of *E. coli*. PBP-As (PBP-A wt, S61A, L158E) expression was carried out from pBAD (pBR322 ori) and/or pBAD43 (pSC101 ori) plasmids with medium- or low-copy numbers, respectively. PBP-As, in both of the plasmids, are expressed under *araBAD* promoter, which is inducible by L-arabinose. To avoid mis-localization of PBP-A, we decided to use a co-translational export to periplasm via signal recognition particle (SRP) pathway by fusing *DsbA* signal sequence (*DsbAss*) (Schierle et al., 2003) on N-terminus of PBP-A. A histidine tag (6xHis) is fused in C-terminal for western blot detection and purification purposes. To estimate fitness cost, growth rate of each sample, after induction by L-arabinose (0.5%, w/v), was calculated from the maximum slope of growth curve and the fitness cost of PBP-A expression was estimated by dividing growth rate under induced expression to growth rate without inducer (basal expression).

Growth arrest upon induction of PBP-A in E. coli

The initial fitness cost is observed as a growth defect when expression of PBP-A was induced. In Figure 3-4 we present the growth curves of the *E. coli* cells harboring PBP-A in low-copy pBAD43 plasmid as well as the growth curves of the wild-type *E. coli* Top10, under basal and induced conditions.

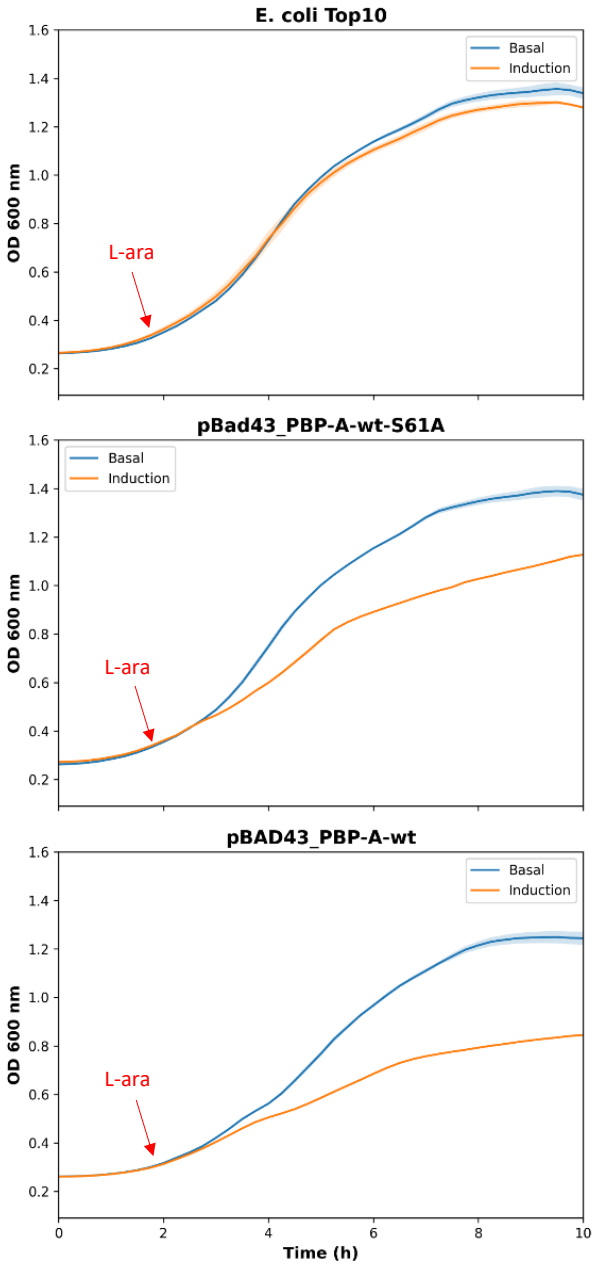


Figure 3-4: Fitness cost upon induction of PBP-A expression in *E. coli*. Induction (0.5% L-arabinose) of PBP-A-wt expression from low-copy pBAD43 plasmid (pSC101 ori) is associated with a significant growth defect, leading to a decreased growth rate as well as decrease in the max OD_{600} within 10 hours of growth.

When expression of PBP-As from low-copy plasmid was induced, a significant defect in growth was observed (Figure 3-4). However, expression of catalytic mutant of PBP-A, S61A, from low-copy plasmid also shows a fitness effect, although significantly less than the wild-type enzyme, which suggests that expression of the inactive enzyme is also costly for *E. coli*. The fitness cost can be quantitatively estimated from the growth curves by measuring either growth rate (after induction) or maximum OD_{600} (carrying capacity) from the growth curves.

Copy number matters; toxicity of pbp-a is related with gene copy number

In attempts to decrease the fitness cost of PBP-A in *E. coli,* we expressed *pbp-a* gene under medium, low and single copy conditions in *E. coli*. Medium- and low-copy expression was carried out in pBAD and pBAD43 plasmids, respectively, as explained in previous section. For single copy expression the *DsbA_pbp-a_His* construct is inserted in *rhaA* locus of the *E. coli* genome. Figure 3-5-A represents the effect of gene copy-number of PBP-A's genetic construct on growth rate of the host *E. coli* cells.

Since the fitness cost of PBP-A is correlated with copy number, it is more likely to be due to the expression level. To confirm that, we decided to evaluate the expression level of periplasmic PBP-A proteins under single- and low-copy number. Thanks to the C-terminal His-tag, we compared the level of periplasmic PBP-As under different copy-number context by western blot analysis of total periplasmic extracts. For that we applied same amount of cells from each culture by normalizing the OD_{600} to a similar value prior to the periplasmic extraction. In Figure 3-5-B we present the results of western blot analysis.

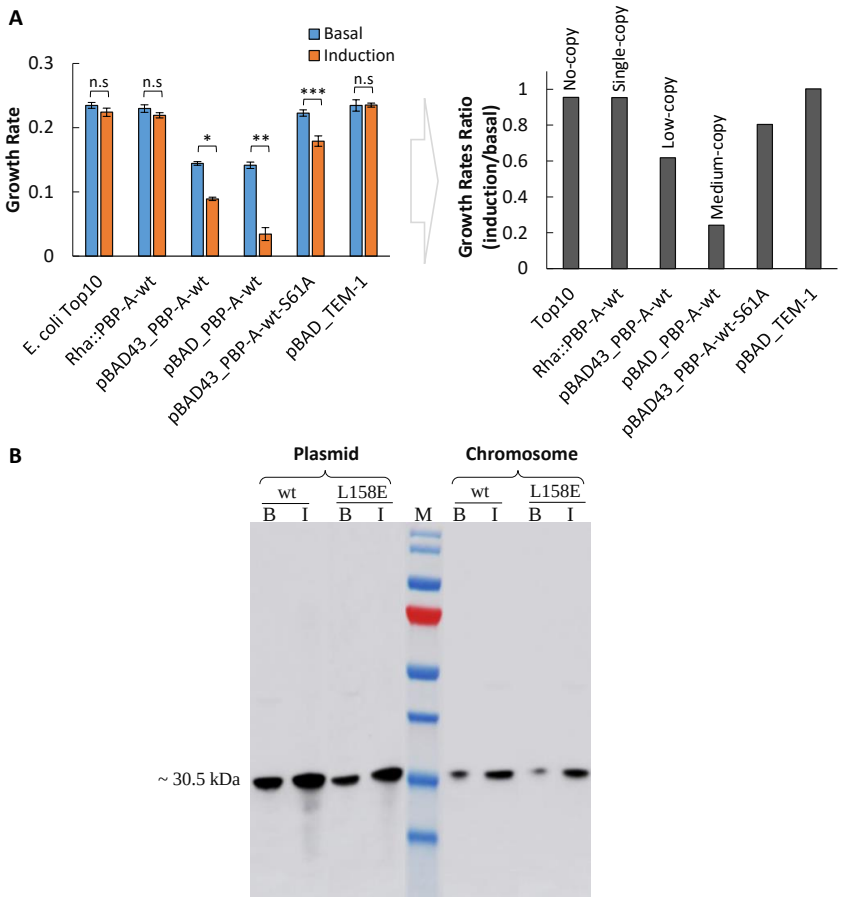


Figure 3-5: Effect of copy number of genetic construct on fitness cost. **A)** Growth rates of different *E. coli* strains. For induction of low- and medium-copy expression from pBAD43 and pBAD plasmids, L-arabinose (0.5 %, w/v)) is used and for induction of single-copy expression from *rhaA* locus, L-rhamnose (0.5%, w/v) was used. The bar graph in the left part represents the mean of three independent replicates and error bars signify the standard deviation (SD) of the replicates. The differences between basal and induction condition of cells expressing PBP-A from plasmids were significant (*P < 0.001, n = 3, t test), while not significant for cells expressing PBP-A from chromosome, nor for the controls (*E. coli* Top10 and pDad_TEM-1). In the right part, the ratio of growth rates in induction over basal condition is presented to compare the growth defect of different clones. *P < 1.4e-05, **P < 7.8e-05, ***P < 0.0013, n.s: not significant, n = 3, t test. Similar growth defect were observed for expression of PBP-A-L158E (data not shown). **B)** Effect of copy number on expression level of PBP-A. Part B represent the western blot analysis of PBP-A expression in periplasm of *E. coli* under low-copy plasmid (pBAD43) and chromosomal context, where we subjected equal amount of cells to periplasmic extraction. wt, PBP-A-wt; L158E, PBP-A-L158E; B, basal (non-induced) expression; I, induced expression; M, pre-stained protein marker. wt, PBP-A-wt; L158E, PBP-A-L158E; B, basal (non-induced) expression; I, induced expression; M, pre-stained protein marker.

Upon induction of PBP-A expression in both of the medium- and low-copy plasmid versions a significant fitness effect is observed as growth rate reduction,

while a minimum growth defect is observed in chromosomal expression (Figure 3-5-A). Similar results were obtained for expression of PBP-A-L158E in *E. coli* (data not presented). To confirm that the growth defect is associated with PBP-A expression, it can be compared to expression of TEM-1 β-lactamase, which even under medium-copy expression from pBAD plasmid does not show any change in growth rates neither under basal nor induced expression.

However, lower-copy expression decreased the magnitude of growth arrest upon induction of PBP-A expression, although not fully recovered compare to wild type *E. coli* cells or cells expressing TEM-1 β-lactamase from a pBAD plasmid (Figure 3-5-A). We also confirmed that expression of TEM-1 under any of these copy number has no significant growth effect upon induction (data not shown).

Expression level of PBP-A is correlated with copy number

Since the fitness cost of PBP-A is correlated with copy number, it is more likely to be due to the expression level. To confirm that, we decided to evaluate the expression level of periplasmic PBP-A proteins under single- and low-copy number. Thanks to the C-terminal His-tag, we compared the level of periplasmic PBP-As under different copy-number context by western blot analysis of total periplasmic extracts. For that we applied same amount of cells from each culture by normalizing the OD_{600} to a similar value prior to the periplasmic extraction. As expected, results in Figure 3-5-B show decrease in periplasmic PBP-A protein level when expressed from chromosome compare to low-copy plasmid. However, even under induced condition in single-copy context in chromosome the expression level of PBP-A protein is lower than the non-induced (basal) expression in low-copy plasmid. Nonetheless, the detectable expression of PBP-A in basal conditions confirms leaky expression of the araBAD and rhaBAD promoters in LB media in absence of the inducers.

This data also confirm an inducible (by L-rhamnose sugar) chromosomal expression system of recombinant proteins in *rhaA* locus of the rhamnose operon, which can be used for expression of toxic proteins such as PBP-A without compromising the fitness of bacteria.

Additionally, western blot analysis of total protein in Figure 3-5-B confirms a single band of mature protein (~ 30.50 kDa) for both plasmid and chromosomal expression, confirming homogenous mature (signal-cleaved) protein exported via the *DsbA* signal to the periplasm of *E. coli*.

3.2.2 PBP-A is correctly localized in periplasm of *E. coli*

The *DsbA* signal sequence exports PBP-As to periplasmic space of *E. coli* through SRP pathway and co-translational translocation (Figure 3-6-A). Such direct export to periplasm obviates any folding or activity in cytoplasm. However, to confirm the export to periplasm and to evaluate any potential folding issues that might contribute to fitness cost in *E. coli* cells, we checked the solubility of PBP-A proteins and its correct localization in periplasm. For that western blot analysis was carried out and the results are presented in Figure 3-6.

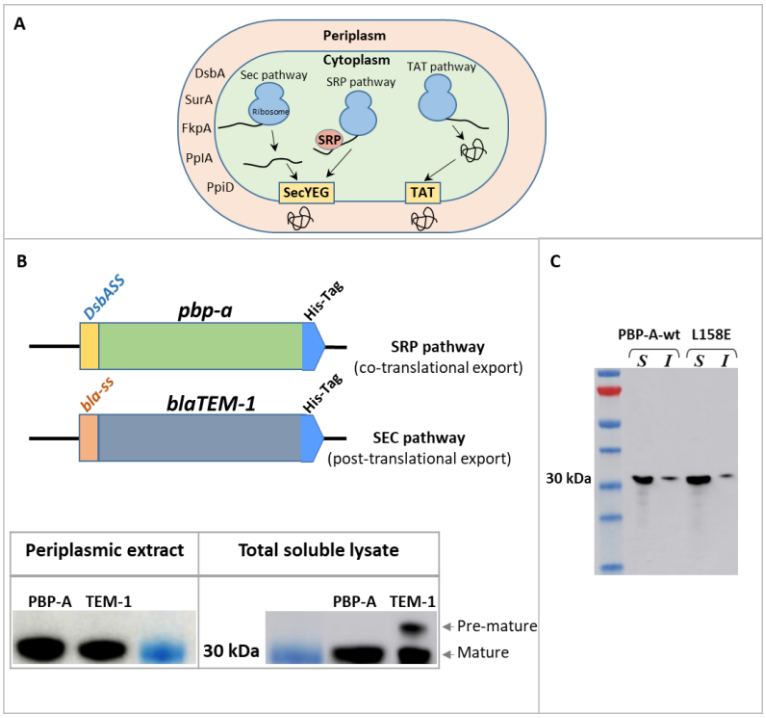


Figure 3-6: Translocation of proteins into periplasm of *E. coli*. **A)** Three main pathways involved in translocation of proteins to periplasmic space in *E. coli*. Sec and SRP pathways handover unfolded proteins and proteins in complex with SRP-ribosome, respectively, to Sec-translocon while Tat pathway transports folded proteins via Tat translocon across the plasma membrane. Both Tat and Sec pathways mediate post-translational export

while SRP-dependent pathway mediates a co-translational export of proteins across inner plasma membrane into periplasm of *E. coli* (adapted from (Liu & Huang, 2018). **B)** In upper part the genetic constructs of *pbp-A* and bla genes are shown in schematic, where a DsbA signal sequence (DsbAss) is fused to N-termini of *pbp-A* gene to facilitate a co-translational export while the TEM-1 gene is fused to its native β-lactamase signal sequence (bla-ss). In the part below, western blot of soluble fractions from periplasmic and total cell lysate of *E. coli* confirms the co-translational expression of mature PBP-A proteins without any pre-mature intermediates, while for a post-translationally mediated TEM-1 protein shows presence of both mature (30.5 kDa) and pre-mature or un-exported (33.0 kDa) proteins. **C)** Western blotting of total soluble and insoluble fractions from total cell lysate of *E. coli*/pBAD43_DsbAss-PBP-As confirms that PBP-A almost entirely exist in soluble form. The small amount of insoluble protein is thought to be due to incomplete cell lysis as we observed (data not shown) that it varies upon efficiency of lysis protocol.

Comparison of western blotting on total soluble proteins of *E. coli* cells expressing PBP-As and TEM-1, which are exported co- and post-translationally (Figure 3-6-A), respectively, confirms expression of only mature PBP-As (exported to periplasm and the signal is cleaved) with a molecular weight of 30.50 kDa, while for TEM-1 protein, both pre-mature (33.00 kDa) and mature (30.42 kDa) proteins are detectable (Figure 3-6-B). Although there is a very minor amount of protein in insoluble fraction, which could be due to inefficient cell lysis, the western blot results indicate that PBP-A-wt and PBP-A-L158E are mainly detected in soluble fractions of the total cell lysates (Figure 3-6-C). Compare to the overexpression of TEM-1 β-lactamase, which uses the post-translational Sec pathway and thus has some fraction of cytoplasmic proteins (pre-mature) that contain signal sequence (Figure 3-6-B), PBP-A proteins are all in mature form (signal cleaved) (Figure 3-6-C).

To evaluate any possible folding problem related to fitness cost of PBP-A, the fitness effect (growth defect) was evaluated during and after heat shock (42 °C) and benzyl alcohol shock (1%, v/v). If the fitness cost is associated with the folding of PBP-A, the induction of chaperonin system could improve the fitness of cells. These indirect results suggest that the induction of chaperonin systems of *E. coli* did not show any alleviation in fitness cost of PBP-A expression from low-copy pBAD43 plasmid (data not shown).

3.2.3 Phenotypic characterization of *E. coli* cells expression PBP-A

To further characterize the effect of PBP-A expression on the host *E. coli* cells, series of experiments were conducted. Since PBP-A is expressed in periplasm and as *in vitro* results in chapter 2 has shown that PBP-A is reducing the cross-linkage of PG in *E. coli*, we hypothesized that the *in vivo* fitness cost is linked to PBP-A activity on PG. To test this hypothesis we carried out series of experiments that are mainly evaluating the effects of PBP-A on the properties of bacterial cell envelope (e.g., sensitivity towards changes in the external environment and/or changes in cell wall properties).

E. coli cells expressing PBP-As are osmosensitive

Hyperosmotic and hypoosmotic stresses cause swelling and stretch in peptidoglycan sacculi leading to plasmolysis and cytolysis (or deplasmolysis), respectively. Rapid increases in water activity and dilution of solutes in external environment is known as a hypoosmotic shock for bacterial cultures, which causes sudden influx of water into the cells and increase in turgor pressure (Ian R Booth & Louis, 1999; Schleyer, Schmid, & Bakker, 1993; Sleator & Hill, 2002). Turgot pressure caused by hypoosmotic stress increases the force within the cell and pushes the plasma membrane against the bacterial cell wall. This phenomenon can be used to estimate the strength of cell wall and its components such as PG or OM.

Osmotic shock causes change in turgor pressure and, therefore, alteration in mechanical forces within the bacterial cell envelope, which can affect growth rate and fitness of the stressed cells (Rojas & Huang, 2018). We aimed to evaluate such potential growth defects that may caused by PBP-A's activity on PG of *E. coli* strains by comparison of their growth parameters under osmotic stress. For that we compared Max OD_{600} of the PBP-A expressing cells with the untransformed *E. coli* cells and the results are presented in Figure 3-7. Hyperosmotic stress was carried out under 0.75 M NaCl in the culture medium and for hypoosmotic stress, solutes in media were diluted (2-fold) by adding sterile distilled water to growth media.

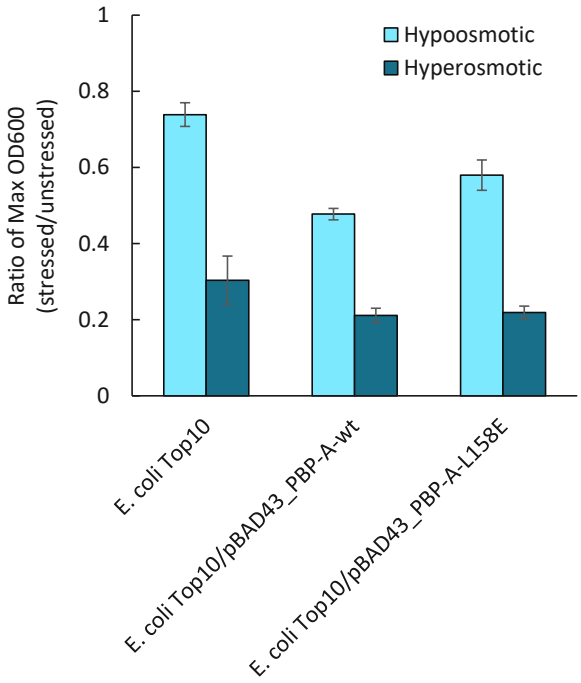


Figure 3-7: Osmosensitivity of *E.coli* cells expressing PBP-A. Maximum OD at 600 nm (Max OD_{600}), achieved during growth of *E. coli*, shows higher sensitivity of cells to osmotic stresses when expressing PBP-A proteins. Data are representing ratio of Max OD_{600} of stressed cells over the unstressed cells of the same strain. All the cultures are under induction of L-arabinose (0.5% w/v). LB media containing 0.75 M NaCl was used to apply the hyperosmotic stress and dilution (1:1) of the LB media by sterile distilled water was used to provide a hypoosmotic stress. Data were obtained from three independent replicates. The difference between *E. coli* Top10 and PBP-A expressing cells were statistically significant for hyoosmotic stress ($P < 0.005$, n = 3, one-way ANOVA), while less significant for hyperosmotic stress ($P < 0.09$, n = 3, one-way ANOVA).

E. coli show more sensitivity to both of the hyper- and hypoosmotic stresses under induction of PBP-A expression, compare to un-transformed Top10 cells. Although sensitivity to hyperosmotic environment is almost similar for both cells expressing PBP-A-wt or PBP-A-L158E compare to wild type *E. coli* Top10 cells, cells expressing PBP-A-L158E appears slightly less sensitive to hypoosmotic shock (Figure 3-7).

E. coli cells expressing PBP-As are more susceptible to vancomycin

The narrow-spectrum glycopeptide antibiotic, vancomycin, is a large hydrophilic molecule (1450 kDa) that is OM-impermeant for *E. coli* cell membrane (Nikaido, 2003b). Compromises in cell envelope of *E. coli* have been shown to affect OM's permeability and therefore, increase the cell's susceptibility against vancomycin (Mathelié-Guinlet, Asmar, Collet, & Dufrêne, 2020; Muheim et al., 2017; Stokes et al., 2016; Vaara, 1993). Here, we investigate the effect of PBP-A expression on susceptibility of *E. coli* cells against vancomycin. For that, we normalized arabinose-induced *E. coli* cultures to the same OD_{600} (OD ~0.6-0.7) by diluting higher density cultures with LB media, and then subjected to spotting assay on inductive media by 10-fold serial dilutions. Results of the spotting assay are presented in Figure 3-8.

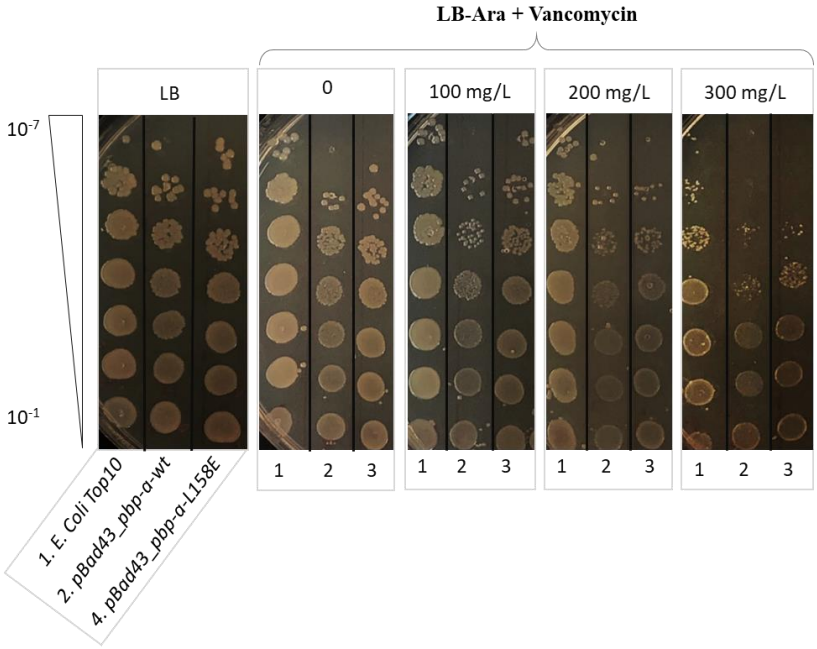


Figure 3-8: Vancomycin susceptibility spotting assay. *E. coli* cells under expression of PBP-As show a significant susceptibility to vancomycin compare to wild-type *E. coli* cells.

Spotting assay for vancomycin susceptibility showed that expression of PBP-As in *E. coli* is associated with increase in susceptibility against vancomycin,

compared to the untransformed *E. coli* Top10 cells (Figure 3-8). These results are indicative towards a damaged cell envelope, caused by PBP-A expression.

Slight increase in sensitivity to ampicillin of ampicillin-resistant E. coli cells upon expression of PBP-As

We further investigated effect of PBP-A expression on sensitivity of *E. coli* cells against ampicillin. This is particularly important for us since many directed evolution attempts of PBP-A to beta-lactamase, which involved selection against ampicillin, have failed. Since the MIC of *E. coli* cells expressing PBP-A against ampicillin was similar to the wild-type *E. coli* Top10 cells (~2 µg/ml), we decided to evaluate the MIC of ampicillin in presence of TEM-1 beta-lactamase expressing from a wild-type conjugative plasmid, RP4, to mimic a natural condition of ampicillin resistance. For that, we conjugated the RP4 plasmid into *E. coli/pBAD43_PBP-A* cells and monitored the growth parameters in presence of ampicillin (0, 512, 1024, 2048, 3000 and 4096 µg/ml) under basal and induced expression of PBP-A. The Max ODs of the cultures were obtained from the growth curves and results are presented in Figure 3-9.

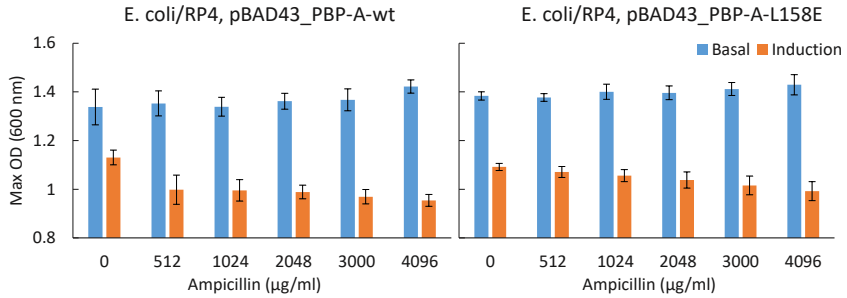


Figure 3-9: Comparison of sensitivity to ampicillin in *E. coli*/RP4, pBAD43_PBP-As under expression of PBP-As. Data represent Max OD_{600} values for ampicillin-resistant *E. coli* cultures under basal and induced (0.5% L-arabinose) expression of PBP-A-wt and PBP-A-L158E. The maximum OD of ampicillin resistant *E. coli* cells is showing a slight decrease at higher ampicillin concentrations, when expressing PBP-A proteins, while under basal condition there is no decrease but a slight increase in max OD. Data were obtained from three biologically independent experiments. The differences between the Max OD of samples under no ampicillin (0 µg/ml) and samples under ampicillin stress, under induction condition, are statistically significant ($P < 0.001$ for PBP-A-wt and $P < 0.02$ for PBP-A-L158E, n = 3, one-way ANOVA).

The sensitivity to ampicillin in ampicillin-resistant *E. coli* cells is increased upon induction of PBP-As expression, while no increase in the sensitivity under basal (non-induced) condition is observed. This susceptibility against ampicillin is detectable by comparing the maximum density of population (max OD_{600}) (Figure 3-9). However, as controls, *E. coli* cells without PBP-A and/or cells expressing PBP-A-S61A mutant are remained to be involved in this experiment.

Fitness cost is higher in acidic growth conditions

Since we discovered that PBP-A activity on non-amidated synthetic peptide as well as non-amidated PG of *E. coli* is increased at acidic pH, we hypothesize that at acidic pH, the PG of *E. coli* would become more vulnerable, hence resulting in a higher fitness cost. To test that, we evaluated the pH effect on the fitness cost by growing the cells in buffered LB media at a pH range from 5.8 to 8. To avoid unexpected effect(s) of changing buffers, for all conditions we used Potassium Phosphate buffer, which buffers well from pH 5.8 to pH 8. The culture of cells was grown until OD ~0.2, induced by arabinose, devided into four, and immediately buffered by adding concentrated buffers.

Additionally, since the PG of *E. coli* is negatively charged and PBP-A is positively charged (pI = 7.56), we also evaluated the impact of lowering the isoelectric point of PBP-A by substituting two surface lysines (Lys, K) into glutamic acid (Glu, E) to construct PBP-A-wt-2K-E (pI 5.99)

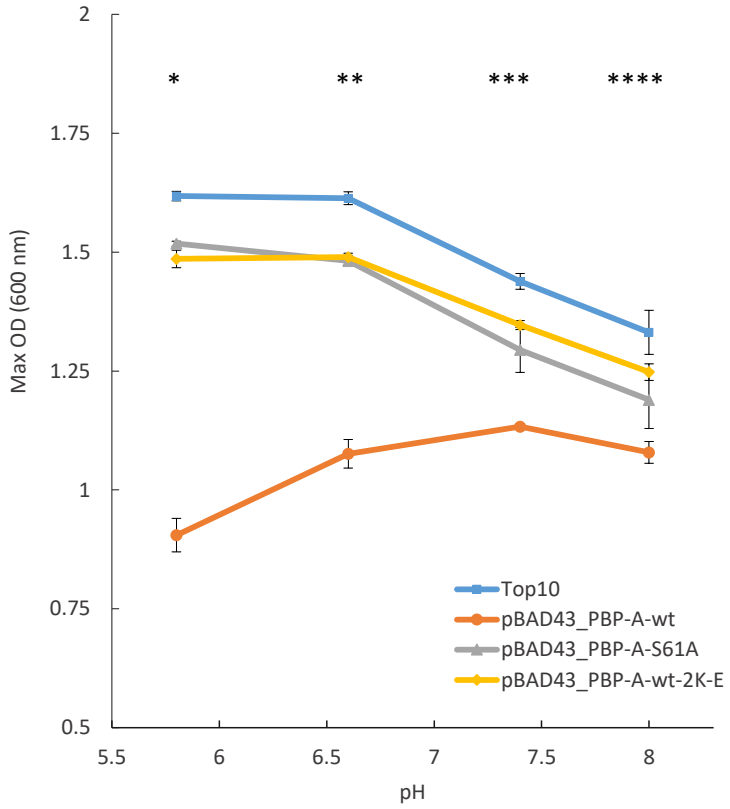


Figure 3-10: Maximum population density (max OD_{600}) of *E. coli* cultures at different pH. Comparison of pH profiles between *E. coli* Top10 wild-type cells and cells expression different PBP-A mutants: PBP-A-wt and PBP-A-wt-S61A (pI 7.56), and PBP-A-wt harboring two substitutions of surface lysines (Lys, K) into glutamic acid (Glu, E), named PBP-A-wt-2K-E (pI 5.99). The maximum fitness cost due to PBP-A expression is detectable at acidic pH and the shift in optimum pH towards basic pH upon expression of PBP-A-wt is significantly detectable. For PBP-A-wt-S61A and PBP-A-wt-2K-E, the pH profile is similar to the wild-type *E. coli* cells. The data represent average of three replicates. $*P < 5.7e\text{-}08$, $**P < 2.3e\text{-}07$, $***P < 3.1e\text{-}04$, $****P < 0.02$, $n = 3$, one-way ANOVA. Similar results were observed for *E. coli*/pBAD43_PBP-A-L158E (data not shown).

As shown in Figure 3-10, optimum pH range for *E. coli* Top10, to achieve maximum density/capacity of population (OD_{max}), is at acidic range (pH < 7), while for cells expressing PBP-As the optimum pH shifted towards basic pH, (pH > 7). Therefore, fitness cost associated with PBP-As expression is more pronounced at acidic pH in agreement with *in vitro* activity data that PBP-A was more active on PG of *E. coli* at acidic pH.

Interestingly, mutating two surface-exposed lysines into glutamates (PBP-A-2K-E, pI=5.99) resulted in a fitness recovery at acidic pH (Figure 3-10), in a similar extent to the inactive enzyme. This suggests that either (i) the 2K-E mutations is compromising the activity of PBP-A, which requires further activity characterization of this mutant, or (ii) a negatively charged PBP-A has less fitness cost, due to less interaction with the negatively charged PG. However, the hypothesis that negatively charged PBP-A has less fitness cost in periplasm of *E. coli* is more plausible when we realized that class A beta-lactamases such as TEM-1 (pI=5.1) has an overall negative charge in neutral and acidic pH. This will be further explained in the discussion section.

3.2.4 Alterations in PG profile of *E. coli* cells upon expression of PBP-As

Following *in vitro* observations in chapter 2, that PBP-A is impairing crosslinking of *E. coli* PG, as well as the phenotypes indicative towards an imperfect cell envelope, we hypothesized that fitness cost related to expression of PBP-A in *E. coli* is caused by its interfering activity(ies) on the bacterial PG. Therefore, we aimed for analyzing the chemical composition of the PG from *E. coli* cells under single- and low-copy expression of PBP-A and PBP-A-L158E. Untransformed *E. coli* Top10 strain was used as reference for PG profile and negative control. To do so, we grow the cells to late-exponential phase (OD ~0.6-0.7) under expression of PBP-A followed by preparation of the cells lysate for PG isolation by boiling in SDS (Glauner, 1988). Analysis of muropeptides were carried out by HPLC-MS and the abundance of muropeptides in the PG of each strain was quantified. In Figure 3-11 we present the abundance of muropeptides that showed most alteration due to PBP-A expression, relative to the abundance of the same muropeptides in wild-type *E. coli* Top10 strain. Abundance of all muropeptides detected in these assays are presented in Appendix 9-21.

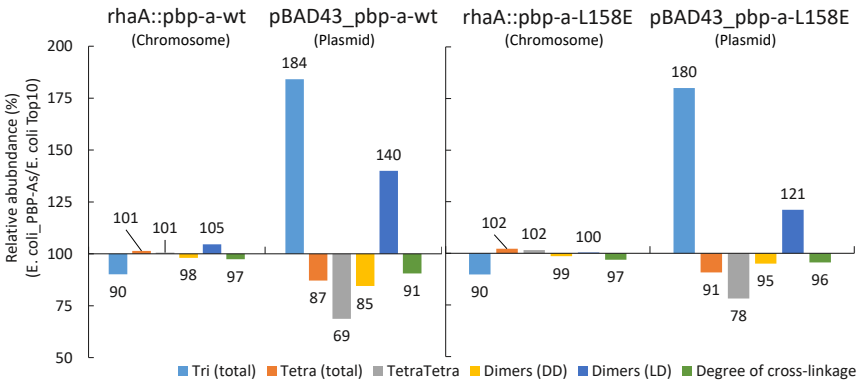


Figure 3-11: Expression of PBP-As is associated with alteration in PG of *E. coli*cells. Bars represent percentage of muropeptides detected in *E. coli* cells expressing PBP-A under single- and low-copy, from *rhaA* locus in genome and pBAD43 plasmid, respectively, relative to untransformed *E. coli* cells. While under single-copy, there is minimum change in PG composition, at low-copy expression from plasmid the changes in PG profile are significant. The significant alterations include: increase in Tri monomers and LD dimers and decrease in Tetra monomers, TetraTetra dimers, DD dimers and an overall decrease in degree of cross-linkage. *E. coli* cells were induced for PBP-As expression at early exponential phase by addition of L-rhamnose (0.5%, w/v) or L-arabinose (0.5%, w/v), for single- or low-copy expressions, respectively, and harvested at late exponential phase for PG isolation. The abundance (%) of all peptides detected in PG isolates of all the strains in this experiment are available in Appendix 9-21.

Comparison of PG profiles of different *E. coli* cells shows that alteration in abundance of muropeptides as well as degree of cross-linkage are associated with expression level of PBP-A, which is in agreement with our results of expression-dependent fitness costs (Figure 3-5). Analysis on PG of *E. coli Top10/pBAD43_PBP-As* cells, under induction of PBP-As expression, shows significant modifications in the peptidoglycan (Figure 3-11), confirming that PBP-A is causing alterations on the PG profile of the host cells.

Overall, PBP-A expression results in a reduction of PG crosslinking, which is mainly detectable as decrease in crosslinks between tetrapeptides (4-3 bridges or DD dimers), confirming the EPase activity of PBP-A that was detected in *in vitro* assays (chapter 3). However, crosslinks between tetra and tri peptides (3-3 bridges or LD dimers) increased, indicating an adaptation of the PG. Similar to *in vitro* assays in chapter 2 (section 2.7.5) an unexpected LD-CPase activity, linked to PBP-A expression, is detectable as decrease in Tetrapeptide monomers, while

increase in Tripeptides, compare to PG profile of wild-type *E. coli* Top10. There are also some other modifications detected including increase in Gly4 containing muropeptides, which are probably due to adaptation and remodeling of PG during the stress caused by PBP-A.

3.3 Discussion

In this chapter, we presented data related to the nature of the fitness cost (toxicity) imposed to *E. coli* cells when expressing the PBP-A proteins. After confirming the involvement of PBP-A activity in fitness cost, we conducted phenotypic characterizations of the host *E. coli* cells which showed evidences (such as increased susceptibility against vancomycin, and/or increased osmosensitivity) that PBP-A is affecting the cell wall of the bacteria. Then, we performed analysis of peptidoglycan of the cells as a potential target of PBP-As' activity/interaction and discovered that indeed PBP-A is causing alteration in PG profile.

Fitness cost of PBP-A is activity-dependent and correlated with expression level

Fitness costs due to overexpression of recombinant proteins are common in molecular biology, which mostly are considered to be related to misfolding and insolubility of the proteins. Thus to confirm that the main cause of the fitness cost of PBP-A is not due to biophysical restrictions on the protein, we evaluated its solubility and folding properties. The analysis of soluble and insoluble fractions of cell lysates from *E. coli/pBAD43_PBP-As* by western blotting, did not show any significant aggregation or insolubility related to PBP-A expression (Figure 3-6). Therefore, under conditions where the protein is correctly exported and mainly soluble in periplasm, the activity of PBP-A is responsible for the main fitness cost. However, an activity-independent contribution of PBP-A to fitness cost was also detected upon expression of inactive mutant of PBP-A (S61A) from low-copy plasmid that was a significantly lower than fitness cost of the active enzyme. This cost is probably due to the use of SRP-dependent signal that may cause saturation in the co-translational translocation capacity of the Sec-translocon (Schlegel et al., 2013). Although we did not observe significant

inclusion bodies in cell lysates, we further (though indirectly) examined that fitness cost of PBP-A is not associated to its folding (e,g., due to requirement of additional chaperonin assistance), by evaluating the fitness cost during and after heat shock and benzyl alcohol shock, which are known as a general strategies to induce chaperonin systems for improving protein folding in *E. coli* (de Marco et al., 2007).

Overall, these data suggest that fitness cost of PBP-A is mainly dependent on its activity. Based on these observations and following the results of *in vitro* activity of PBP-A on peptidoglycan, we hypothesized that the fitness cost is dependent on activity of PBP-A against PG of *E. coli*. This was confirmed by analyzing the PG composition of the fitness-impaired *E. coli* cells and observing damaged PG with a reduced crosslinking. In agreement with *in vitro* assays of PBP-A on PG of *E. coli* (chapter 2), this data suggest that the origin of the fitness is mainly due to DD-EPase activity of PBP-A by directly affecting crosslinking of the PG. However, DD-CPase activity of PBP-A may also indirectly affect the biosynthesis/remodeling of PG by reducing pentapeptide (Penta) substrates for endogeneous DD-TPases of the *E. coli*.

Such expression-related fitness cost was previously reported as even overexpression of *E. coli's* own DD-peptidases contributes to growth defect upon induction of their expression from a medium-copy plasmid, that was suggested to be due to cell lysis (D. E. Nelson & Young, 2001). Of course, *E. coli* carries its own endogeneous murein hydrolases such as DD-CPase, LD-CPase and DD-EPase enzymes involved in biosynthesis, maintenance and recycling of peptidoglycan. However, depending on phase of bacterial growth, these enzymes are regulated in multiple levels to play their vital role in cell growth and division (Egan et al., 2020; Garde, Chodisetti, & Reddy, 2021; Typas et al., 2011, 2010; Verheul et al., 2020; Vermassen et al., 2019), and overexpression of murein hydrolase enzymes would disturb this spatio-temporal regulation of PG biosynthesis (Yang et al., 2019) and affect the cellular fitness (D. E. Nelson & Young, 2001). Among eight endopeptidase paralogues in *E. coli*, PBP4, PBP7and AmpH are serine hydrolase enzymes, which confer DD-CPase and DD-EPase

activities and involved in maturation, remodeling and recycling of PG (Sauvage et al., 2008b; Voedts et al., 2021). Although these DD-EPase enzymes possess cleavable signal sequence and are considered periplasmic proteins, it was shown that they remain membrane associated (González-Leiza, de Pedro, & Ayala, 2011; Henderson, Templin, & Young, 1995; van Heijenoort, 2011). On the other hand, PBP4 and AmpH express at a very low level in wild-type *E. coli* cells (Henderson, Young, Denome, & Elf, 1997) and overexpression of PBP4 imposes a fitness cost in *E. coli* (D. E. Nelson & Young, 2001). Our observations are consistent with these data, suggesting that expression of peptidoglycan hydrolase enzymes is tightly regulated while their overexpression causes a fitness cost, which is shown in this study to cause disturbance on PG profile mainly by reduction in PG crosslinking. Nevertheless, the membrane anchorage of *E. coli* DD-EPase enzyme (and many hydrolases) might play a regulatory role and secure PG sacculus against their activity, unlike PBP-A in this study, which is expressed as a soluble (not membrane associated) periplasmic protein, therefore, accessing PG cross-links without constraints of the membrane anchoring.

Increase in Gly4 peaks and LD-cross-links such as TetraTri(DAP) observed in PG analysis, could be due to stress induction of LdtD transpeptidase (LD-TPase) under stress caused by PBP-A expression. LdtD is known to be a stress response LD-TPases able to exchange D-Ala by Gly at position 4 (Morè et al., 2019).

Additionally, characterization of the genetic copy number revealed that the level of expression and, thus the level of toxicity due to expression, is correlated with the copy number of the genetic construct of *pbp-a* in *E. coli*. Even under induced expression of PBP-A from single-copy chromosomal context, the fitness cost is significantly lower than the basal (non-induced) expression in low-copy plasmid context, a promising approach to control toxicity due to protein expression in *E. coli*.

Phenotypes related to an imperfect cell wall

PBP-A-expressing *E. coli* cells showed increased sensitivity against osmotic stress and vancomycin, which points toward a destabilized outer membrane (OM). The

OM layer has a protective role against the external environment by acting as a selective barrier, and as a components of bacterial cell wall, it is involved in bearing turgor pressure and preventing cell lysis. The changes in turgor pressure caused by osmotic shock are effective in mechanical forces within the bacterial cell envelope (Rojas & Huang, 2018), which are controlled by the PG and the OM as well as the proteins involved in PG-OM interactions (Mathelié-Guinlet et al., 2020). The effect of osmotic shock on these forces can be more critical if the integrity of cell wall is compromised. As PBP-A has impairing effect(s) on the PG of the host *E. coli* cells, it can (i) directly impact the integrity of cell wall, (ii) and/or indirectly effect the adaptation against the osmotic stress.

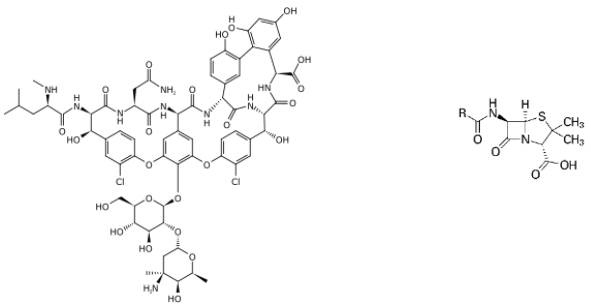
Susceptibility against the OM-impermeant antibiotic, vancomycin, can be increased by compromises in the integrity of OM (Mathelié-Guinlet et al., 2020; Muheim et al., 2017; Stokes et al., 2016; Vaara, 1993). A recent study by Mathelié-Guinlet et al (2020) showed that vancomycin is able to traverse an OM bilayer whose integrity has been mechanically disrupted through modifications on Lpp (Braun's lipoprotein), which covalently crosslinks the peptidoglycan sacculus to the outer membrane bilayer. Another study shown that disturbing the covalent interaction between PG and OM by disruption of the lipoprotein traffic to the outer membrane increases the permeability of *E. coli* against large-scaffold antibiotics such as vancomycin (Muheim et al., 2017). Lpp-peptidoglycan linkage is mediated through a tripeptide in the PG and catalyzed by three LD-TPase in *E. coli*, LdtA, LdtB and LdtC, which use Tetra muropeptide as donor substrate and Lpp C-terminal lysine as the acceptor (Magnet et al., 2007). Although we did not observe decrease in total Tri-Lys-Arg, expression of PBP-A in *E. coli* is associated with decrease in abundance of Tetra (Figure 3-11), which are essential substrates in Lpp-PG linkage. However, the effect of PBP-A on PG-OM linkage could also be indirect due to the damages in the PG.

Furthermore, we performed single-cell atomic force microscopy (AFM) experiments to evaluate effect of PBP-A expression on mechanics of *E. coli* cell envelope. Although the data were indicative towards alteration in mechanical properties of the cell envelope, results were highly variable (not shown),

suggesting that the bacteria may be adapting to the PBP-A stress and slight (impalpable) changes in culture conditions seem to generate large variations.

Increased sensitivity of ampicillin-resistant *E. coli/RP4* cells, under expression of PBP-A, against ampicillin could be explained by two hypotheses. First, the higher sensitivity against ampicillin may be due increase in the influx of ampicillin caused by expression of PBP-A (due to damage in the cell envelope), and thus effecting the MIC of ampicillin resistance provided by RP4/TEM-1. Second, this effect may be due to inhibition of endogeneous PBPs that are important in the PG adaptation/metabolism, hence their inhibition will be more deleterious in the context of a less reticulated PG upon PBP-A expression.

Here we further discuss the first hypothesis. Relatively small-scaffold antibiotics (less than 600 Da) such as β-lactams (Figure 3-12) can generally permeate across non-specific OM porins by passive diffusion and enter to the periplasmic space (Figure 3-2). However, level of this diffusion may be altered through an imperfect OM caused by PBP-A stress.



Vancomycin (1450 Da) **Penicillin (334 Da)**

Figure 3-12: Comparison of molecular weight of large scaffold antibiotic vancomycin with small-scaffold β-lactam antibiotic penicillin.

As discussed earlier, expression of PBP-A increases the permeability of OM against large-scaffold vancomycin, most probably by effecting Lpp anchorage. Depending on the nature and degree of this permeability, is might be much higher for the smaller smaller-scaffold ampicillin. Thus, the increase of ampicillin

sensitivity in ampicillin-resistant *E. coli/RP4* could be explained by higher periplasmic concentration of ampicillin as a result of higher permeability of cell wall. Nikaido and Normark (1987) have well discussed the factors involved in MIC of β-lactams antibiotics against Gram-negative bacteria. Therefore, based on the assumption that periplasmic β -lactam concentration is at equilibrium between diffusion and hydrolysis (by periplasmic β-lactamases), they proposed a model for estimation of minimum inhibitory concentration that is formulated by Knies et al., (2017) as bellow:

$$MIC = [S]_{lethal} + \frac{[S]_{Lethal} \cdot k_{cat} \cdot [E]_{active}}{P.A.(K_M + [S]_{Lethal})}$$

where, $[S]_{lethal}$ is the periplasmic concentration of β-lactam that inhibits the PBPs and blocks the cellular growth, $[E]_{active}$ is the concentration of active β-lactamases in periplasm, *P* is the permeability coefficient of the β-lactam across cell's OM and periplasm, and *A* is the surface area of the cell (Knies, Cai, & Weinreich, 2017; Nikaido & Normark, 1987). Therefore, according to this reasonable model, MIC is negatively correlated with permeability (P) of the cell wall against antibiotics while positively correlated with the k_{cat} of β-lactamases in periplasm. This suggests that our observation of higher sensitivity of *E. coli/RP4* cells against ampicillin is presumably due to higher permeability of ampicillin, because the rest of the factors are expected to remain constant during our experiments. However, TEM-1 is a highly efficient β-lactamase and a slight increase in sensitivity of resistant cells against ampicillin would mean a much higher sensitivity for non-resistant cells. Hence, in ampicillin-sensitive *E. coli/ pBAD43_PBP-A* cells (expressing no β-lactamase), it is imaginable that the effect of increased permeability of β-lactams would become much stronger (more lethal), which could be a reason behind failure of selecting resistant clones in libraries of PBP-A in directed evolution attempts of converting PBP-A into a β-lactamase.

Effect of pH and isoelectric point on fitness cost of PBP-As

Fitness cost(s) due to PBP-A activity was higher at low pH. We propose two non exclusive reasons for that: first, lowering the pH is changing the overall charge of PBP-A to positive and the enzyme is then attracted by the negative PG.

Second, as PBP-A is shown (chapter 2) to be active (i) on amidated peptidic substrate (in D-isoGlu) and not on non-amidated peptide, and (ii) on non-amidated peptide and PG of *E. coli* at acidic pH, therefore, muropeptides with protonated D-isoGlu (at low pH) are better substrates of PBP-A compared to anionic substrate (high pH).

In contrast to cytoplasmic pH of neutralophilic bacteria (grow best around neutral pH) such as *E. coli,* which is always maintained almost neutral (pH 7.0 to 7.5) by the force of homeostasis phenomenon, the pH of periplasmic space is dependent to the pH of external environment (Slonczewski, Rosen, Alger, & Macnab, 1981; Wilks & Slonczewski, 2007). This maintenance and regulation of cytoplasmic pH is due to control mechanisms over the permeability of the inner cell membrane to protons, (I R Booth, 1985), while there is no such control for periplasmic proton balance. Therefore, periplasmic proteins' overall charge is dependent on pH of the external growth environment (Stancik et al., 2002; Wilks & Slonczewski, 2007). Similarly, the net charge of periplasmically expressed PBP-As is also dependent on pH of external media and will be affected by alterations in the pH of external environment. Interestingly, we realized that the isoelectric point of, structurally similar, TEM-1 β-lactamase and PBP-A were very different (Table 3-1). At neutral pH, the global charges of the proteins were opposite, negative for TEM-1 and positive for PBP-A. The positive charge of PBP-A at neutral or acidic conditions may result in electrostatically binding of PBP-A to the negatively charged plane of peptidoglycan or its abundant peptide crosslinks or stem peptides. To investigate whether the charge of PBP-A has any role in its toxicity against *E. coli* cells, we shifted the charge of PBP-A by (i) changing the pH of growth media and (ii) by mutating positively charged lysine residues at the surface of PBP-A, to decrease its pI. Both of basic pH (7-8) and lower pI, which changed the positively charged PBP-A macromolecules to negative, showed a beneficial effect on the fitness and partial restoration of the fitness effect. These observations suggest that vulnerability of PBP-A-expressing *E. coli* cells to acidic pH could be due to electrostatic interactions of PBP-A with PG components such a cross-links and/or stem peptides. We further confirmed by *in vitro* assays

that PBP-A has higher activity at acidic pH (pH 5) on PG of different *E. coli* strains compared to more basic pH (pH 7.5) (section 2.2.2 chapter 2).

Nonetheless, the second reason (sensitivity to pH is related to the protonation state of the non-amidated muropeptides of *E. coli*) is probably the more plausible scenario as amidation of carboxy groups in peptidoglycan of cyanobacteria is reported (Jürgens et al., 1983), which suggest that the original substrate(s) of PBP-A is probably an amidated muropeptide.

3.4 Conclusions

Our results show that overexpression of the cyanobacterial PBP-A in *E. coli,* is associated with high fitness cost(s) by altering the PG composition of the bacteria. The effect is mostly reduction in PG crosslinking, which is caused by the EPase activity of PBP-A.

Our data on protein expression analysis confirm an inducible expression system in chromosome of *E. coli*, which we developed by coupling CRISPR-Cas9 and λ-Red recombination technologies (Dorrazehi et al., 2020). We observed a correlation between expression level and the intensity of fitness cost and reported a minimum fitness cost under single-copy expression of PBP-A.

In addition, lower gene-copy number, alkaline pH and negatively charged PBP-A can minimize the fitness cost imposed by its expression to the host *E. coli* cell. Finally, we recommend that the conversion of PBP-A into a beta-lactamase in *E. coli* host should be carried out under fitness-improving conditions such as keeping the pH of growth media above 7 (avoid acidification due to metabolism), a starting PBP-A variant with low pI and/or addition of osmo-protectants to the media.

3.5 Contribution of authors:

W.V generously collaborated on HPLC-MS analysis of peptidoglycan isolates as well as discussions on results. The rest of the works including development of theory, preparation of mutants of PBP-A, assays, discussions on data and preparation of manuscript are carried out by G.M.D and supervised by P.S.

3.6 Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722610 (ES-Cat).

3.7 Material and methods

3.7.1 Strains and plasmids

Escherichia coli strain Top10 was purchased from Invitrogen, which has a genotype of F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-.

pBAD43 plasmid is a low-copy (~5 copies/cell) expression vector (Thompson et al., 2018), which is stringently replicated under pSC101origin (T. S. Lee et al., 2011). It carries tightly regulated pBAD expression system under araBAD promoter and AraC and catabolite activator protein (CAP)-cAMP regulators from araBAD (arabinose) operon (Guzman, Belin, Carson, & Beckwith, 1995). The expression of gene of interest (GOI) is tightly repressed in the presence of glucose and absence of L-arabinose and highly induced in absence of glucose and presence of L-arabinose. pBad43 plasmid was kindly provided by Michael Deghelt from Institut de Duve, UCLouvain, Brussels.

3.7.1.1 Single-copy expression

E. coli genome engineering and insertion of PBP-A in rhaA locus

We have already published the detailed protocol of genome engineering for 'Building Scarless Gene Libraries in the Chromosome of Bacteria', which is attached at the end of general methodology chapter and available online: https://doi.org/10.1007/978-1-0716-0720-6_11. Briefly, *pbp-a* gene comprising *DsbA* signal sequence and 6×His tag fused at the amino- and carboxy-terminal, respectively, was replaced with *rhaA* gene at the second position of the polycistronic operon of rhamnose in the genome of *E. coli* using CRISPR-Cas9 technology and λ-Red recombineering system (Figure 3-13).

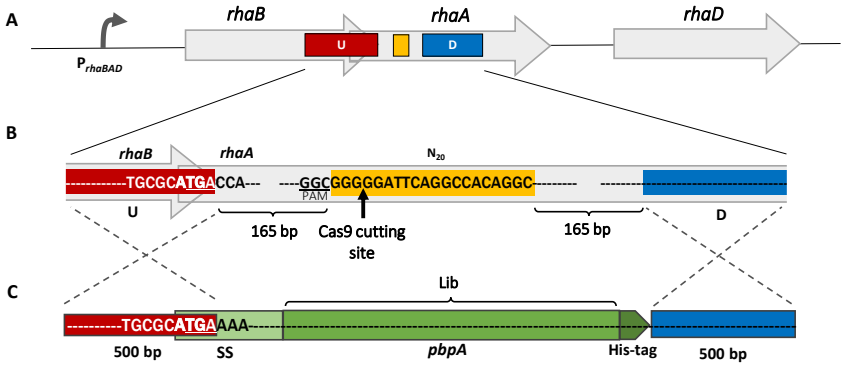


Figure 3-13: Schematic of *php-A* insertion in *rhaA* locus of *E. coli* genome. **A)** Rhamnose operon. We decided to introduce *php-A* gene exactly at the position of the second gene (*rhaA*). **B)** Nucleotide resolution of N_{20} target of guide-RNA, Cas9 cut site and up and down homology regions (U and D). **C)** Donor DNA cassette of *php-A* gene (or library). The gene of interest will replace exactly the start of *rhaA* gene far enough from promoter (see Note 1 in section 7.2.12). The insertion will be facilitated by homologous arms (500 bp) via λ-Red recombination, leaving no modification on rhaB while using the start codon of the *rhaA* gene. The *php-A* ORF encodes the N-terminal DsbA signal sequence (SS) for periplasmic localization and a C-terminal 6×His-tag for purification. (See section 7.2 for the detailed protocol of this genome engineering).

3.7.1.2 Cloning of pbp-a gene into low- and high-copy plasmids

Both of *php-a-wt* and *php-a-L158E* were amplified from *RhaA* locus from a previously engineered genome of *E. coli* Top10 (Chapter 3) and assembled into pBAD43 plasmid thanks to the overlapping sequences integrated in primers followed by Gibson assembly. The protocol of cloning *php-a* into low-copy pBAD43 plasmid is explained in section 7.1.3 Cloning into medium-copy plasmid was previously done and kindly provided by Julia Wessel. However, to make sure that revertant and adaptation against toxicity of PBP-A has not happened in clones, the analysis were confirmed by fresh transformants with minimum freezing-thowing cycles. Map of the low-copy pBAD43_PBP-A and high-copy pBAD_PBP-A is provided in section 7.1.2 of the general methodology of this thesis.

3.7.1.3 Replacement of DsbA signal sequence with bla-ss

Signal sequence of TEM-1 β-lactamase (MGIQHFRVALIPFFAAFCLPVFA) was amplified by PCR from pBAD43_TEM-1 together with the pBAD43 plasmid backbone by designing primer to exclude TEM-1 (reaction 1). In another

PCR reaction, the PBP-A gene without *DsbAss* was amplified. After DNA cleanup, PCR product of linearized pBAD43_bla-ss backbone and PBP-A gene were mixed in 1:3 ration of plasmid:gene and assembled using Gibson assembly method, thanks to overlap sequences incorporated into primers. On ice, 1 µl of the assembly mix was added to 50 µl of *E. coli* Top10 competent cells followed by transformation by electroporation.

Following primer pairs were used for the two reactions. Overlapping sequences, at 3' and 5' end of PBP-A are indicated as bold and underlined, respectively:

Reaction 1:

pBAD43_bla-ss fw: CTTGGCTGTTTTGGCGGATGAG

pBAD43_bla-ss rv:

TTGCAGATTCGTCAGCGGGCGTTCCGGCGGCAGCGTGCTCGTCG

GGGCAAAAACAGGAAGGCAAAATGC

Reaction 2:

Pbb-a_noSignal fw: **AACGCCCGCTGACGAATCTGCAA**C

Pbb-a_noSignal rv: TCTCATCCGCCAAAACAGCC

3.7.2 Periplasmic lysate preparation and western blot analysis

Cells were normalized to the same OD600, by diluting the more dense cultures, to subject similar numbers of cells to the periplasmic extraction. Protocols for periplasmic extraction and the western blot are available at section 7.1.14 and 7.1.19 respectively, in general methodology.

3.7.3 Bacterial growth rate measurements

For all the bacterial cultures in this study, pre-cultures are started by inoculation of cells from a single colony into 5 ml of media containing appropriate antibiotics and 0.4% (w/v) glucose (for repression of leaky recombinant protein expression) and incubated overnight at 37 °C under agitation of 180 rpm.

To start the growth measurement experiment, 1 ml of pre-culture was inoculated into 100 ml of LB media without any antibiotic in a sterile 500 ml erlenmeyer flask and incubated at 37 °C under agitation of 180 rpm. When cultures reached early exponential phase (OD600 ~ >0.2), cultures were divided into half by transferring each 50 ml to a sterile 250 ml erlenmeyer. One of the culture of each

strain was induced for expression of PBP-As by adding L-arabinose (0.5% w/v) followed by incubation for 4-5 hours. OD_{600} of the cultures were measured over time from the start of the cultures.

To estimate fitness effect associated with expression of PBP-As in *E. coli* Top10, we measured growth rate after induction of PBP-As expression and maximum optical density at 600 nm (max OD_{600}) from growth curves of induced and basal expressions. Growth rate of each culture was estimated by measuring the slope of the growth curve in exponential phase and max OD of the cultures were simply determined as highest value(s) of absorbance at 600 nm during the growth. Fitness cost of each strain are quantified by dividing average values of growth rate and OD_{max} of induced cultures to the basal (non-induced) cultures as shown in following formula:

$$Ratio\ of_{\text{growth rate, OD max}} = \frac{induced\ cultures_{\text{growth rate, OD max}}}{basal\ cultures_{\text{growth rate, OD max}}}$$

The growth rate, can be estimated by the slope of the tangent line drawn to the inflexion of the growth curve at exponential phase, which can be estimated by following equation in Microsoft excel:

$$b = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2}$$

where *b* is the slope of the regression line and *y* values are absorbance at 600 nm (OD_{600}) measured at time points *x*.

3.7.4 Osmotic stress assays

To change the osmolarity of external environment to the bacterial strains and, therefore, the turgor pressure within the cells, two osmotic stress were employed, hyperosmotic and hypoosmotic stress.

Hyperosmotic stress: To increase the osmolarity of bacterial external environments, LB growth medium was replaced with LB supplemented with high concentration of NaCl (0.75 M) and the growth was monitored over time. Growth rate and max OD_{600} were estimated as explained in previous section.

Hypoosmotic shock: To decrease the osmolarity in bacterial growth environment, solutes in media were diluted by addition of sterile distilled water to growth media

(1:1 dilution) at mid-exponential growth (OD_{600} ~ 0.4-0.5) followed by monitoring growth over time and calculation of growth rates and max OD_{600}.

3.7.5 Vancomycin susceptibility spotting assay

Cultures of different bacterial strains were induced for expression of PBP-As by adding L-arabinose (0.5% w/v) at OD600 ~ 0.2-0.3 and grown to late exponential phase (OD600nm ~ 0.6-0.7). Then the cultures were normalized to the same OD_{600} by diluting higher density cultures with LB media. From the OD_{600} normalized samples, cultures were diluted by series of sequential 10 fold dilutions in sterile LB media up to 10^{-7} and then 4 µL of each dilution were spotted on LB agar supplemented with 0, 100, 200 and 300 mg/L of vancomycin and L-arabinose (0.4% w/v). Plates were incubated 16-20 hours at 37 °C and analyzed by calculating cfu/ml.

3.7.6 Preparation of cell lysates

Periplasmic extraction: After induction of PBP-A expression for 3-4 hours, *E. coli* cultures (10 ml) were harvested at 4400 *xg* for 10 min, supernatants were discarded, cells were resuspended in 0.33 mL of sucrose 20% (w/v) in 20 mM tris-HCl, pH 8 and 33 µl of EDTA 0.1 M, pH 8 was added. Samples was incubated at room temperature for 10 min followed by 10 min centrifugation of cells at 6 000 rpm. Supernatants were discarded and the pellets subjected to osmotic shock by resuspending in 0.5 mL of 5 mM $MgSO_4$. The mix was incubated at 4 °C for 10 min followed by 10 min centrifugation at 10 000 xg. The supernatants were recovered as periplasmic extracts. The pellets could be continued to cytoplasmic extraction.

Complete cell lysis: Pellet of 10 ml bacterial culture was resuspended in 500 µl of lysis buffer containing 10% glycerol, 0.1% Triton X-100, 100ug/ml lysozyme, 1 mM EDTA and 3 U DNAse. The cell suspension was incubated at room temperature for 30 min and sonicated 5 times for 20 seconds until the samples are no longer viscous and a clear lysate was obtained. After centrifugation at 4 °C and 12 000 xg for 10 min, the supernatant collected as the soluble cell lysate.

Insoluble fraction: The pellet remaining from the cell lysis preparation was considered as insoluble fraction and it was suspended in 500 µl of PBS 1X pH 7 for further analysis.

3.7.7 Purification of PBP-A proteins

Thanks to a histidine tag (6x HisTag) at C-terminal of the PBP-As, purification of these protein was carried out by affinity chromatography using nickel-charged resins. Details are available in (section 7.1.20) of general methodology chapter 6.

3.7.8 Conjugation of RP4 plasmid into *E. coli/pBAD43_PBP-As* cells

Pre-cultures are started from single colonies of *E. coli* CM140/RP4 (carrying the conjugative plasmid RP4) and *E. coli* Top10/*pBAD43*, as donor (F+) and recipient (F-) strains, respectively, in LB media containing appropriate antibiotics and incubated overnight at 37 °C. Next day, fresh cultures are started by inoculation of 150 µl of pre-cultures into 15 ml LB media with appropriate antibiotics and incubated under agitation at 37 °C until OD_{600} is reached to ~ 0.5. Cultures were centrifuged at 4000 rpm and the pellets were resuspended in 1 ml of PBS buffer (1x) and centrifuged again at 4000 rpm to wash the antibiotics. The washing step is repeated and the pellets were resuspended into 0.5 ml of LB media without antibiotics, followed by 30 min incubation at 37°C to regenerate the pili. The conjugation is carried out by spotting 10 µl of recipient strain on LB agar without antibiotics and 10 µl of donor strain on top of the spot of recipient strain, followed by incubation for 4-6 hours at 37 °C. The cells are then scraped and resuspended into 0.5 ml of LB media, 50-100 µl was plated on LB agar containing antibiotics for selection of transconjugant strain (spectinomycin and ampicillin for selection of pBAD43 and RP4, respectively), and incubated overnight at 37 °C.

3.7.9 MIC of ampicillin resistance from RP4 plasmid in PBP-A expressing context

Pre-cultures are started from single independent colonies of *E.coli* Top10/[pBAD43_PBP-As, pRP4] in LB media containing kanamycin (30 µg/mL) and spectinomycin (30 µg/mL) for maintenance of RP4 plasmid and pBAD43_PBP-As, respectively, and glucose (0.4% (w/v), for repression of any

leaky expression of PBP-A) and incubated overnight at 37 °C under 180 rpm agitation. To minimize bias related to initial population, prior to starting the growth measurement experiment, OD_{600} of different cultures were normalized by dilution of more dense cultures with LB media. OD_{600}-normalized cells were diluted 1:100 into fresh LB media containing different concentration of ampicillin (0, 512, 1024, 2048, 3000 and 4096 μg/mL), spectinomycin (30 μg/mL) and L-arabinose (0.5%, w/v), in a sterile 96-well plates (Greiner Bio-One). Plates are incubated at 37 °C, 180 rpm gentle agitation in Tecan (Tecan Group Ltd) microplate reader and OD_{600} was measured every 15 minutes for 16-20 hours. Growth rate each culture was estimated by measuring the slope of the growth curve in exponential phase, after induction of PBP-A expression, as well as the maximum OD at 600 nm (max OD_{600}) from the highest value of absorbance reached during the growth.

3.7.10 Growth of *E. coli* cultures under different pH

From pre-cultures, 250 ml cultures were started in LB media, grown until OD ~0.2, induced by L-arabinose (0.5%, w/v), devided into 50 ml in different erlenmeyers, and immediately buffered by adding mixture of monobasic dihydrogen phosphate (KH_2PO_4) and dibasic monohydrogen phosphate (K_2HPO) of concentrated (1 M) stock buffers in order to achieve LB medium with 50 mM potassium phosphate buffer at different pH (pH 5.8 to pH 8). The cultures were incubated at 37 °C under 180 rpm gentle agitation in and OD_{600} was measured every 20 minutes. The maximum OD_{600} was measured from as the highest value of absorbance reached during the growth in 18 hours of growth.

3.7.11 Engineering isoelectric point of PBP-As

Isoelectric point (pI) of PBP-A proteins was reduced by either removing the C-terminal 6xHis tag or introducing two nonsynonymous substitution of lysine (K) to glutamic acid (E) at positions K104 and K212 located on surface of PBP-A protein. These substitutions would decrease the pI of PBP-A-wt-His and PBP-A-L158E-His for 1.5 and 0.96 units (Table 3-1).

Table 3-1: Theoretical isoelectric point of proteins were calculated using *ExPASy Server (https://web.expasy.org/compute_pi/)*.

Protein	**pI values**		
	+6xHis	No 6xHis	2K>E (+6xHis)
PBP-A-wt	7.56	7.2	5.99
PBP-A-L158E	6.78	6	5.82
TEM-1	5.55	5.1	

3.7.11.1 Removing 6xHis-tag from PBP-As

His-tag from N-terminal PBP-As was removed using QuickLib protocol (Galka et al., 2017). Briefly, deletion was introduced using primers upstream and downstream the target site (see primers bellow) and a single PCR amplification of pBAD43_PBP-A. The PCR product of linearized plasmid+gene is purified and then circularized using Gibson assembly method, thanks to overlap sequences incorporated into primers, followed by transformation into *E. coli* competent cells.

Primers used for His-tag removal from pBAD43_PBP-A (overlapping sequences at 3' end of primers are indicated as underlined):

PBP-A NoTag fw:

ATCAAGCCTTTGAAAAACTGTCGCCGCCGTGAGTTGAATGGGCGAAAGCC

PBP-A NoTag rv:

CGGCGGCGACAGTTTTTCAAAG

Substitution of two Lysines (K) with Glutamic acid (E) on surface of PBP-As

Lysine at positions 104 and 212 were mutated to Glutamic acid by incorporating mutations on primers followed by two PCR reactions on pBAD43_PBP-As template. Reaction 1 to amplify *pbp-A* gene from position 104 to 212 and reaction 2, for amplification of the remaining backbone of pBAD43_PBP-As. The PCR products of reaction 1 and 2 were purified and then assembled (thanks to overlap sequences incorporated into primers), by mixing 3-4 folds molar excess of the shorter fragment (reaction 1) with 50-100 ng of longer backbone (reaction 2) and

0.5 µL of DpnI (20 U/µL) into 15 µL Gibson master mix (see section 7.2.14 for composition), followed by transformation into *E. coli* competent cells.

Following primer pairs were used for the two reactions. Overlapping sequences, at region 104 and 212 are indicated as underlined and bold, respectively:

Reaction 1:

Pbb-a_K104E fw:

ATTGCCCCG<u>GAAGCAGGCACCCTGCAGTAT</u>CAAGAACCGAATTCA

CAATACGCAG

Pbb-a_K212E rv: **CGGCAGCAGGGTATTGGTAACGGT**G

Reaction 2:

Pbb-a_K212E fw:

ACCGTTACCAATACCCTGCTGCCGGCCGGTCTGGGTGAAGGTGC

AACGATCGCTCATAAAACC

Pbb-a_K104E rv: <u>ATACTGCAGGGTGCCTGCTTC</u>

3.7.12 Isolation of *E. coli* cells lysate for peptidoglycan analysis (boiling in SDS)

From overnight pre-cultures, 4 ml was inoculated into 400 ml of LB media without any antibiotic in a sterile 2 L erlenmeyer flask and incubated at 37 °C under agitation of 180 rpm. At early exponential phase (OD600 ~ 0.2), cultures were induced for expression of PBP-As by adding L-arabinose (0.5% w/v) followed by incubation to the late exponential state (OD_{600} ~ 0.7). The cultures were rapidly cooled in an ice/water bath to 4°C and then cells were harvested by centrifugation at 4°C, 5000 rpm. The supernatants were discarded and each pellet was resuspended in 6 ml of ice-cold water. Each cell suspension was slowly (drop-wise) added into 6 ml of pre-boiling 8% SDS solution under gentle stirring on a magnetic stirrer in an Erlenmayer flask. Note that drop-wise addition of cell suspension into boiling SDS is important to ensure rapid desintegration of the cells that is required for the fast inactivation of the endogenous autolysins. Samples were boiled under stirring for 30 minutes. If excess foam developed, addition of some drops of water would help. Also water is added if the volume(s) of the boiling solution decreased too much. The suspension should boil at low

heat all the time but do not too much heating to not turn dark. Finally, samples were cooled down to room temperature and transferred to Falcon tubes and kept at room temperature until further treatment and HPLC-MS analysis. HPLC-MS analysis was carried out as explained in section 2.7.4.2, chapter 2.

Chapter 4

Role of disulfide bond in the evolutionary emergence of serine β-lactamases

Gol Mohammad Dorrazehi, Hervé Degand, Damien Evrard, Waldemar Vollmer and Patrice Soumillion

Abstract

Although disulfide (S-S) bonds are predominantly known for their role in protein stability, their impact on evolution of conformational dynamic and activity of enzymes is overlooked. S-S bonds can narrow the conformational spectrum in enzyme structure and, therefore, result in shifting their global dynamic space. This is particularly important if the S-S bond is formed in the active core of the enzyme or between motifs that are directly or indirectly involved in shaping the active site. In many class A β-lactamases including TEM-1, a S-S bond is holding two α-helices together, which are ending in active pocket of the enzyme by providing the catalytically important SXXK and SXN motifs. In this chapter, we investigate the impact of introducing the conserved Cys77-Cys123 S-S bridge from TEM-1 into the PBP-A DD-peptidase from *Thermosynechocchocus elongatus*, a phylogenetically related enzyme that we have repeatedly failed to convert into a β-lactamase. We previously showed that PBP-A expression in *E.coli* generates a fitness cost due to its activity on peptidoglycan. Here, we report that introduction

of S-S bond in PBP-A results in fitness improvement. This improved fitness is related to a decrease in activity rather than increase in stability of the enzyme. Interestingly, the reactivity of PBP-A towards penicillin is not affected, suggesting adaptive role of this S-S bond during evolution of β-lactamases from PBPs. Our findings highlight an overlooked role of S-S bonds as molecular devices allowing to tune the dynamics and activity spectrum of proteins.

4.1 Introduction

Protein dynamics is considered as the heart of enzyme catalysis. Promiscuous activities in the same structural fold of enzymes originate from conformational versatility in their active core, which can provide organization of atoms in different orientations to accommodate and react with secondary substrate(s) (Babtie et al., 2010a; Ben-David et al., 2012; James & Tawfik, 2003a; Khersonsky & Tawfik, 2010; Tokuriki & Tawfik, 2009a). Although such plasticity can provide a good starting point for artificial evolution towards neofunctionalization of enzymes (Khersonsky & Tawfik, 2010; Pandya et al., 2014) by providing a malleable state to be shaped by mean of natural selection, the capacity of cellular context to accommodate generalist intermediate(s) in evolutionary trajectories has to be considered with caution. Indeed, fitness-costly intermediates of enzymes associated with promiscuity will require urgent adaptation either in organismal or protein level to bypass a low-fitness valley(ies) in rugged evolutionary trajectories (Goldsmith & Tawfik, 2017). On the other hand, it is widely accepted that protein stability and mutational robustness also promote evolvability by providing tolerance to destabilizing but functionally beneficial mutations, i.e., by providing a robust scaffold to maximize tolerance against destabilizing mutations (Bloom et al., 2006; Masel & Trotter, 2010; Tokuriki & Tawfik, 2009b). Therefore, to promote evolvability of proteins, a concession between conformational diversity in catalytic core and robustness of structural scaffold would be a plausible scenario.

Role of conformational dynamics in enzyme evolution is receiving increasing attention (Bunzel, Anderson, & Mulholland, 2021; Galdadas et al., 2021; Otten

et al., 2018; Schenkmayerova et al., 2021) (see also section 1.4 in introduction), mostly on taking advantage of molecular dynamics to understand the relationship between conformational movements and the catalytic behavior of enzymes. Studies on conformational dynamics of proteins are providing new insights on intramolecular mechanisms driving catalysis in enzymes. For example, such insights are proposing mechanisms in controlling dynamics of active core of enzymes by (i) proximal or distal loops (Liao et al., 2018), (ii) covalent intramolecular interactions by disulfide (S-S) bond formation (Astudillo, Bernad, Derrien, Sebban, & Miksovska, 2010; Silvers et al., 2012; Simakov et al., 2017; Haoming Zhang, Kenaan, Hamdane, Hoa, & Hollenberg, 2009), and (iii) by movement of dynamic networks from surface all the way to the active center of enzymes (Agarwal, 2019; Bunzel et al., 2021; Galdadas et al., 2021). These studies are increasing our understanding on allosteric regulation of proteins and how distal modifications can communicate/propagate an effect to the catalytic core of enzymes.

Since conformational diversity of the catalytic core in enzymes is the origin of promiscuity and extension of substrate spectrum (James, Roversi, & Tawfik, 2003; James & Tawfik, 2003a), alterations in conformational space of this core would affect the spectrum of such unspecific activities. Such suboptimal unwanted activities generally prove problematic in cellular pool by introducing a high toll on fitness of the cell or organism (O'Loughlin et al., 2006). Illustrating this, metabolic repair enzymes have evolved to minimize the costs due to promiscuous activities of various enzymes in the cell (Bommer, Van Schaftingen, and Veiga-da-Cunha 2020; Veiga-da-Cunha, Van Schaftingen, and Bommer 2020). On the other hand, minimizing deleterious side activities should obviously be a driving force for the evolution of highly specific enzymes. Therefore, either through intragenic or intergenic (extragenic) mutations, natural evolution has managed to cope with the costly and undesired activities.

Post-translational modifications can provide new windows of folding and dynamics in conformational space of proteins, which are not predictable from primary sequences of proteins. S-S bonds are unique post-translational

modifications in proteins' fold that requires two appropriately positioned cysteine residues and their thiol group has to be oxidized to form the covalent S-S bond. By reticulating the polypeptide chain, an S-S bond will permanently pull two parts together, with a strong impact on the folding state and dynamics of the protein.

Even though proteins evolve by acquiring mutations one at a time, emergence of a disulfide bond requires the acquisition of a pair of cysteines. Phylogenetic data shows that, once a disulfide bond is acquired, it will be rarely lost. It is believed that disulfide bond formation is positively correlated with increase in functional sophistication of proteins in more complex species (Wong, Ho, & Hogg, 2011). Function of some secreted soluble proteins and cell-surface receptors in complex organisms is suppressed by formation of disulfide bonds during their folding and activated by cleavage of the bonds (Hogg, 2003; Sun, Skorstengaard, & Mosher, 1992). Disulfide bonds are also known be important in signal transduction, thiol protection, and regulation of redox homeostasis through thiol–disulfide exchange reactions (Messens & Collet, 2013; Van Laer, Hamilton, & Messens, 2013). Nonetheless, gaining (as well as loosing) disulfide bond in evolution of proteins is a costly process because intermediates with a single cysteine or unpaired cysteines may lead to incorrect disulfide bond formation, protein aggregation or increased sensitivity to oxidation (Mehlhoff et al., 2020).

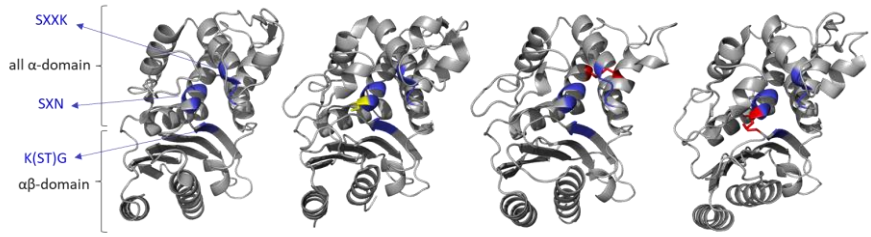
Formation of disulfide bond as a well-known intramolecular reaction, which can narrow the conformational spectrum in proteins' structure, can result in variation in protein dynamic (Astudillo et al., 2010; Silvers et al., 2012; Haoming Zhang et al., 2009) and, in case of enzymes, may impact activity by altering the dynamics of catalytic motifs and residues. Even a distal S-S bond has been shown to stabilize the catalytic center on a β-lactamase enzyme by changing the allosteric network across the protein's fold thus leading to preorganization of the catalytic residues' orientation and altering the conformation of the Ω loop (Simakov et al., 2017). Growing evidences are showing that evolutionary role of disulfide bonds can be more than just for stability purposes, such as roles in regulation of conformational dynamics and/or acting as switches for protein's function

(Astudillo et al., 2010; Gilbert, 1990; Hogg, 2003; Messens & Collet, 2013; Pace, Grimsley, Thomson, & Barnett, 1988; Silvers et al., 2012; Wetzel, 1987; Haoming Zhang et al., 2009). However, the role of disulfide bond is mostly regarded to be linked with stability of proteins (Fass, 2012; Hatahet, Boyd, & Beckwith, 2014) and, although some studies pay attention on effect of S-S bond on activity and substrate spectrum of enzymes, mainly β-lactamases, (Majiduddin & Palzkill, 2003; Raquet, Lamotte-Brasseur, Bouillenne, & Frere, 1997; Schultz, Dalbadie-M~farland, Neit~el, & Richards', 1987; Shimizu-Ibuka, Matsuzawa, & Sakai, 2004; C. A. Smith et al., 2016), there is still a gap in literature to fully uncover the evolutionary role of disulfide bonds with regard to promiscuity and for fitness adaptation of proteins.

Disulfide bonds are rarely formed in intracellular (cytoplasmic) proteins and they are mostly known to be abundant in extracellular proteins, such as antibodies. This is because of the low redox potential in the cytoplasm. However, in the oxidizing periplasmic compartment of bacterial cell envelope, the majority of cysteines are consistently oxidized into disulfides (Arts, Gennaris, & Collet, 2015a). Enzymes catalyzing disulfide bond formation and isomerization are found in all living cells. Well known examples are protein disulfide isomerases (PDI) in the endoplasmic reticulum of eukaryotes or DsbA in the periplasm of *E. coli*.

The evolution of disulfide bonds in class A β-lactamases becomes more interesting when within the same structural scaffold of these enzyme a disulfide bridge is very often found at two distinct structural positions with a cysteine just before the catalytic Ser70 (in SxxK motif) (C69) or just after it (Cys77) (Figure 4-1) (Philippon et al., 2016). Previous studies have reported a conserved disulfide bridges mainly at two different structural positions of class A β-lactamases, between Cys69-Cys238 (in SME- and KPC-types) or Cys77-Cys123 (in TEM- and SHV-type enzymes) while some members of this class has no S-S bond (such as Toho-1 and CumA) (Figure 4-1) (Philippon et al., 2016). Among those class A β-lactamases with no S-S bond, either one cysteine residue is present (e.g., in Toho-1), or no cysteine is present (e.g., in CumA). On the other hand, PBP-A

family, which has closest sequence similarity, among known PBPs (Fröhlich et al., 2021; Urbach et al., 2008), and high structural similarity to class A β-lactamases, does not harbor any disulfide bond (Figure 4-1).



Position of the S-S bond	No SS bridge	No SS bridge (Cys69)	Cys77-Cys123	Cys69-Cys238
Enzyme	Murein hydrolase	Class A β-lactamase		
	PBP-A	Toho-1	TEM-1	SME-1
Function(s)	DD- and LD-CPase, DD-EPase	Activity on cefotaxime, (Extended Spectrum)	Activity on Penicillins and cephalosporins	Carbapenemase

Figure 4-1: Absence or conservation of a disulfide bond in different structural position in class A β-lactamases compared with PBP-A DD-peptidase. PBP-A DD-peptidase has no cysteine residue in its primary sequence and, therefore, contain no S-S bond (PDB: 2J7V). While among class A β-lactamases, a conserved disulfide bond exists at two different structural positions or is absence in the structure. In TME- and SHV-type enzymes (e.g., TEM-1, PDB: 1BTL) a S-S bond is formed between Cys77-Cys123 crosslinking α2 and α5 helices, in SME- and KPC-types (e.g., SME-1, PDB: 1DY6) a S-S bond is formed between Cys69-Cys238 in the active site and acting as a third hinge between all α and αβ domains, and in some others such as CumA and Toho-1 (PDB: 1IYS), no S-S bond is formed. S-S bonds are shown in red, free cysteine residues in yellow and three catalytic motifs (SXXK, SXN and K(ST)G) in blue. Structures are visualized and modified by PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

Vicinity of both of the S-S bridges in class A β-lactamases to the active site of enzymes make them potentially important targets for evolutionary studies. The 69-238 disulfide bridge, which is located at the active site of carbapenemases, was shown to be important for their β-lactamase activity (Majiduddin & Palzkill, 2003; Philippon et al., 2016; Raquet et al., 1997). As reported by Majiduddin and Palzkill (2003) codon randomization of Cys69 and Cys238 (individually, as Cys69X or Cys238X, and together, as Cys69X-Cys238X) in SME-1 enzyme has led to selection of functional clones that only possessed cysteines at positions Cys69 and Cys238, suggesting the essential role of the disulfide bond in β-lactamase activity. Such effect is expectable because, in addition to critical location of 69-238 bridge in the active site, which is constraining the overall

topology of the active site, more importantly, the bridge is introducing a drastic conformational change in the general fold of the protein by acting as a third hinge between the all α-domain and αβ-domain in serine β-lactamases. On the other hand, the 77-123 bridge in TEM enzymes was reported as not essential for the kinetic properties on β-lactam hydrolysis (Laminet & Plückthun, 1989; Schultz et al., 1987).

Thanks to the β-lactamase DataBase (BLDB, available at http://www.bldb.eu) that has been developed by Iorga and colleagues (Naas et al., 2017), we checked all the known sequences (1378 sequences) of class A β-lactamases for presence as well as position of cysteine residues in their corresponding protein sequence. Interestingly, majority of class A β-lactamases (59%) contain a disulfide bridge at one of the two known positions (77-123 or 69-238) (see Figure 4-2). Among class A β-lactamases that do not carry both cysteines for these bridges (559 enzymes), single cysteine in position 69 is highly conserved (72%), less conserved (11%) at position 135 (Ambler numbering (Ambler et al., 1991)) next to the SDN motif in the α_6 helix and much less in position 123 (6%) and position 77 (4%).

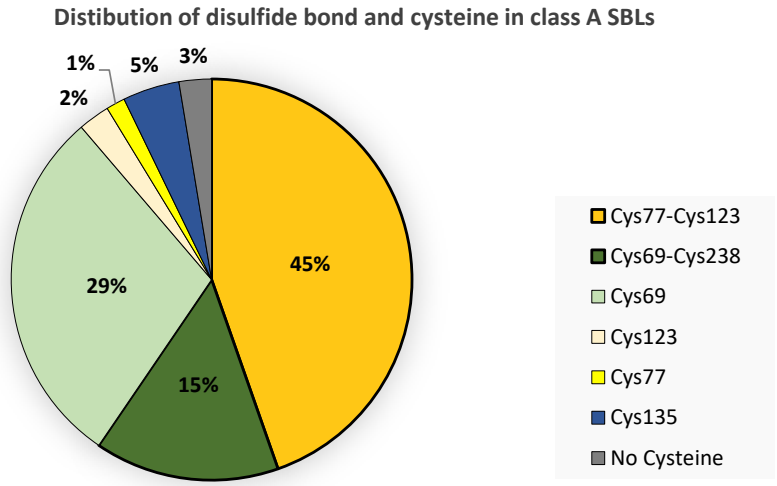


Figure 4-2: Distribution and positions of disulfide bonds and cysteine residues in class A β-lactamase. All the currently known sequences (1378 enzymes) of class A serine β-lactamases (SBLs) available at BLDB database are analyzed. Amino acid sequences of mature proteins (excluding putative signal peptide) were based on Ambler numbering. Number of enzymes in each category are available in Appendix 9-22.

Another disulfide bridge, distal to active site (allosteric site), is conserved in OXA-1 clade of class D β-lactamases (Szarecka, Lesnock, Ramirez-Mondragon, Nicholas Jr, & Wymore, 2011) and was shown to significantly affect the organization of catalytic residues and conformational plasticity of the active site pocket (Simakov et al., 2017). Another study by Sakai *et al.,* (2006) has shown that although introduction of 77-123 disulfide bond in Toho-1 (CTX-M-44) class A β-lactamases, which originally has no disulfide in its structure, did not impact its catalytic properties, it granted a significant *in vivo* advantage for drug resistance when expressed in *E. coli* (Shimizu-Ibuka et al., 2004; Shimizu-Ibuka, Matsuzawa, & Sakai, 2006).

Class A β-lactamases are structurally similar and studies are suggesting that conformational dynamics of active site loops in these enzymes play an important role in evolution of their activities (Galdadas et al., 2021). Highly conserved and catalytically essential residues, such as Glu166 and Asn170, are reported to influence the conformational plasticity of the omega-loop (Banerjee, Pieper, Kapadia, Pannell, & Herzberg, 1998). Furthermore, recent studies show that conformational changes in active site pocket are responsible for differences in the spectrum of activity between KPC-2 and KPC-4 carbapenemases (Galdadas et al., 2018; Tooke et al., 2021).

It is generally accepted that introduction of an intramolecular covalent bond in proteins' fold will shift the conformational versatility by holding different regions of the protein together. In this study, we evaluate the effect of introducing the conserved Cys77-Cys123 bridge from class-A β-lactamases into a phylogenetically related DD-peptidase (or Penicillin Binding Protein, PBP), called PBP-A. PBP-A has high structural similarity with class A β lactamases, which suggest a common ancestral PBP as an evolutionary intermediate between PBP-A and class A β-lactamases (Urbach et al., 2009). Several attempts in our laboratory to convert PBP-A to a β-lactamase, through directed evolution, has failed and a fitness cost associated with expression of PBP-A has been suggested as the main obstacle behind this failure. Therefore, we first evaluate the fitness effect upon introduction of the S-S bond, and then we will assess evolvability in

presence of the bridge by conducting directed evolution experiments. We further characterize the effect of S-S bond on stability as well as catalytic properties of PBP-A and its reactivity towards penicillin. Our results suggest a fitness-improving role of the S-S bond in PBP-A by reducing the peptidase activity without compromising the penicillin acylation capacity. Based on these observations, we suggest an adaptive role, rather than (only) stability, of the S-S bonds in evolution of class A β-lactamases from PBPs.

4.2 Results

4.2.1 Disulfide bond formation in PBP-A mutants

We initially aligned crystal structures of PBP-A with TEM-1 β-lactamase to evaluate the position of the S-S bond in PBP-A protein. The Cys77-Cys123 bridge exists between α_2 and α_5 helices in class A β-lactamases like TEM-1 and since the α_2 and α_5 of PBP-A are perfectly superimposed with that of TEM-1, the missing cysteines were introduced at the aligned positions of Leu68 and Ala115 on PBP-A by site directed mutation (Figure 4-3).

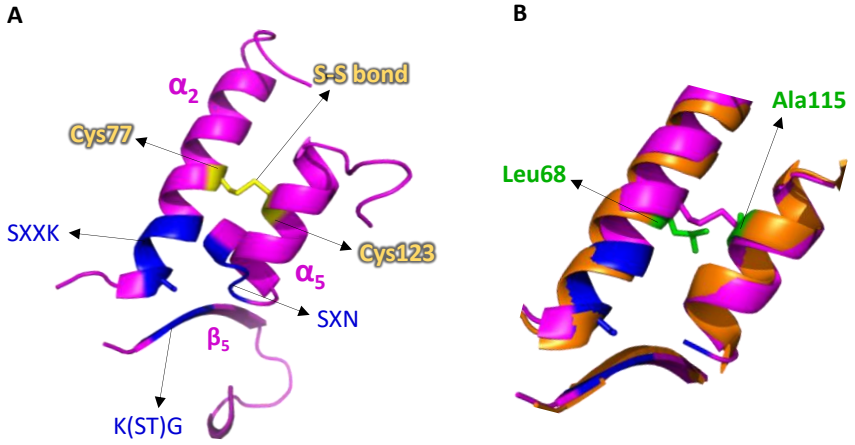


Figure 4-3: Position of disulfide bridge (S-S bond) in TEM-1 β-lactamase and its corresponding position in PBP-A. **A)** A conserved disulfide bond is bridging α_2 and α_5 helices via Cys77 and Cys123 in structure of TEM-1 β-lactamase (PBD: 1BTL). **B**) Structural alignment of PBP-A (PDB: 2J7V) in orange and TEM-1 in magenta shows a good superimposition of α_2 and α_5 helices, where the Leu68 and Ala115 of PBP-A are superimposed with the Cys77 and Cys123 of TEM-1. Three conserved catalytic motifs (SXXK, SXN, K(ST)G) are colored blue in TEM-1, where the side chain of the active

Ser70 is shown by sticks. (Structures are visualized and modified by PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC).

Prior to any characterization, the formation of S-S bond between the newly introduced cysteines in PBP-A need to be confirmed. For that we compared cysteine-containing PBP-As with no-cysteine variants based on their migration on non-reducing SDS-PAGE gel and mass analysis.

Detection of disulfide bond on non-reducing gel

Because proteins with disulfide bond(s) in their structure will have a more compact state than their non-bridged counterparts, they would migrate slightly faster on SDS-PAGE gels under non-reducing conditions. We compared migration of cysteine-containing PBP-As (wt and L158E) with their no-cysteine counterparts, as well as TEM-1 with TEM-1-ΔSS (S-S bond in TEM-1 was eliminated by mutating Cys77 and Cys123 to alanine), under reducing and non-reducing conditions, by subjecting soluble periplasmic extracts to SDS-PAGE, followed by western blotting (Figure 4-4).

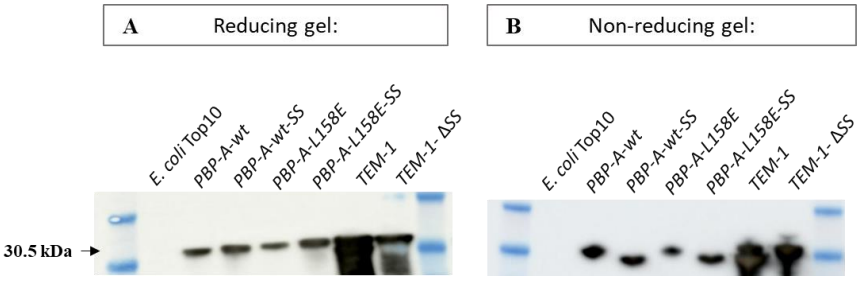


Figure 4-4: Western blotting analysis of His-tagged PBP-As and TEM-1, present in periplasmic extracts of corresponding *E. coli* cells, migrated on SDS-PAGE gels using either reducing (A) or non-reducing (B) sample buffers and detected with anti-Histag antibody. **A)** When samples are treated with reducing and denaturing sample buffer (containing β-mercaptoethanol and SDS) prior to loading on SDS-PAGE, the disulfide bond is reduced and thiol groups in cysteins will remain in their free form. Under such reducing circumstances, there will be no different between migration speed of proteins with cysteines and without cysteins. **B)** When a non-reducing sample buffer is used, the disulfide bond has remained oxidized and, therefore, due to the more compact state of the proteins they migrate faster than the same protein without S-S bond. The S-S bond in TEM-1 was eliminated by mutating Cys77 and Cys123 to alanine.

Results in Figure 4-4 shows a faster migration of the cysteine-containing proteins under non-reducing SDS-PAGE condition, which confirms the formation of S-S bond in their structural fold.

Detection of disulfide bond by mass analysis of intact proteins

To further confirm SS bond formation, we compared total mass of purified PBP-A-wt and PBP-A-wt-SS obtained from liquid chromatography–mass spectrometry (LC-MS). After substitution of Leu68 (113.15 Da) and Ala115 (71.07 Da) by two cysteines (Cys, 103.13 Da) in the PBP-A-wt (30574 Da), a mass of 22 daltons will be added to the total protein mass. If the two cysteines are oxidized to form the SS bond, two hydrogens will be released, therefore the mass of oxidized PBP-A-wt-SS will be 30594.

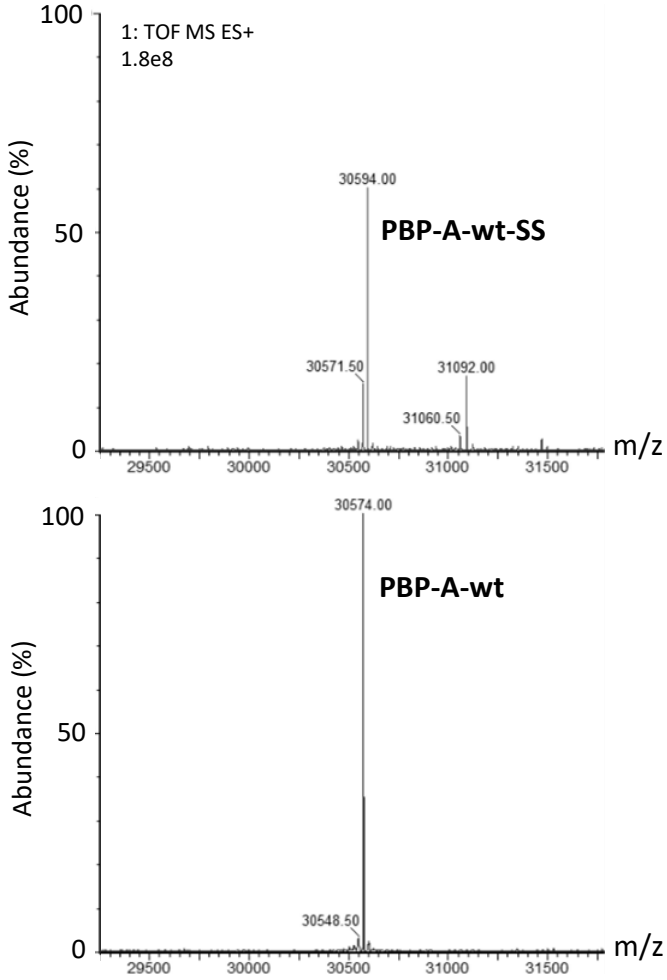


Figure 4-5: Disulfide bond confirmation in PBP-A by total mass of protein. Mass of intact proteins are presented for PBP-A-wt (30574 Da) and PBP-A-wt-SS (30594 Da).

The difference of 20 daltons corresponds to substitutions of Leu68 and Ala115 to cysteins, while the thiol (SH) groups of the cysteines are oxidized and a disulfide bond is formed.

MS analysis in Figure 4-5 confirms presence and reductions of the two cysteine residues into oxidized cystines in the PBP-A-wt-SS, and thus formation of the Cys68-Cys115 S-S bond.

4.2.2 Expression level of PBP-A-SS is not significantly altered

Since introducing cysteines (or disulfide bond) in periplasmic proteins may affect their expression (as cysteine is a relatively rare amino acid) and/or export to the periplasm of *E. coli* (e.g., by effecting the traffic in export pathway and/or requiring periplasmic chaperonin systems such as Dsb proteins to catalyze the disulfide bond formation), we decided to evaluate the periplasmic expression level of PBP-A proteins before and after introduction of cysteines/disulfide bond. For that, we compared amount of PBP-A in periplasmic extracts of corresponding *E. coli* cells from cultures with the same cell densities (i.e., normalized to the same OD_{600}). Amount of PBP-A proteins were estimated semi-quantitatively by western blot analysis using ImageJ image processing software and the results are presented in Figure 4-6.

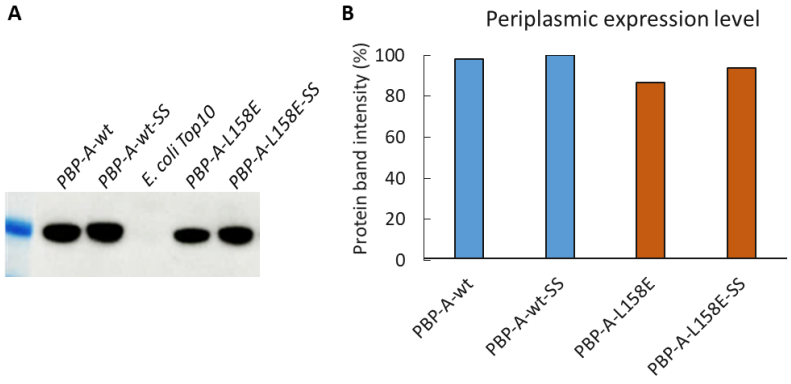


Figure 4-6: Periplasmic expression level of different mutants of PBP-A. **A)** Western blot analysis of different PBP-A in periplasm of *E. coli*. All the samples were normalized to the same OD_{600} prior to periplasmic extraction. **B)** Band intensity of different PBP-A proteins in western blot image were quantified using ImageJ image processing software. Introducing a disulfide bridge between positions 68 and 115 of PBP-A does not significantly change the expression level of periplasmic protein.

Results from western blot analysis of periplasmic extracts from equal cells (Figure 4-6) numbers show that expression level of all PBP-A variants are similar.

4.2.3 Effect of disulfide bridge on fitness cost of PBP-A expression in *E. coli*

Previously in chapter 3, we observed fitness costs related to expression and activity of PBP-A-wt and L158E in *E. coli*. The fitness cost was phenotypically detectable as growth defect, sensitivity against vancomycin and alteration in peptidoglycan (PG) profile of the host *E. coli* cells. Here, we evaluate effect of the PBP-A's Cys68-Cys115 disulfide bond on these phenotypes.

Disulfide bridge introduction in PBP-A partially suppress growth defect of E. coli expression host

To evaluate effect of the S-S on the growth of host bacteria, we compared growth parameters of the *E. coli* strains expressing different mutants of PBP-A from low copy number expression plasmid (pBAD43). Results in Figure 4-7-A represent the growth curve of *E. coli* cultures under induction for PBP-A expression (0.5% L-ara) and in Figure 4-7-B, the relative maximum population density (Max OD_{600}) of cultures under induced expression of PBP-As (0.5% L-arabinose) to the basal (non-induced) expression are presented.

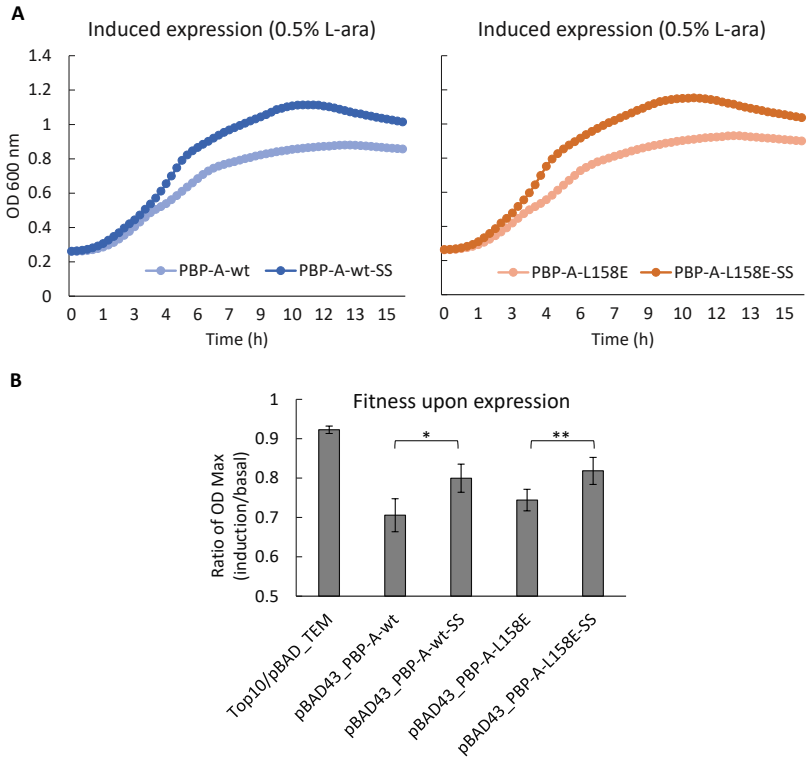
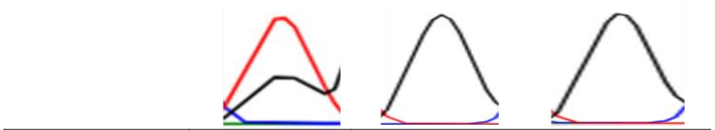


Figure 4-7: Advantage of introducing a disulfide bond in PBP-A proteins on growth of the host *E. coli* cells. **A)** Growth curves of *E. coli* Top10 under expression of PBP-A-wt, PBP-A-wt-SS, PBP-A-L158E and PBP-A-L158E-SS shows a significant growth advantage of disulfide containing PBP-As compare to variants without disulfide bond. **B)** Comparison of OD max of different PBP-A variants. The bar chart illustrates the relative OD max between induction and basal conditions, which can be considered as the fitness of cells. The growth of S-S containing variants show improvement in OD max (fitness) upon introduction of S-S bond in PBP-A-wt and PBP-A-L158E that are obtained as 13.3% and 10.0%, respectively. $*P < 0.02$, $**P < 0.04$, $n = 3$, t test.

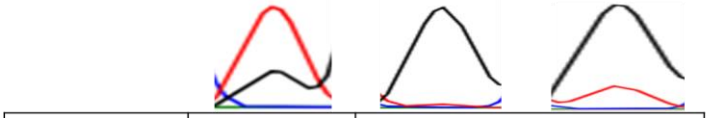
We observed a significant improvement on growth of *E. coli* cells expressing disulfide bonded PBP-A proteins in comparison to no-disulfide variants (Figure 4-7). Introduction of the Cys68-Cys115 disulfide bond, has improved the maximum OD of the *E. coli* cells expressing PBP-A-wt and PBP-A-L158E for up to 13.3% and 10.0%, respectively.

Disulfide mutants of PBP-A win growth competition

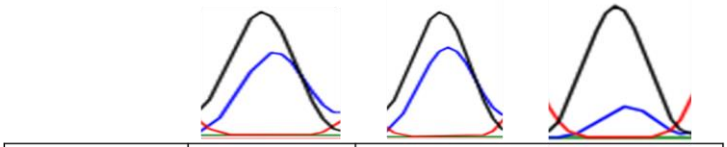
To further confirm the fitness-improving effect of the S-S bond on PBP-As, we conducted competition assays. Fitness of a bacteria expressing different proteins or different mutants of the same protein can be determined by fitness competition assay and sequencing of a diagnostic PCR product corresponding to the gene of the proteins, and then sequencing chromatograms of the mutated region can be compared. If there is an absolute winner among competitor mutants, only chromatogram of the winner mutants are detectable in sequencing results, but if there is no strict winner or no winner at all, then the chromatograms of both (or all) competitors will be present and the more fit clone can be determined from the ratio of the respective chromatograms, compare to the starting day. Therefore, to evaluate fitness advantage of introducing the S-S bond into PBP-As, growth competition for ~30 generations in LB-media under basal and induced expression were carried out and the dominant mutants (mutants with higher fitness) of PBP-A were determined by analyzing chromatograms of sequencing results. *E. coli* Top10 cells expressing PBP-A-wt, PBP-A-L158E and TEM-1-ΔSS (containing C77A, C123A mutations) were subjected to competition against their SS-containing counterparts. The fitness competitions were conducted for 3 days (~30 generations), by every 24 hours serial passages with 1/1000 dilutions (~10 generations per day (Knöppel et al., 2018)). The ratio of competitors was estimated by analyzing sequencing results based on chromatogram peaks of mutations between two constructs.



	t = 0	t = day 3	
		Basal	Induction
PBP-A-wt/wt-SS	≈ 50/50	≈ 0/100	≈ 0/100



	t = 0	t = day 3	
		Basal	Induction
pBAD_PBP-A-L158E/L158E-SS	≈ 50/50	≈ 0/100	≈ 12/88



	t = 0	t = day 3	
		Basal	Induction
TEM-1-wt/TEM-1-ΔSS	≈ 50/50	≈ 50/50	≈ 82/18

Figure 4-8: Fitness competition between S-S containing and non-S-S proteins. Outgrowing *E. coli* cells expressing PBP-A-SS proteins against cells expressing PBP-As without S-S bond in 3 days, indicate a significant fitness advantage of the S-S bond in these proteins. Although not as significant as PBP-A proteins, absence of S-S bond in TEM-1 β-lactamase (TEM-1-ΔSS) shows lower fitness than the wild-type TEM-1, which contain the S-S bond. Ratio of bacterial strains in culture were estimated from DNA chromatogram traces in a competition culture. A dominant mutation, estimated from height of peaks, indicated the protein with higher fitness (see text). Each day (24 hours) of serial passage of *E. coli* cultures is approximately 10 generations. In sequencing chromatograms, a red peak is a T (thymine) from CTG (Leu68) of non-bridged PBP-As, a black peak is a G (guanine) either from TGC (Cys68) of PBP-As with S-S bond or TGT (Cys77) from wild type TEM-1 and a blue peak is a C (cytosine) from GCC (Ala77) from TEM-1-ΔSS. For PBP-A mutants, analyses were focused on L68C mutation, where CTG (Leu) represents non-bridged proteins and TGC (Cys) corresponds to PBP-As with S-S bond. Based on trace colors of chromatograms, a red peak is a T (thymine) from CTG (Leu) of non-bridged PBP-As while a black peak is a G (guanine) from TGC (Cys) of PBP-A-wt-SS or PBP-A-L158E-SS. For the starting mix (t=0), which contains 50/50 of PBP-A-wt/wt-SS or pBAD_PBP-A-L158E/L158E-SS, the ratio is considered 50/50 of T/G. For the samples of day 3, the height of the T peak was converted to percentage relative to its height at t=0. Similar measurements were applied for TEM-1 mutants based on C77A mutations, where a black peak represents G from TGT (Cys) and a blue peak represents C (cytosine) from GCC (Ala).

For competition between PBP-A mutants, under both conditions (basal and induction), the clones expressing PBP-A with S-S bond had outgrown the mutants without the S-S bond, therefore, confirming the fitness improvement upon introduction of the S-S bond in PBP-As. The outgrowth was absolute for PBP-A-wt-SS, while for PBP-A-L158E-SS the non-bridged mutants were detectable after 3 days of competition (Figure 4-8). For TEM-1 competitions under induction conditions, the cells expressing wild type TEM-1 which contains a S-S bond has outgrown the clones with TEM-1-ΔSS, while in basal conditions both competitors survived equally (Figure 4-8).

Introducing S-S bond in PBP-A recovers host cells' susceptibility against vancomycin

We have shown in chapter 4 that *E. coli* cells expressing PBP-A or PBP-A-L158E were more sensitive to vancomycin compare to wild-type *E. coli* cells. Here we compared the vancomycin susceptibility of cells expressing the different PBP-A variants.

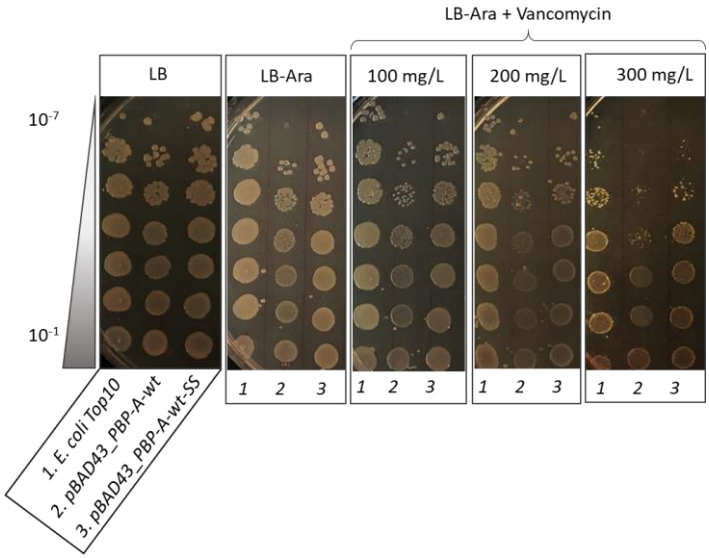


Figure 4-9: Effect of disulfide bond in PBP-As on vancomycin susceptibility of the host *E. coli* cells. Cells from *E. coli* cultures expressing different of PBP-A from pBAD43 plasmid (PBP-A-wt, PBP-A-wt-SS, PBP-A-L158E and PBP-A-L158E-SS) are normalized for similar OD_{600} prior to serial dilution and spotting on agar plates containing different concentrations of vancomycin.

The results of spotting assay under vancomycin stress shows that *E. coli* cells expressing PBP-A-SS or PBP-A-L158E-SS are less sensitive to vancomycin compared to cells expressing enzymes lacking the S-S bridge (Figure 4-9).

SS bridge introduction suppress the defects in peptidoglycan

As shown in chapter 3, the fitness cost associated with expressing PBP-A and its mutant, PBP-A-L158E, in periplasm of *E. coli* was linked to alteration in PG profile. Here, we analyze PG isolates of *E. coli* cells expressing PBP-A-wt-SS and PBP-A-L158E-SS in comparison with cells expressing PBP-As lacking the S-S bond, as well as with the wild-type *E. coli* cells. In Figure 4-10 we present abundance of muropeptides in the PG of different *E. coli* cells that were obtained by HPLC. Detailed quantitative data of all detected muropeptides are available in Appendix 9-23.

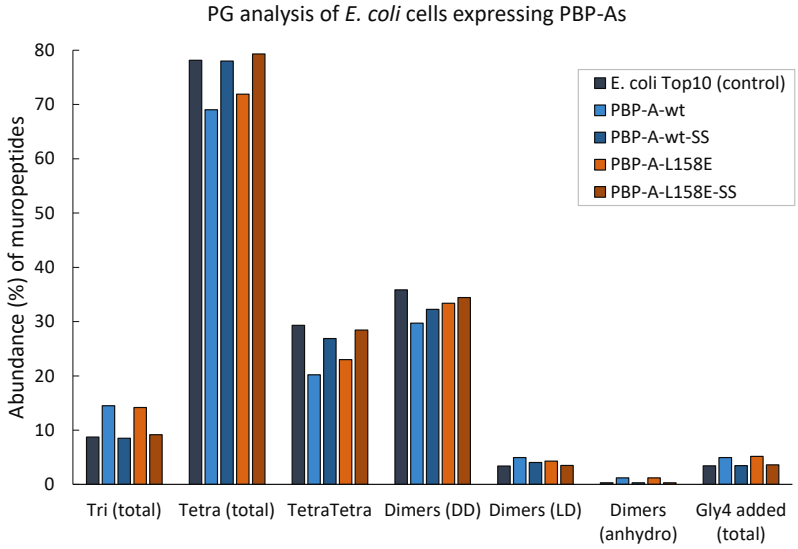


Figure 4-10: Disulfide bond in PBP-A prevents its effect on PG profile of host *E. coli* cells. Abundance of different muropeptides in PG of cells expressing disulfide-bonded PBP-As is more identical to the PG of wild-type *E. coli* cells, compare to cells expressing PBP-As with no S-S bond (i.e., the . S-S bond in PBP-As prevent their disturbing effect on PG of *E. coli*). While expression of PBP-A-wt and L158E lead to increase in total Tri, LD-dimers, anhydro-dimers and Gly4-added muropeptides, and decrease in total Tetra, TetraTetra and DD-dimers, expression of disulfide-bonded PBP-As shows no change or less change on the abundance of these muropeptides in PG. The effect of S-S bond is similarly detectable for both PBP-A-wt and L158E. For PG isolation, cells under induced expression of PBP-As were harvested at early exponential phase. HPLC-MS analysis was

carried out as explained in 2.7.4. The abundance (%) of all peptides detected in PG isolates of all the strains in this experiment are available in Appendix 9-23.

The results indicate a recovery of PG profile towards the wild-type *E. coli* cells upon introducing a S-S bond in PBP-As, compare to non-bridged PBP-As (Figure 4-10). Expression of PBP-A-wt and L158E lead to increase in total Tri, LD-dimers, anhydro-dimers and Gly4-added muropeptides, and decrease in total Tetra, TetraTetra and DD-dimers, while expression of disulfide-bonded PBP-As shows no change or less change on the abundance of these muropeptides in PG, compare to the PG from wild-type *E. coli* cells.

4.2.4 Effect of disulfide bond on properties of PBP-A enzyme

As alteration in PG profile is mainly due to activity of the PBP-A enzyme, the results in previous section suggest that improvement of PBP-A's fitness upon introduction of the S-S bond is due to effects on activity. To confirm this, we conducted *in vitro* assays to compare properties of disulfide-bonded PBP-As with the no-cysteine counterparts. We first compare kinetic parameters of PBP-A-wt and PBP-A-wt-SS for DD-peptidase activity, then we evaluate effect of the S-S bond on reactivity of PBP-A towards penicillin, and finally, we perform stability assay to evaluate the effect of S-S bond on stability of PBP-A proteins.

Disulfide bond is reducing activity of PBP-A

For kinetic characterization of PBP-A enzymes, we performed activity assays on S2d, a surrogate thiolester substrate mimicking D-alanyl-D-alanine (Urbach et al., 2008) as well as a synthetic pentapeptide mimicking muropeptides of *E. coli* peptidoglycan.

In assay with S2d at 25°C, hydrolysis of S2d by PBP-A-wt and PBP-A-wt-SS was followed spectrophotometrically at 250 nm and kinetic parameters were extracted by numerical simulation of the progression curves using Michaelis–Menten equation. Figure 4-11 represents the numerically fitted curves with the experimentally obtained values of velocity over enzyme concentration [E] measured from in assay with S2d and Table **4-1** represents kinetic parameters of each variant extracted from this assay.

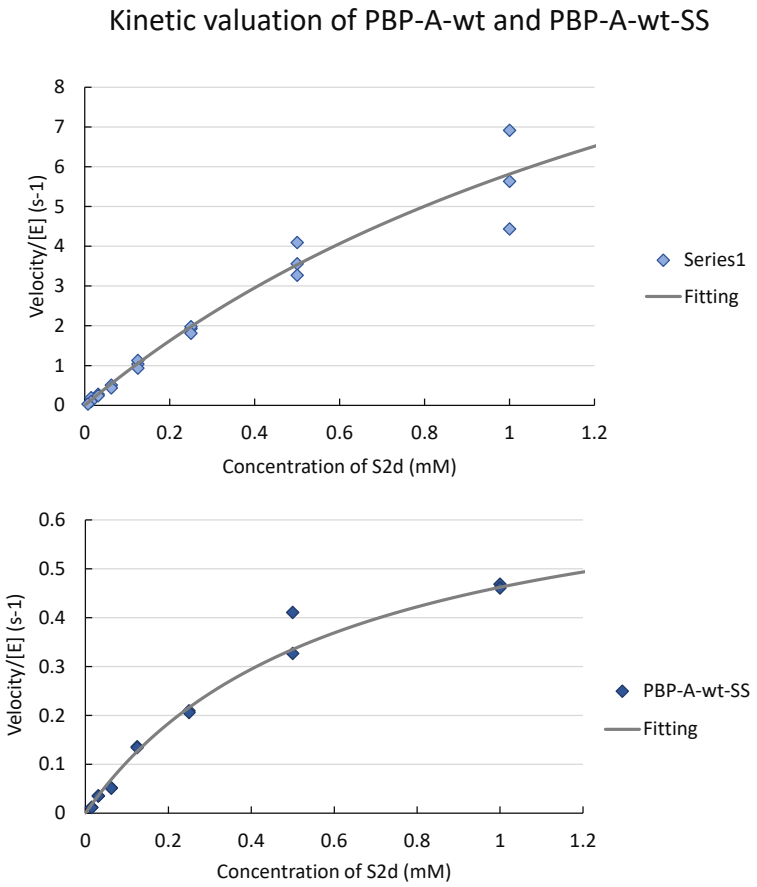


Figure 4-11: Assessment of kinetic parameters of PBP-A enzymes by fitting theoretical values obtained from Michaelis–Menten equation to experimental values obtained from assay on S2d substrate. At each concentration of S2d, at least two data points are collected.

Due to saturation of the signal from higher concentration of S2d, we were not able to collect data for concentrations higher than 1 mM. Kinetic parameters were extracted by fitting the simulated curves from the Michaelis–Menten equation to our experimental values (Figure 4-11). Because substrate range was far from saturation, only kcat/KM can be evaluated with good precision (Table 4-1).

Table 4-1: Kinetic parameters for hydrolysis of S2d by PBP-A-wt and PBP-A-wt-SS. Experiments were carried out at 25 in a microplate reader and values of k_{cat} and K_M were

calculated by fitting the simulated progression of curves, using Michaelis–Menten equation, to the experimental data.

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($M^{-1}.s^{-1}$)
PBP-A-wt	> 9	> 1	6.50E+03
PBP-A-wt-SS	> 0.65	> 0.5	0.84E+03

Kinetic parameters obtained from the activity assay on S2d, show that efficiency (k_{cat}/K_M) of disulfide-bonded PBP-A is significantly reduced (7.7 fold).

In another assay, we incubated enzymes with substrates mimicking a muropeptide of the peptidoglycan precursor. For that we used a synthetic pentapeptide that was acetylated at N-terminal L-Ala**,** with an amidated gamma-D-Glu at position 2, a mimic of mDAP in position 3 and a C-terminal (D)Ala–(D)Ala) (structure of the peptide is available in Figure 2-1, chapter 2). A DD-peptidase activity would cleave the bond between the two D-Ala at position 4 and 5, producing tetrapeptide by releasing a D-Ala. The substrate (penta) and product (tetra) in each assay were observed by mass spectrometry (MS) presented in Figure 4-12 as a semi-quantitative analysis.

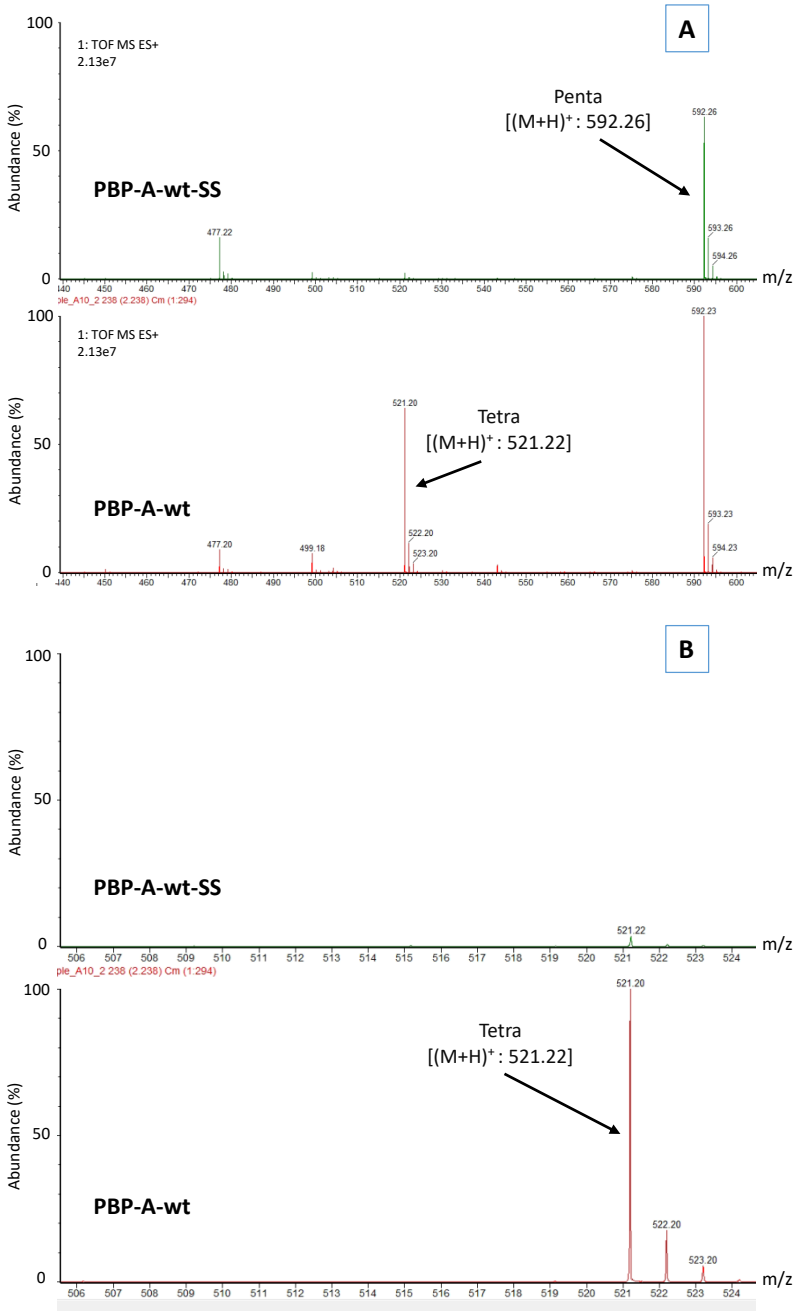


Figure 4-12: Effect of S-S bond on activity of PBP-A against synthetic pentapeptide. **A)** MS spectra represents abundance of substrate (pentapeptide, Acetyl–Ala-gamma(D)Glu$_{(amidated)}$–mDAP*–(D)Ala–(D)Ala) [(M+H)$^+$: 592.26]) and product (amidated tetrapeptide, Acetyl–Ala-gamma(D)Glu$_{(amidated)}$–mDAP*–(D)Ala [(M+H)$^+$: 521.22]) in assays of PBP-A-wt and PBP-A-wt-SS. **B)** MS spectra for detection of

product (amidated tetrapeptide). Production of Tetra is significantly decreased in assay with disulfide-bonded PBP-A. The slight reduction of Penta (substrate) is probably due to higher affinity (lower K_M) of PBP-A-wt-SS (i.e., more substrates are in form of acyl-enzyme, which were not detected in MS).

Results of both assays with S2d substrate as well synthetic pentapeptide confirm a lower DD-CPase activity of disulfide-bonded PBP-A, in agreement with the reduction of PG defects observed *in vivo*.

Disulfide is not affecting the penicillin binding capacity of PBP-A

As the S-S bond significantly impairs the peptidase properties of PBP-A, we wondered if it was correlated with a decreased reactivity with penicillin, especially since the S-S bond is conserved in class A β-lactamases, suggesting its compatibility with penicillin recognition. To assess the effect of S-S bond in reactivity of PBP-A towards penicillin, we measured the rates of acylation and deacylation of PBP-A with BOCILLIN FL penicillin, a fluorescent penicillin (Figure 4-13). Bocillin FL is a sensitive reagent allowing the detection of as low as 2 ng of acylated protein from SDS-PAGE using a fluorescence scanner (Zhao, Meier, Kahl, Gee, & Blaszczak, 1999).

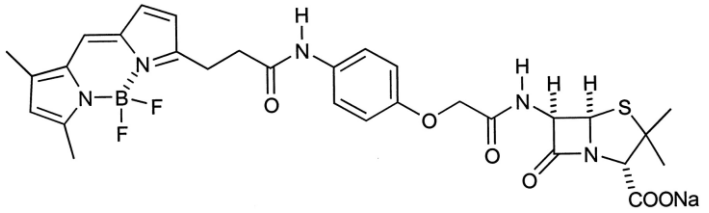


Figure 4-13: Structure of BOCILLIN FL penicillin (BOCILLIN-FL).

The reaction is started by mixing purified PBP-A enzymes and Bocillin FL. For the acylation assay, the increase in intensity is immediately followed over time, while for the deacylation assay the PBP-A is incubated with BOCILLIN to allow acylation then TEM-1 β-lactamase is added to the mix to consume the unbound BOCILLIN, then deacylation can be evaluated by sampling over time. The samples were rapidly inactivated by mixing with denaturing sample buffer and run on SDS-PAGE followed by analysis using a fluorescence scanner. This would allow us to quantify acyl-enzyme intermediates from a fluorescence signal

detectable from SDS-PAGE, from which we can measure rate constants of binding/dissociation of PBP-A with penicillin. Interaction of PBP enzymes with β-lactam antibiotics follows as the following scheme:

Scheme 4.1

where $E.S$ is the noncovalent Michaelis complex, $E-S$ is the covalent acyl-enzyme complex, and P is the released product. The rates of acylation (k_{ac}or k_2/K_S , where $K_S = k_{-1}/k_1$) and deacylation (k_3) were determined from intensity of fluorescent bands by fitting numerical simulation of the corresponding equations. For that, k_3, was derived from the first order decay of BOCILLIN-PBP complex, and using k_3 from the deacylation experiment, the optimized k_{ac} constants were determined by fitting numerically simulated values obtained from the differential equation (4.1) to the experimental data.

$$d[E-Ac] = (k_{ac}[E][S] - k_3[E-Ac])dt \quad (4.1)$$

The results of properties against penicillin are presented in Figure 4-14 and Table 4-2.

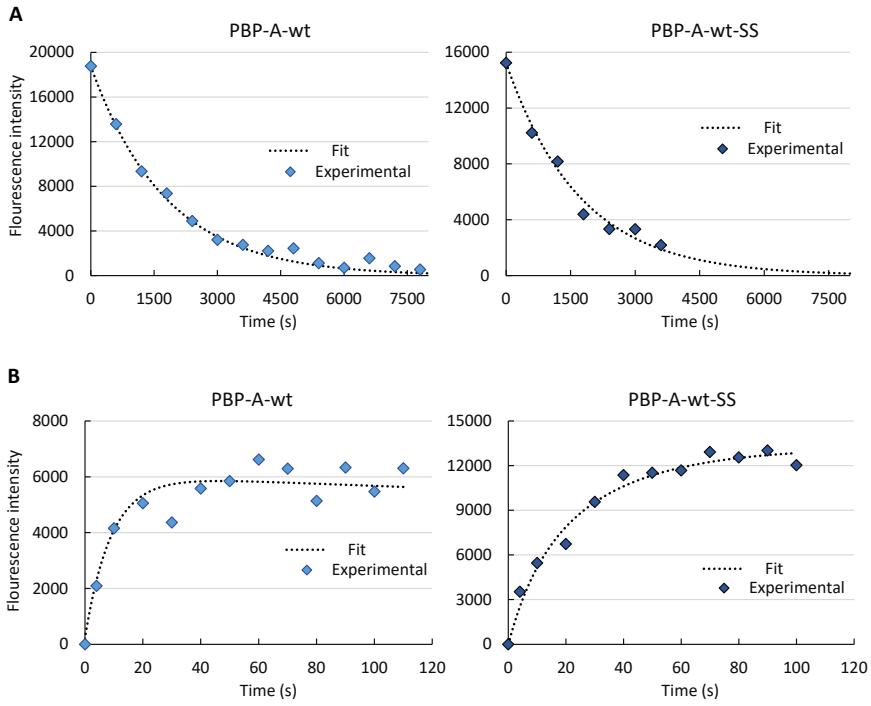


Figure 4-14: Assessing rate constants of deacylation and acylation of PBP-A-wt and PBP-A-wt-SS by florescence intensity of BOCILLIN-labelled PBP-A over time. A) Deacylation over time can be calculated by numerical fitting of an exponential decay curve to the experimentally observed values. B) Acylation rate can be extracted by numerical simulation of the differential equation 4.1 to the experimental data. The calculated rates of deacylation (k_3) and acylation (k_{ac}) from these experiments are presented in (Table 4-2).

Table 4-2: Constants of acylation (k_2/K_S or k_{ac}) and deacylation (k_3) of BOCILLIN FL penicillin at room temperature.

	k_3 (s^{-1})	k_{ac} ($M^{-1}s^{-1}$)
PBP-A-wt	5.6×10^{-4}	9.0×10^{5}
PBP-A-wt-SS	5.8×10^{-4}	5.2×10^{5}

Results from assays with BOCILLIN show that the disulfide bond in PBP-A is almost not affecting the enzyme's reactivity against penicillin. Both acylation and diacylation rate constants are in the same range when compared to the wild-type PBP-A.

Disulfide bridge is not improving the stability of PBP-A proteins

Since disulfide bonds are predominantly known for their role in increasing stability of proteins, we carried out protein stability assay using Differential Scanning Fluorimetry (DSF) to evaluate effect of the S-S bond on stability of PBP-A protein. The denaturation transition of PBP-A proteins was monitored by measuring fluorescence generated by binding of SYPRO Orange dye to hydrophobic region of proteins using a real-time quantitative PCR machine (Kellner et al., 2021). Result in Figure 4-15 present normalized fluorescence intensity of SYPRO orange as obtained through DSF assays on different PBP-A variants.

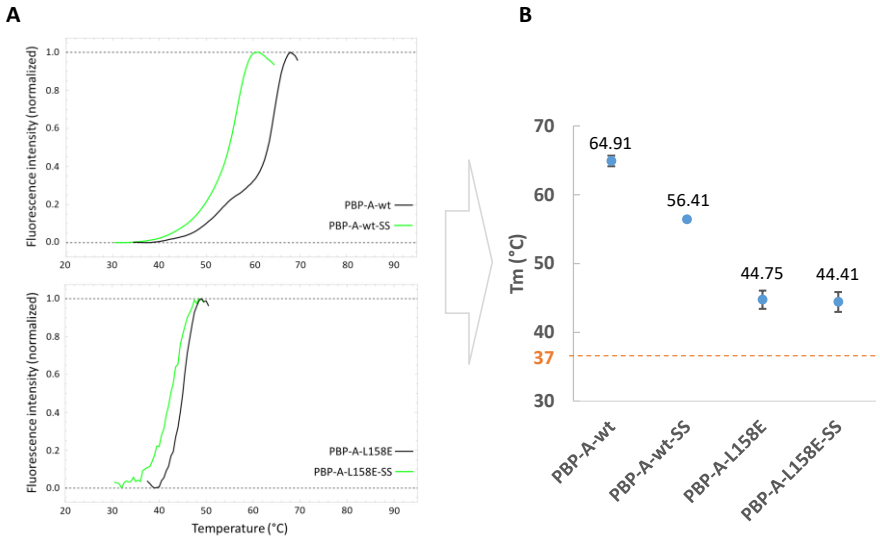


Figure 4-15: Effect of disulfide bond on the thermal stability of the PBP-A protein. Protein stability of PBP-A variants was determined by DSF, where fluorescence intensity of SYPRO Orange was monitored at different temperature (20 to 95 °C). **A)** Normalized fluorescence intensity of SYPRO orange as obtained through DSF assays on different PBP-A variants. **B)** Comparison of T_m values for the different variants of PBP-A proteins. Dashed line highlights the growth temperature of *E. coli* cultures (37 °C). T_m values of the unfolding transition were calculated by the maximum of the first derivative of the fluorescence intensity signal. The T_m values represents the mean of three replicates and error bars show the SD between the replicates.

Results of stability assay using DSF show that the disulfide does not improve the stability of proteins. Melting temperature (Tm) of PBP-A-wt-SS was estimated as ~ 56.5 °C, which is 8.5 degree lower that Tm of PBP-A-wt (Tm = ~ 65 °C).

While Tm of PBP-A-L158E-SS (Tm = ~ 44.4 °C) show no significant change compare to PBP-A-L158E (Tm = ~ 44.7 °C). These results also show a significant decrease (20 °C difference) in thermal stability of PBP-A-L158E compare to PBP-A-wt. Nevertheless, Tm of all PBP-A variants in this study are significantly higher than growth temperature (37 °C) of *E. coli* cultures, thus causing least no stability-related issues under bacterial growth condition, at least for PBP-A-wt and PBP-A-wt-SS. These results further indicate that the fitness improvement of S-S variants of PBP-A is not related to stability.

4.2.5 Directed evolution of PBP-As in presence of the S-S bond

Since disulfide is improving fitness without impairing penicillin reactivity, it may facilitate the conversion of PBP-A into a β-lactamase. We therefore subjected error-prone libraries of PBP-A-wt-SS (2 $\times 10^6$ individual clones) and PBP-A-L158E (1.8 $\times$ 10^6 individual clones), expressed from low-copy pBAD43 plasmid, to selection against ampicillin using gradient plate as well as ampicillin test strips. Prior to selection, expression of PBP-A was induced. As shown in Figure 4-11, interestingly, some bacteria were able to grow at ampicillin concentrations reaching 16-24 μg/ml, i.e. about 10-times higher than the starting level. This was never observed in previous attempts with PBP-A without disulfide bridge.

Top resistant clones were collected at the border of the halo and re-plated on a medium with arabinose and ampicillin strip. As shown in Figure 4-16, the improved resistance phenotype was not reproduced even though presence of PBP-A in the second round of selection was confirmed by PCR and presence of cysteine mutations were verified by sequencing.

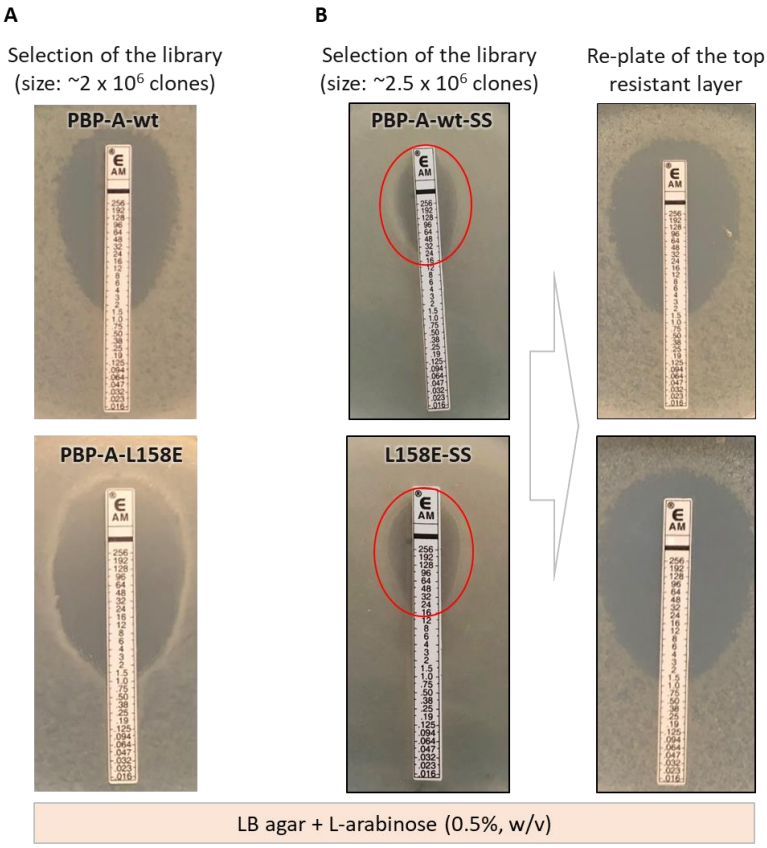


Figure 4-16: Selection of disulfide bonded PBP-A libraries. Compare to libraries of no-disulfide PBP-As **(A)**, an increase in MIC (up to 24 μg/ml) against ampicillin is observed by selection of PBP-As-SS libraries **(B)**, harboring up to ~2 x 10^6 individual clones, using a gradient strip of ampicillin. This resistant (less susceptible) phenotype was not stable upon recovery and re-selection of the top resistant population (indicated by a red circle) close to the inhibitory zone. Note that same amount of cells spread on each plate after normalizing the OD_{600} of each culture to a similar value. Experiment on re-plating of top clones was repeated 2 times and the same results were obtained. Selection and re-plating were carried out on inductive agar plates (containing 0.5% L-arabinose).

These results indicate that the disulfide bridge seems to facilitate evolutionary emergence of β-lactamase activity but may not be sufficient to reach a stable phenotype.

4.3 Discussion

Results in this chapter emphasizes on fitness-improving role of introducing Cys77-Cys123 disulfide bond of class A β-lactamases into PBP-A. The disulfide

bridge reduces the DD-peptidase activity of PBP-A without compromising the reactivity against penicillin. Additionally, the S-S bond seems to be beneficial for PBP-A evoltution into β-lactamase, as bacteria were selected with higher M.I.C of ampicillin for the first time. However, this higher M.I.C was not stable and somehow disappeared upon re-growing in absence of ampicillin followed by second round of selection.

In class A β-lactamases, there is a conserved S-S bond either (i) between Cys77-Cys123, as in TEM-type β-lactamases, or (ii) between Cys68-Cys238, as in SME-type carbapenemases, or there is no S-S bond in the structure such as in Toho-1 and CumA (Figure 4-1 and Appendix 9-24) (Philippon et al., 2016). Single cysteines are also found to be conserved in position 69, 135, 123 and 77 (Figure 4-2). Interestingly, we found that among DD-carboxypeptidases of E. coli, PBP5, PBP6 and PBP6b contain a cysteine residue at position 135 (ambler numbering). Since in the oxidative environment of periplasm single cysteine residues are vulnerable to oxidation and/or undesired misfolded disulfides, carrying such residue would be costly unless protection mechanisms are involved to rescue the single cysteine (Depuydt et al., 2009). Since the single cysteines are mainly located in the active site pocket of these enzymes, such protection mechanisms may contribute in controlling the (costly) activity(ies) of the enzymes.The emergence of the S-S bonds during evolution of β-lactamases has not been investigated and it is not clear whether this emergence happened before gaining hydrolytic power against β-lactams, in ancestral PBPs, or after becoming β-lactamases. Here, we investigate effects of introduction of the 77-123 disulfide bond into PBP-A, which is the closest known LMW PBP to the class A BLs, hence mimicking evolutionary emergence of this S-S bond in this ancestral-type PBP.

Initially we observed that introducing the 77-123 S-S bonds of class A BLs into PBP-A proteins (PBP-A-wt and PBP-A-L158E) has significantly reduced the fitness cost associated to their expression in *E. coli*. This fitness advantage was detectable as significant recovery in growth defect of *E. coli* cells (Figure 4-7), as well as absolute outgrowing of *E. coli* cells expressing proteins with the S-S bond in competition experiments against cells expressing PBP-As without S-S bond

(Figure 4-8). Interestingly, the disulfide bridge in TEM enzymes is not important for activity (Laminet & Plückthun, 1989; Schultz et al., 1987) but our results indicate that its removal generated a slight fitness cost in *E. coli,* suggesting that its role may be to control promiscuity. Sparse observations in the literature point in the same direction. For example, the expression of class A BLs, such as CumA family, that has no S-S bond is tightly regulated in Gram-negative *Proteus vulgaris* and only inducible in presence of the antibiotic substrates (Datz et al., 1994). Such regulation may protect the bacteria from deleterious side activity. More interestingly, researchers failed at expressing a disulfide-lacking class A BL, called bla_{ARL}, in *E. coli* and cloning assays resulted in inactivation by nonsense mutations in *blaARL* gene in all of the constructed plasmids, again suggesting a deleterious effect (Andreis, Perreten, & Schwendener, 2017). On the other hand, many disulfide containing β-lactamases are constitutively expressed with no sign of deleterious impact in the absence of β-lactams. In this context, it is worth noting that extremely slow peptidase activity has been reported for TEM-1 (Rhazi, Galleni, Page, & Frère, 1999). Therefore, we hypothesized that evolutionary acquisition/conservation of S-S bond in the similar structural fold among members of class A BLs, might have a security role against interfering activities and/or their substrate spectrum rather than only on stability of these proteins.

We investigated this hypothesis by comparing the activity and stability of disulfide-bonded PBP-A with their disulfide-lacking counterparts. Since we know (from chapter 3) that PBP-A is effective on PG of the host *E. coli* cells, we compared the PG profile of *E. coli* cells expressing the disulfide-bonded PBP-As with those lacking the S-S bond. We also compared their *in vitro* activities against S2d and peptidic substrates. Both *in vivo* and *in vitro* DD-peptidase activities of PBP-A was constrained upon introduction of the S-S bond. The lower k_{cat}/K_M of the disulfide-containing PBP-As Table **4-1** could probably originate from a rigidified active site while limiting the dynamics that may be necessary to perform the hydrolysis. Additionally, activity on penicillins is barely affected, indicating that the active site is still well structured. These results support our hypothesis that the S-S bond has an adaptive role by securing the interfering DD-peptidase

activity(ies) of PBP-A. Moreover, looking at a possible evolutionary trajectory of PBPs to class A BLs, our results also indicate that acquisition of the S-S bond would not compromise the capacity of PBPs to bind the β-lactam antibiotics such as penicillin.

Interestingly, removing the Cys69-Cys238 S-S bond in SME-1, NMCA and SFC-1 carbapenemases has been reported to cause arrest of bacterial growth upon expression, hindering further study on these enzymes (C. A. Smith et al., 2016). However, despite the postulation that this growth arrest is cause by severe compromise in stability, the exact cause behind the growth is not reported. Smith et al., (2016) has shown that the 69-238 disulfide bond plays a direct effect on the conferred level of resistance of GES-type class A carbapenemases. They substituted the Cys69 in these enzymes into glycine, alanine, methionine and leucine, and shown that only C69G substitution partially generates the resistance phenotype against β-lactam, while other substitutions did not confer any resistance. This suggest that the S-S bond is not essential for the activity against β-lactams. Analysis of the steady-state kinetics of the C69G mutants of GES-5 enzyme showed that the catalytic efficiency (k_{cat}/K_M) of the mutant is very similar to the disulfide-bonded parental enzyme in assays with ampicillin, penicillin G, cephalothin and imipenem, while as mentioned the antibiotic resistance had a drop. They further determined the crystal structure and showed that the disulfide-bonded parental GES-5 has essentially a similar structure to the disulfide-deficient C69G variant (rmsd = 0.14 Å for 265 matching C_α atoms). Since the Cys69-Cys238 bridge is acting as a third hinge between two domains of class A β-lactamases, they suggested that the essentiality of 69-238 S-S bond for hydrolytic power of class A carbapenemases depends on intensity of the hydrogen-bonding network between the two structural domains (all-α and αβ domains) of the enzyme to buffer the destabilizing effect on conformation of the active site (i.e., if there are more hydrogen bonding interaction between domains, the effect of removing disulfide would be less). Although this research concluded that the disulfide bond in GES-type enzyme is important for stability and structural integrity, based on their observation of growth arrest upon expression

of bridge-deficient carbapenemase enzymes as well as our observation on fitness-improving role of S-S bond in PBP-A, we suggest that there may be a role in managing promiscuous activities that can prove problematic against the host bacteria.

On the other hand, carbapenemase-resistance is mostly conferred by class B or class D β-lactamases, and a few class A β-lactamases are capable of hydrolyzing carbapenems (such as SME-1, KPC-1 and NMC-A that are known as class A carbapenemases), which contain the 69-238 disulfide bond in their active site, suggesting that narrowing conformational versatility of the active site in these enzyme is essential for the hydrolytic power against carbapenems. Although the impact of S-S bond of these β-lactamases on cellular fitness and evolvability of these enzymes is not reported, a similar role in securing the active site against deleterious activities may be hypothesized.

Several studies highlighted a role of internal S-S bond in regulation of conformational dynamics of proteins (Astudillo et al., 2010; Silvers et al., 2012; Haoming Zhang et al., 2009). Despite the prevailing view that disulfide bonds have emerged during evolution to enhance the stability of proteins, more evidences are showing that internal disulfide bonds can act as switches for protein function (Butera et al., 2018; Hogg, 2003; Yiming Liu, Carlsson Möller, Petersen, Söderberg, & Hederstedt, 2010; Messens & Collet, 2013). Such role of S-S bond as functional switch is well described in a HMW-PBPs of *Bacillus subtilis*: two cysteine residues are conserved among some class B HMW-PBPs, known as SpoVD, in endospore-forming bacteria, such as *Bacillus subtilis* and *Clostridium acetobutylicum*. Hederstedt and colleagues have shown that a disulfide bond is formed in SpoVD of *B. subtilis*, which is inactivating the enzyme and, thus, activity of the enzyme is dependent on cleavage of the S-S bond by a disulfide reductase, called StoA (Bukowska-Faniband & Hederstedt, 2017; Yiming Liu et al., 2010). SpoVD is a sporulation-specific transpeptidase homologous to PBP3 of *E. coli* (Daniel, Drake, Buchanan, Scholle, & Errington, 1994). It is suggested that SpoVD remains 'off' during the growth of bacteria and turns 'on' when spore peptidoglycan synthesis for sporulation is required (Figure 4-17)

(Bukowska-Faniband & Hederstedt, 2017; Eichenberger, 2010). However, the role of S-S bond in security of SpoVD enzymes, to control their unwanted interference/activities in cellular reactions is not reported. It is possible that early activity of this enzyme (prior to sporulation) impair the cellular fitness, and therefore, the red-ox switch via the S-S bond has been selected during evolution to maintain the fit state during growth of the bacteria.

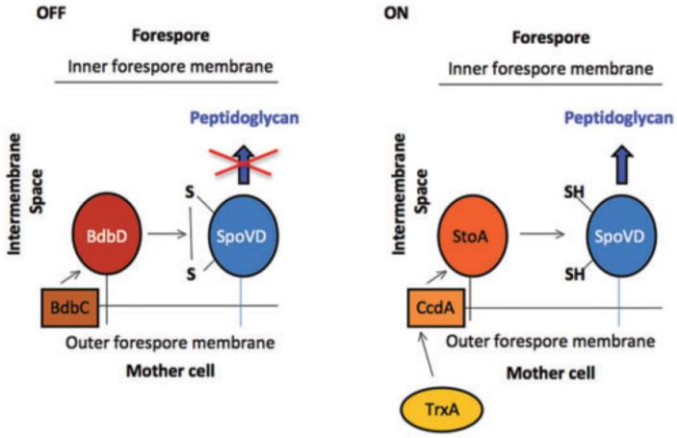


Figure 4-17: Disulfide bond as a switch for activation of, SpoVD, a HMW-PBP in *B. subtilis*. The synthesis of spore peptidoglycan, and thus, sporulation is dependent on the redox state of the SpoVD. **OFF)** upon secretion of SpoVD into the intermembrane space, a disulfide bond is formed in the active site of in the enzyme, switching 'off' its activity. S-S bond oxidation in SpoVD is catalyzed by the a thiol-disulphide oxidoreductase, called BdbD. Oxidation of BdbD is maintained by the membrane protein BdbC. **ON)** The S-S bond in SpoVD is reduced by production of StoA thioredoxin-like protein, which lead to activation of SpoVD, and thus, switching on spore peptidoglycan biosynthesis. The electrons flow from TrxA in the mother cell cytoplasm, reducing CcdA in the outer forespore membrane and then StoA in the intermembrane space.

In our directed evolution experiments, unlike in most of the PBP-A libraries, initial selection of PBP-As-SS libraries shown an increase in MIC against ampicillin (up to 24 μg/ml), but this phenotype was not stable upon reselection of the bacteria growing in the resistance zone (i.e., resistant clones were not viable, Figure 4-16). Our main hypothesis to explain this result is that there are a few clones in our libraries that are expressing PBP-A variants endowed with β-lactamase activities but the increased hydrolytic activity may also be associated with increasing toxicity. Also, it is reasonable to imagine that ampicillin itself is

protecting against such toxicity by competing with PG and the resistant bacteria grow until they die upon ampicillin depletion leaving the place to non-resistant clones. More work will be necessary to test this hypothesis. However, despite the significant improvement in fitness of PBP-A, introduction of the disulfide bond did not significantly improve evolvability, which make sense as the disulfide bridge reduces the conformational plasticity. Therefore, the flexibility of the active core is restricted and will provide less access to the conformational space, and thus, reducing evolvability. Presumably the acquisition of S-S bonds in evolution of β-lactamases has happened after gaining the β-lactamase activity to secure the hydrolytic active site, and not during PBP to not compromise the required plasticity for evolvability towards β-lactamases. Nonetheless, after gaining the S-S bond, a robust and secure scaffold is provided for the β-lactamases for both, coping with interfering promiscuities as well as increasing robustness against destabilizing mutations (Schultz et al., 1987). The latter could explain the extension of substrate spectrum in TEM-type β-lactamases, without compromising the general structural fold. In a general sense, we could imagine that once a secondary function in a promiscuous enzyme becomes beneficial for the cell, acquisition of catalytic-fixing disulfide bonds can act as a switch by selective pressure to evolve minor (plastic) conformers into major (rigid) conformers. Protein switches are currently attracting more attentions in designing novel proteins as they are known to play role as regulatory mediators as well as primers of energy conversion and mechanical dynamics in molecular machines (Alberstein, Guo, & Kortemme, 2022).

4.4 Conclusions

We have shown that introducing a S-S bond between α_2 and α_5 of PBP-A protein, improves its cellular fitness by minimizing the interfering activity(ies) of this enzyme on PG of *E. coli* cells.

Our results also confirmed that introducing the S-S bond in PBP-A enzymes has a confining effect on peptidase activity limiting the hydrolytic efficiency without compromising its reactivity towards penicillin. These are most likely due to a

constraining effect of the S-S bond on dynamics of the catalytic core. However, molecular dynamic (MD) studies on these mutants remain to be carried out, which would help us to understand more about the effect of S-S bond on dynamics of the PBP-A active site.

In the evolutionary trajectory between PBPs and serine β-lactamases, the introduction of a disulfide bridge may have played a very important role in protecting the bacteria against unleashed activity against the peptidoglycan while keeping the reactivity towards β-lactams, a role of positive intragenic epistasis. Overall, this work is shedding light on an overlooked role of disulfide bonds in proteins which is to act as securing elements against deleterious promiscuity.

4.5 Contribution of authors:

H.D kindly carried out the MS analysis of PBP-A assays with synthetic peptides. D.E helped in enzyme purifications. W.V generously collaborated on the assays of PBP-A on peptidoglycan and muropeptides as well as HPLC and MS analysis and discussions on results of these assays. The rest of the works including development of theory, preparation of mutants of PBP-A, assays, discussions on data and preparation of manuscript are carried out by G.M.D and supervised by P.S.

4.6 Acknowledgements

We are thankful to Prof. André Matagne and his student Romain Malempré, University of Liège, Belgium, for helping us with the Differential Scanning Fluorimetry assay.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722610 (ES-Cat).

4.7 Material and methods

4.7.1 Site-directed mutagenesis

To construct L68C/A115C mutant of *pbp-a-wt and pbp-a-L158E genes*, which were already cloned in low-copy pBAD43 plasmid, (3.7.1.2) two PCR reactions were performed. First to introduce L68C and A115C mutations by amplification of *pbp-A* genes from 68 to 115 positions (178 bp) using primers that contained

mutations (primer pair 1, cysteine mutations are underlined)). Second to amplify remaining plasmid_*pbp-A*_cys backbone with primers complementary the first primer pair (primer pair 2, cysteine mutations are underlined and overlapping complementary ends between two pairs of primers are indicated as bold or italic.

Primer pair 1:

*pbp-A*_L68C fw:

ACCATTAAATTCCCGATCTGCGTGGCGTTTTTCAAAGCCG

*pbp-A*_A115C rv:

*TGGTAATCATCAGTTCGCAAACTTCCAGAGCTGCGT*ATTG

Primer pair 2:

PBAD43_*pbp-A*_cys fw:

*ACGCAGCTCTGGAAGTTTGCGAACTGATGATTACCA*TCTC

PBAD43_*pbp-A*_cys rv:

GCTTTGAAAAACGCCACGCAGATCGGGAATTTAATGGTAG

The two PCR products were purified and subjected to Gibson assembly by adding 50-100 ng of vector backbone with a 2-3 folds molar excess of the insert into 15 µL of Gibson Master Mix (100 mM Tris–HCl pH 7.5, 10 mM MgCl2, 0.2 mM of each four dNTPs, 1 mM NAD+, 15% (w/v) PEG-8000, T5 exonuclease (2 U/mL), Phusion DNA polymerase (33 U/mL), and Taq DNA ligase (1666 U/mL)). To digest the original (methylated) plasmid template, 0.5 µL of DpnI (20 U/ml) was added to the mix followed by addition of ddH2O up to 20 µL. The assembly mix is then incubated for 1 h at 50 °C. One µL of the assembly mix is electroporated into 50 µL of *E. coli* Top10 electrocompetent cells (Invitrogen). Transformants are incubated at 37 °C, for 1 hour before plating all the cells on LB plates containing spectinomycin and glucose (for repression of expression) followed by overnight incubation 37 °C.

4.7.2 Disulfide bond detection by protein's migration

Native (non-reducing and non-denaturing) sample buffer (50 mM Tris base, 10% glycerol, H2O, at pH 6.8), Periplasmic extracts of each mutant is mixed with 5x native sample buffer in 5:1 ratio and loaded into SDS-PAGE gels. The migrated

proteins on gel were subjected to western blotting for detection and comparison of bands corresponding to PBP-A-wt proteins and PBP-A-SS.

4.7.3 Estimation of expression level using band intensity quantification

To compare expression level of SS-bridged and non-bridged PBP-As, the intensity of their corresponding bands from western blotting images were analyzed by ImageJ image processing software, which is publicly available online: https://imagej.net/software/imagej/ (Schneider, Rasband, & Eliceiri, 2012). Western blot images are obtained from analyzing periplasmic extracts of 10 ml PBP-A expressing cultures, which were normalized to the same OD_{600} prior to periplasmic extraction. Intensities estimated from ImageJ were normalized by dividing intensities of the bands to the intensity of background (non-blotted white area of the image).

4.7.4 Fitness competition assay

From a single colony on a recently inoculated plate, pre-cultures for each strain are started into 5 ml of LB media containing appropriate antibiotics in a 50 ml falcon tube and incubated for overnight at 37°C under agitation of 180 rpm. A fresh culture is started for each strain under repression conditions (0.5% glucose) and growth to reach exponential phase (OD_{600} ~ 0.5-0.6). To start each competition, the cultures for both competitors are normalized for the same OD_{600} and mixed at a ratio of 1:1 and 100 µl of the mix was inoculated into 100 ml LB medium containing appropriate antibiotic and inducer or repressor. The culture was incubated at 37 °C under 180 rpm agitation and after 24 hours, 100 µl of the cultures we inoculated into fresh media with similar conditions. Competition assay was conducted for 4 days by repeating the inoculation after every 24 hours. From the remaining volume of cultures on each day, plasmid was extracted and used as template for PCR to amplify the region containing the gene of interest, followed by sequencing. Winner of the each competitions are determined by analyzing sequencing data.

4.7.5 Enzymatic assay of PBP-A on S2d substrate

In a flat bottom and clear UV-Star® half area 96 well microplate (Greiner Bio-One), 5 µL of enzymes is mixed with 75 µL of different concentrations (0 - 1 Mm) of S2d thiolester substrates that were prepared in 20 mM HEPES buffer, pH 7.4. Hydrolysis of S2d was performed at 25 °C and monitored spectrophotometrically at 250 nm, using a microplate reader (Tecan Spark). Kinetic parameters were calculated by fitting experimental values to numerically simulated values of the progression curves using Michaelis–Menten equation:

$$v = \frac{V_{max}[S]}{K_M + [S]} \qquad 4.1$$

where, v = velocity (rate) of reaction, V_{max} = maximum rate achieved, [S] = concentration of substrate K_M = Michaelis constant.

4.7.6 Penicillin Binding Assay and Rate Measurements

To compare the acylation/deacylation of different PBP-A with penicillin, we used, BOCILLIN FL, a fluorescent penicillin, to measure the rates of formation (k_{ac}, acylation) and decay (k_3, deacylation) of the $E - S$ complex.

Deacylation rate: k_3, was derived from the first order decay of BOCILLIN-PBP complex. For that, PBP-A protein (65 nM) was incubated for 2 min at room temperature with 10 µM BOCILLIN FL (Invitrogen) in 20 mM HEPES, 140 mM NaCl, pH 7.5. Then, 1 µl of TEM-1 β-lactamase (at a concentration to hydrolyze the remaining free BOCILLIN in a few seconds) was added, immediately and every 10 min aliquots of 15 µl were removed and mixed into 10 µl denaturing sample buffer for SDS-PAGE (50 mM Tris base, 10% glycerol, 0.01% bromophenol blue, 2% mercaptoethanol, 2% SDS and H2O at pH 6.8) and immediately frozen at -80 °C to stop the deacylation reaction. The reaction mixtures then boiled at 98 °C for 5 minutes and subjected to SDS-PAGE followed by fluorescence imaging using Amersham Typhoon (GE Healthcate). The intensity of fluorescent bands was measured using ImageJ software (Schneider et al., 2012) and k_3 constants were determined by fitting an exponential curve to the experimental values.

Acylation rate: 100 nM PBP-A was mixed with 65 nM BOCILLIN in 20 mM HEPES, 140 mM NaCl, pH 7.5. Every 10 seconds, aliquots of 15 µl were removed and mixed into 10 µl denaturing sample buffer and continued similarly as explained in deacylation protocol. Since in acylation of PBPs to β-lactams, the $(k_{-1} + k_2)/k_1$ (or K_S) is most probably much higher than the experimental concentrations of enzyme and substrate, the acylation follows a second order reaction with a rate constant $k_{ac} = k_1 k_2/(k_{-1} + k_2)$. Using k_3 from previous deacylation experiment, the optimized k_{ac} constants were determined by fitting numerically simulated values using the differential equation 4.2 to the experimental data.

$$d[E - Ac] = (k_{ac}[E][S] - k_3[E - Ac])dt \qquad 4.2$$

4.7.7 Differential scanning fluorimetry (DSF)

Differential scanning fluorimetry (DSF) was used as described by (Kellner et al., 2021). SYPRO Orange (Invitrogen Inc.) was used as the dye in DSF assay to monitor the denaturation transition of PBP-A proteins by measuring fluorescence intensity using a real-time quantitative PCR machine (Biorad MJ Mini Thermal Cycler equipped with Mini Opticon real-time PCR System). In 48-well PCR plates (white, with optical lids) and a final volume of 30 µL, 28 µL of PBP-A protein samples at 0.2-0.4 mg/ml (6-12 µM) concentrations, were incubated with the SYPRO orange dye (1000x diluted from the commercial solution) and incubated at 20 °C for 5 min. Fluorescence intensity was measured under gradient temperature from 20 to 95 °C under 0.5 °C steps and 20 s incubation at each step. Thermal unfolding curves were obtained T_m values of the unfolding transition were calculated by the maximum of the first derivative of the fluorescence intensity signal.

4.7.8 Selection of PBP-As-SS libraries

Prior to selection, 200 µl of the prepared and pooled libraries of each variant were inoculated into 20 ml of LB media and inducted by L-arabinose (0.5%, w/v, at OD_{600} 0.3-0.4) for PBP-A expression, followed by incubation for 1.5-2 hours at 37 °C under agitation. Expressed libraries were inoculated on LB agar plates

containing L-arabinose and a MIC Test Strip of ampicillin (bioMérieux) was applied on top followed by incubating overnight at 37 °C. To confirm the MIC of top variants in the library, cells that were growing at higher concentration of ampicillin (close to inhibition zone) were pooled, grown, induced for PBP-A expression and re-selected as described.

Chapter 5

Directed evolution and genetic drift of PBP-A gene libraries in plasmid and chromosomal contexts

Gol Mohammad Dorrazehi, Steve Alarcia and Patrice Soumillion

Abstract

Directed evolution of recombinant proteins in microbial system requires careful attentions towards minimizing potential fitness costs associated with protein expression. An ideal library would have the best compromise between fitness burden and emergence of new phenotypes in the protein under study. In chapter 3 and 4 we studied the fitness cost associated with the activity of a cyanobacterial DD-peptidase (PBP-A) expressed in the periplasm of *E. coli* and demonstrated improving impact of (i) lowering gene copy-number and isoeletric point of the enzyme, (ii) fixing the pH of growth media above 7 and (iii) introducing a disulfide bond in the protein. Here, we attempt the conversion of PBP-A into a beta-lactamase under fitness-improving conditions. After developing a novel efficient method for the construction of single copy and scarless gene libraries in chromosome of *E. coli*, we performed directed evolution experiments with PBP-A under both single- and low-copy genetic contexts. Although directed evolution

was not successful, in order to evaluate mutational landscape, we further subject the libraries of PBP-A variants to prolonged serial passage under no selection pressure, referred as 'neutral drift'. Along the neutral drift experiment, we monitored the expression level of PBP-A proteins by western blot as well as the fate of individual clones by next generation sequencing (NGS) of libraries before and after drift. The expression level remained constant during drifts performed at the lower expression levels while it was gradually lost during the drift under induction of PBP-A in plasmid library. NGS data analysis did not reveal any fitness-improving mutations other than insertion and deletions (InDels) in the case of higher expression conditions. Finally, we subjected the plasmid-borne libraries to successive selection on plates containing ampicillin gradient and were able to gradually increase resistance against ampicillin in bacteria up to 1500 mg/L. However, the ampicillin-resistant phenotype was not due to PBP-A but rather to mutations leading to the increase in expression level of the chromosomally encoded AmpC beta-lactamase. Interestingly, the presence of *pbpA* gene appears necessary for the phenotype emergence as well as two other chromosomal mutations in *pbp3* and *mar*R genes.

5.1 Introduction

Directed evolution is by now a foundational technology in enzyme and protein research, used to improve a wide array of characteristics such as stability, activity and selectivity. Directed evolution aims to sidestep the problem of predicting the relationship between the primary sequence of a peptide or protein and its function by harnessing the power of evolution through sequential rounds of diversity generation and artificial selection or screening (Romero & Arnold, 2009a) (see also section 1.2.3 in Introduction). Genetic libraries of gene variants, generally cloned in plasmids, are used as the input directed evolution campaigns. Working with plasmid libraries has several advantages, notably the ease of DNA cloning, recovery and sequencing as well as the high expression level of protein, thus facilitating downward purification and characterization. These advantages are even stronger with plasmids featuring high copy number that can reach several hundred molecules per cell. However, it also comes with some potential

disadvantages. Maintaining plasmids requires the use of extra antibiotics in the culture medium. In addition, the plasmid carrying the library might be incompatible with other plasmids eventually required for screening or selection. On the other hand, at the protein level, although high expression may be advantageous from an experimental point of view, it is also known that highly expressed proteins evolve slowly due to associated fitness costs such as selection against misfolding (Cherry, 2010; Drummond et al., 2005; Serohijos, Rimas, & Shakhnovich, 2012b). Because of the variation in copy number, the tight and tunable control of gene expression level is difficult to achieve with plasmids. It has been shown that having finely regulated expression system is important, which may otherwise skew the results of experiments, e.g. when doing genotype-phenotype mapping studies (Rockah-Shmuel, Tóth-Petróczy, & Tawfik, 2015). Additionally, leaky promoter on high-copy plasmids can also be an impediment for the expression and laboratory evolution of cytotoxic proteins, cytotoxicity being obviously a matter of dose. Finally, for evolution campaigns where the goal is to evolve an endogeneous enzyme to simulate a natural evolutionary trajectory, cloning a library in a natural genomic environment and, if possible, without the need of antibiotic selection, is a better choice to mimic the native expression and regulations conditions. There is therefore an increasing interest in methods that can generate chromosomal gene libraries to provide tighter control and repression, easier maintenance of the gene, greater compatibility with plasmid-borne system and better mimicking of the natural context, even at the cost of a more intricate process of DNA manipulation (Rockah-Shmuel et al., 2015).

Several recently published methods allow the diversification of genomic loci. The pEvolvR system recently published by Halperin et al. couples an error-prone PolI polymerase mutant with a nickase variant of Cas9 to diversify a specific region of the genome in a continuous manner (Halperin et al., 2018). However, the target region is quite small (a short window up to 350 bp) and the method does not allow for the insertion of designed libraries. These methods also suffer from having a mutation rate that depends on the distance from the cutting site. The Multiplexed Automated Genomic Editing method invented by George Church

allows efficient insertion and screening of libraries at different loci, but requires advanced and expensive equipment to overcome the low efficiency of recombination through continuous transformation with a library of synthetic DNA fragments (H. H. Wang et al., 2009). In this chapter, first we present an efficient method for the creation of scarless chromosomal libraries in *E. coli*, which relies on the combination of CRISPR-Cas9 with the λ-Red recombineering systems (Jiang et al., 2015). Then we apply this system in directed evolution of PBP-A.

Directed evolution of PBP-A, which is phylogenetically the closest family to class A β-lactamases, was previously attempted in our team in order to generate β-lactamase resistance in *E. coli*. Despite of a very similar structural fold, all the previous attempts based on rational design and random mutagenesis strategies have failed to generate an active enzyme to confer resistance against penicillin. Evidences suggest that converting a DD-peptidase into a β-lactamase follows a counter-intuitive evolutionary trajectory and fitness cost(s) related to expression of PBP-A in *E. coli* were suggested as the main obstacle. As shown in Chapter 3, a high fitness cost is associated with expression of PBP-A in *E. coli*, originating from the DD-peptidase activity of PBP-A on the host's peptidoglycan. We observed a correlation between fitness cost and expression level and reported a minimum fitness cost under single-copy expression of PBP-A.

Despite the similarities in structural scaffolds and conserved active site motifs between PBP-As and class A β-lactamases, an important difference is the presence of a six-residues shorter Ω-loop flanking the active site of PBP-A Figure 5-1. An essential catalytic base (Glu166) is carried by this loop in β-lactamases, which is responsible for water activation in the deacylation of the penicilloyl-enzyme. In PBP-A, the structurally aligned position is occupied by a leucine (Leu158). Substitution of this leucine into a glutamate results in a 90-fold increase of the β-lactamase activity with a k_{cat} of 10^{-2} s^{-1}, which is still about 10^5-fold less active than a natural β-lactamase and this variant is unable of conferring any penicillin resistance to *E. coli* (Labarbe, 2006; Urbach et al., 2008).

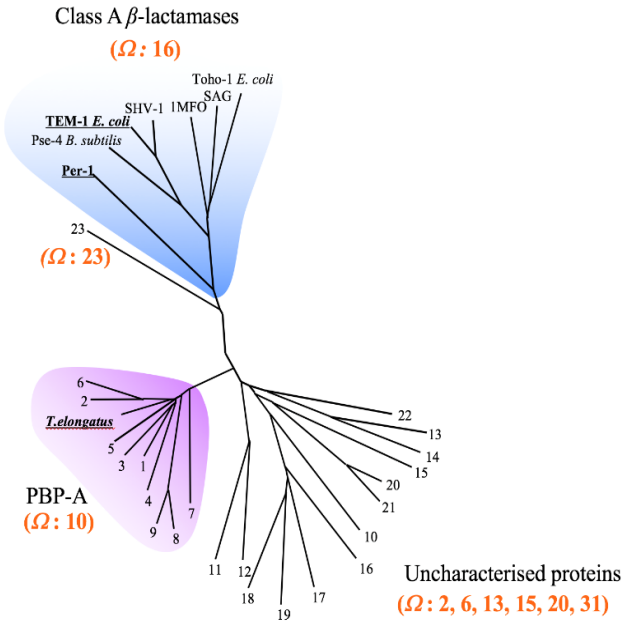


Figure 5-1: Variable lengths of Ω-loop between PBP-A and class A β-lactamases. The Ω-loop of PBP-A family is six amino acids shorter than those of the class A β-lactamases.

An ancestral duplication of a heptapeptide may be at the origin of the omega loop in beta-lactamases (Joachim, 2016). Results (unpublished) suggested that while the longer length of Ω-loop in TEM-1 is not essential for hydrolysis of penicillin, its shortening generates a fitness cost to *E. coli* host at low temperature under low expression conditions.

In this chapter, we carry out directed evolution experiments of PBP-A-wt and the L158E mutant under the fitness-improving conditions that were previously identified, including generating libraries under single and low gene copy-number, selection at higher pH of growth media, introducing the disulfide bond or inserting an Ω-loop from a beta-lactamase. We also subjected PBP-A libraries to drifting without any imposed selective pressure and used NGS to evaluate the fate of mutations under various expression conditions.

5.2 Results

Here we first describe our methods to randomize the genes encoding different PBP-A by error-prone PCR (epPCR), as well as expression of ep-libraries from single- and low-copy contexts from chromosome and plasmid, respectively, in *E. coli*. Then we present data related to all attempts of directed evolution towards beta-lactamase followed by the drift experiments.

We subject error-prone libraries (ep-libraries) of *pbp-A* genes, under low- and single-copy number contexts, to selection against β-lactams (mainly ampicillin). For single-copy, we take advantage of above-mentioned method of chromosomal library construction, which allow us for generation of libraries up to million independent clones. However, the *in vivo* selection was not successful to isolate any PBP-A variant hydrolyzing ampicillin.

Additionally, for toxicity assessment of the generated diversity, we subject the ep-libraries to neutral serial passage under no selection pressure (neutral drift) to perform fitness competition between variants. In neutral drift experiment, we evaluate the fate of periplasmic protein expression along the drift by western-blot, followed by evaluation of the fate of individual variants by next generation sequencing (NGS) analysis and comparison of their redundancy in initial and final populations. This would allow us to not only assess the diversity and compositions of the starting libraries, but also estimate the role of variants in the fitness of protein in *E. coli*. NGS data analysis show that fitness competition between variants is highly relevant to expression level as majority of PBP-A variants remained neutral during non-selective bulk competition (neutral drift) in chromosomal context, while in drift under plasmid context the competition is stronger that is leading to fixation or purge of some variants for basal (leaky) expression and surprisingly complete loss of PBP-A proteins for induced expression as a result of selection of insertion-deletions (InDels) and nonsense mutations in the gene.

Finally, we subject the plasmid-borne libraries to non-neutral drifting under selection for sublethal concentration of ampicillin to benefit from the stress response activated by the antibiotic treatment (Kohanski, Dwyer, Hayete,

Lawrence, & Collins, 2007). In the non-neutral drift experiment, we increased the resistance against ampicillin in bacteria by drifting the antibiotic-stressed population over time, and because the resistant phenotype was not linked to plasmid, we carried out whole genome sequencing (WGS) and analyzed the genotype of the resistant clones.

5.1.1 Directed evolution of PBP-A to β-lactamase

As the main objective of this thesis, and taking benefit from results of previous chapters, directed evolution of PBP-A to β-lactamase was carried out under conditions that improve organismal fitness associated with PBP-A expression, including lower expression level under single- and low-copy genetic constructs, growth at basic pH and introduction of a disulfide bond in PBP-A.

We also used multiple variants of PBP-A as the starting point of the evolutionary experiments including the *pbp-a-wt* and its mutant *pbp-a-L158E,* which endowed with increased hydrolytic activity on penicillin (Urbach et al., 2008) as well as, disulfide-bonded *pbp-a-wt-SS* and *pbp-a-L158E-SS* (see chapter 4), and two chimeric versions named *pbp-a_TEM-1-Ω-loop* and *pbp-a_TEM-1-Ω-loop-KTG_fit* that contain Ω-loop of TEM-1 β-lactamase and a modified active site floor in addition to the Ω-loop, respectively. To avoid possible steric clashes between the longer loop of TEM-1 in PBP-A, the floor of active site (adjacent to KTG motif) in *pbp-A* was modified in 6 residues to make *pbp-a_TEM-1-Ω-loop-KTG_fit*. We prepared single- and low-copy number libraries of these four variants in *E. coli*, expressing the variants from chromosome and pBAD43 plasmid, respectively.

Controlled diversity of ep-libraries

For generating libraries in *pbp-a* genes, we used mutazyme II DNA polymerase (GeneMorph II Random Mutagenesis Kit, Agilent Technologies), where the mutation rate can be controlled by adjusting the number of epPCR cycles and the amount of initial template. We decided to build libraries with a low mutation frequency (between 0 and 5 mutations/kb) by applying 100 ng template DNA and 20–25 cycles of epPCR amplification as recommended by user manual. This relatively high input of DNA as template for epPCR (100 ng corresponding to

8.9×10^{10} copies of 1094 bp *rhaB_pbp-a_rhaA* template (see Appendix 9-25) and/or 1.37×10^{10} copies of the 7074 bp *pBAD43_pbp-a* plasmid) would also prevent bottlenecking of the diversity. All the libraries used in this study had a low mutation rate (the mean number of mutations per gene) of 0 to 6 mutations/gene, which was estimated by Sanger sequencing of the full-length *pbp-a* gene from at least 20 random individual clones in each pool of library.

Preparation of pbp-a libraries under single- and low-copy genetic constructs

As previously described (see section 3.2.1), the fitness cost of expressing PBP-A in *E. coli* was correlated with copy number of genetic constructs and a minimum growth defect was observed under single-copy expression from chromosome of *E. coli* (Figure 3-5). Based on this, we developed an efficient and scarless method for construction of genetic libraries in chromosome of *E. coli* (Dorrazehi et al., 2020): *pbp-a* gene was inserted into *rhaA* locus (second gene in rhamnose operon), which was inducible by addition of rhamnose into the culture media (Figure 5-2). The reason that we decided to insert *pbp-a* in place of second gene in rhamnose operon (*rhaA*) was to avoid the presence of a promoter in the donor DNA cassette, which requires relatively long homology regions up- and downstream of the cassette. The absence of promoter in *in vitro* prepared DNA ensures that *pbp-a* genes won't be expressed before their chromosomal integration. The method is based on a combination of CRISPR-Cas9 cleavage and λ-Red recombination technologies, allowing for generation of libraries containing up to million independent clones. Since the expression level was decreased in single-copy context, we also constructed the ep-libraries of *pbp-a* under low-copy pBAD43 plasmid to benefit from the higher level of expression and we achieved libraries of up to $\sim2\times10^{6}$ individual clones.

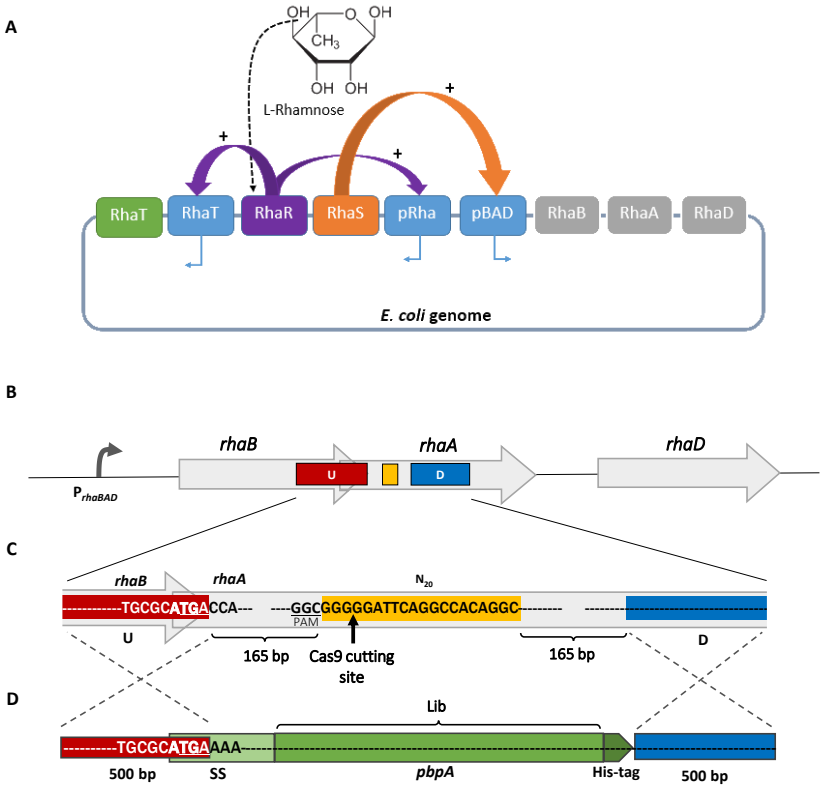


Figure 5-2: Insertion of *pbp-A* library in rhamnose operon of *E. coli* genome. A) **A)** L-rhamnose operon of *E. coli* and its positive regulation by L-rhamnose sugar. Generally, the catabolism of L-rhamnose is inactive in *E. coli* but inducible by availability of L-rhamnose for the bacteria. L-Rhamnose activates the transcription of the genes *rhaR* and *rhaS*, which positively activate the uptake of rhamnose by RhaT transport system, and its metabolization by rhaBAD. **B)** Schematic of *pbp-A* insertion in *rhaA* locus of *E. coli* genome. We decided to introduce *pbp-A* gene exactly at the position of the second gene (*rhaA*). **C)** Nucleotide resolution of N_{20} target of guide-RNA, Cas9 cut site and up and down homology regions (U and D). **D)** Donor DNA cassette of *pbp-A* gene (or library). The gene of interest will replace exactly the start of *rhaA* gene far enough from rhamnose-inducible rhaBAD promoter. The insertion will be facilitated by homologous arms (500 bp) via λ-Red recombination, leaving no modification on rhaB while using the start codon of the *rhaA* gene. The *pbp-A* ORF encodes the N-terminal *DsbA* signal sequence (SS) for periplasmic localization and a C-terminal 6×His-tag for purification. (See section 7.2 for detailed protocol of this genome engineering).

In vivo selection of PBP-A libraries against ampicillin

The clones in each library were scrapped from initial plates and pooled, induced for PBP-A expression by L-rhamnose (0.5%, w/v) or L-arabinose (0.5%, w/v), for chromosomal and plasmid libraries, respectively, and subjected to selection against ampicillin by plating on gradient concentration of ampicillin.

Nonetheless, to avoid competition between mutants prior to selection, selection was also carried out directly on the pool of library, without pre-induction of PBP-A expression. The top resistant clones on the gradient plate were selected for further verification and the pool of top clones were subjected to DNA isolation and to next round of library construction (Figure 5-3).

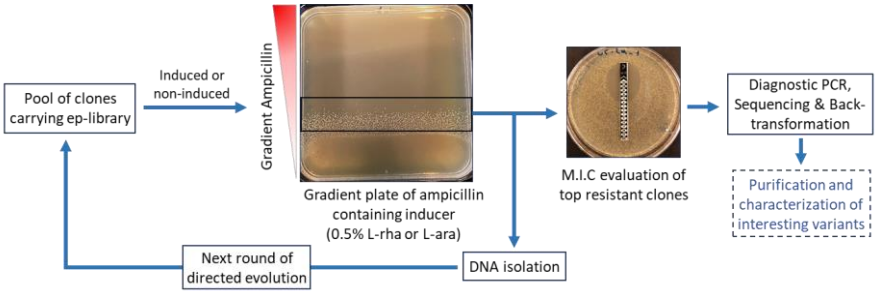


Figure 5-3: Schematic workflow of selection strategy in directed evolution of *pbp-A*. A gradient concentration of ampicillin will provide a competition environment for different variants of *pbp-A*, which allows us to detect the best mutants, conferring highest resistant to ampicillin, in each round of directed evolution. Pre-induced or non-induced libraries were sunjected to selection on the gradient plate containing either L-rhamnose or L-arabinose for induction of chromosomal or plasmid libraries, respectively. While the pool of top clones can be subjected to next round of directed evolution by isolation of the *pbp-A* gene and epPCR, the M.I.C of top mutants is verified using ampicillin test strips, followed by confirming presence of *pbp-A* gene by a diagnostic (colony) PCR using the clones directly as DNA template for the PCR. Mutations on *pbp-A* gene are analyzed by sequencing and desired mutants are selected for further characterization.

Unfortunately, among 22 ep-libraries of PBP-A enzyme, no clones conferring resistance to ampicillin at concentrations higher than 8 μg/ml could be selected. We have also carried out the selection of libraries under fitness-improving environment such as osmoprotective media or media with basic pH and observed no advantage of these conditions on the selection. For the osmoprotective condition, we prepared media with 50 mM $MgCl_2$ to increase osmolarity of the media, and for the basic pH, we buffered the agar media with phosphate buffer (50 mM, pH 8.0). Furthermore, we carried out selection on liquid media containing sublethal concentration of ampicillin (4-16 μg/ml) and selected no resistant clones. Additionally, to evaluate possible evolution against other β-lactams, we carried out the selection using ETEST Strips of different β-lactams other than ampicillin (AM), including benzylpenicillin (PG), cefotaxime

(CT), imipenem (IP), amoxicillin (AC), ceftazidime (TZ). However, compared to the wild-type *E. coli* Top10, no clone with significant increase in the M.I.C against these antibiotics were detected (data not presented).

From selection of each library on gradient plates we isolated some clones with slightly higher M.I.C than the wild-type *E. coli* cells, and the PBP-A gene in each clone was verified by PCR and sequencing. The verified and sequenced clones with higher M.I.C in the first round of directed evolution in chromosomal context are presented in Figure 5-4 and Table 5-2. However, we could not get higher resistant clones upon next rounds of directed evolution on the pool of top clones. Nonetheless, in libraries of disulfide-bonded PBP-As, we observed higher M.I.C of 16-24 μg/ml, but unlike the other libraries, this resistant phenotype was not stable in the clones after re-selecting on ampicillin strip (see section 4.2.4), therefore no further analysis of individual clones was possible in these libraries. In Table 5-1 we present the summary of all libraries prepared in this study and selected against ampicillin.

Table 5-1: Directed evolution attempts on multiple mutants of *pbp-A* gene. Libraries of PBP-A variants, under single- and low-copy genetic constructs, were subjected to selection against gradient ampicillin concentrations followed by next rounds of directed evolution on pool of top resistant clones.

Starting *pbp-a* gene	Rounds of directed evolution/library size (clones)	Copy number /Expression context	Highest M.I.C (μg/ml)
pbp-a-wt	3 rounds/~1×10^6	Single-copy/Chromosome	8
pbp-a-L158E	3 rounds/~0.8×10^6	Single-copy/Chromosome	8
pbp-a_TEM-1-Ω-loop	2 rounds/~0.6×10^6	Single-copy/Chromosome	2
pbp-a_TEM-1-Ω-loop-KTG_fit	2 rounds/~0.5×10^6	Single-copy/Chromosome	2
pbp-a-wt	2 rounds/~2×10^6	Low-copy/pBAD43 Plasmid	6
pbp-a-L158E	2 rounds/~2.1×10^6	Low-copy/pBAD43 Plasmid	6
pbp-a_TEM-1-Ω-loop	2 rounds/~0.8×10^6	Low-copy/pBAD43 Plasmid	4
pbp-a_TEM-1-Ω-loop-KTG_fit	2 rounds/0.8×10^6	Low-copy/pBAD43 Plasmid	4
pbp-a-wt-SS	2 rounds/~2.5×10^6	Low-copy/pBAD43 Plasmid	16-24
pbp-a-L158E-SS	2 rounds/~2.5×10^6	Low-copy/pBAD43 Plasmid	16-24

The mutations of top clones were identified by sequencing after verification of the phenotype by determining the M.I.C on ampicillin strip. In Figure 5-4 we present the M.I.C of top clones in chromosomal libraries and in Table 5-2 the sequence of PBP-A variant in each clone. The variants are named for number of missense mutations in PBP-A sequence (e.g., PBP-A-wt-4M contains 4 missense mutations).

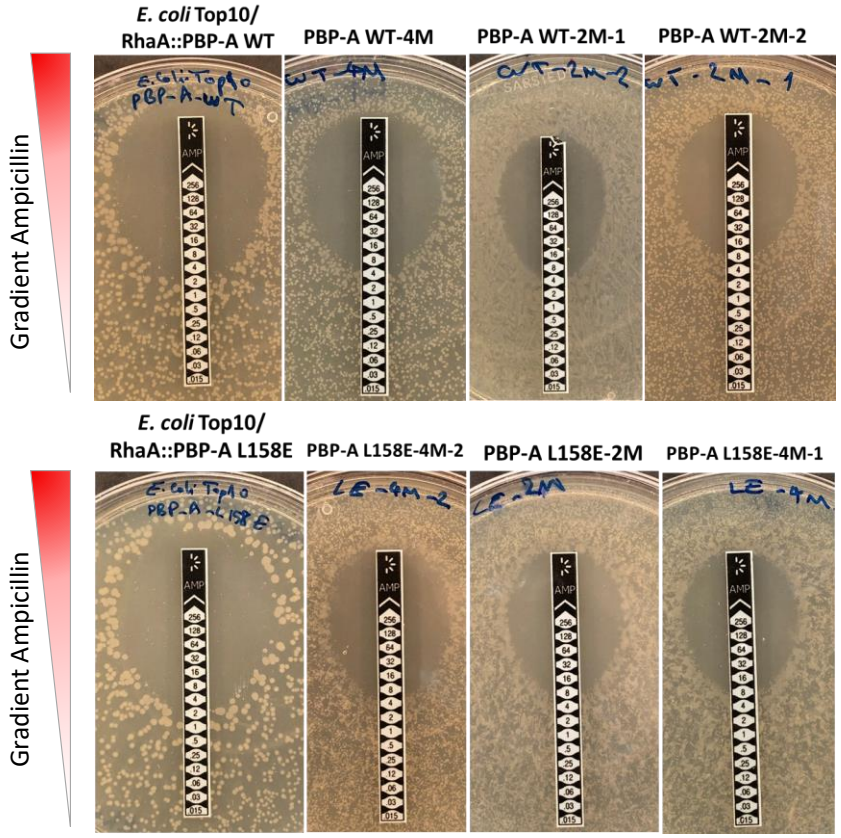


Figure 5-4: Minimum inhibitory concentration (M.I.C) of top resistant clones in chromosomal libraries of PBP-A-wt (top) and PBP-AL158E (bottom). When libraries are subjected to selection on gradient ampicillin (0-32 µg/ml), colonies growing at higher ampicillin concentrations (top colonies) were isolated and re-selected on ampicillin strip to quantify their M.I.C. Sequences of the PBP-A variants in these clones are available in Table **5-2**.

Table 5-2: Selected top variants in first round of directed evolution of PBP-A under single-copy context with the corresponding mutations and redundancy in among 20 sequenced clones. The variants are named based on the number of missense mutations in PBP-A sequence (e.g., PBP-A-wt-4M contains 4 missense mutations).

Library	Clone	Mutations	M.I.C	Redundancy	Selection condition
PBP-A-wt	WT-4M	G53D, V56M, D91G, K212I	~4 μg/ml	× 7	basal & induction
	WT-2M-1	D222G, F271Y	~8 μg/ml	× 3	basal & induction
	WT-2M-2	M176T, I216N,	~4 μg/ml	× 1	induction
PBP-A-L158E	L158E-4M-2	P105L, T204I, G231S, V263 Del	~4 μg/ml	× 5	basal & induction
	L158E-4M-1	S61T, P66L, Q83H, T167S, Stop codon to Arg	~8 μg/ml	× 4	basal & induction
	L158E-2M	L17P, L205P	~8 μg/ml	× 2	induction

The slight increase in resistance was neither observed by back-transformation nor increased upon next rounds of directed evolutions. Therefore, we assume that the slight increase in M.I.C is (i) either not due to PBP-A, (ii) or, if the resistance is provided by the PBP-A variants, it requires other complementary/compensatory adaptation(s) from the cell.

5.1.2 Evolution of *E. coli/pBAD43_PBP-A-L158E* cells under sublethal concentration of ampicillin

Selection of PBP-A-L158E and PBP-A-wt libraries, expressed from low-copy pBAD43 plasmid in *E. coli*, on ampicillin gradient strip (0.016-256 μg/mL) showed no increase in the minimum inhibitory concentration (MIC) of the cells (i.e. showing the same MIC as wild type *E. coli* Top10, ~2 μg/ml). We scraped and recovered the cells on the edge of the inhibition zone (sublethal concentration of ampicillin) around the ampicillin strip and spread the selected cells on a new plate with another ampicillin strip. This was repeated for 5 days and we observed a progressive MIC increase with the PBP-A-L158E library but

no increase PBP-A-wt library. Therefore, we continued the selection experiment on cells expressing PBP-A-L158E for more days and after 18 days of drift under stress, we observed gradual increase in the MIC and eventually obtained cells with MIC of more than 1500 μg/mL (MIC >1500 μg/mL)(see Figure 5-5).

We also conducted similar selections on PBP-A-L158E and PBP-A-wt libraries in liquid media under sublethal and gradually-increased concentrations of ampicillin and observed similar results (data not presented). Again, only the L158E library evolve to higher MIC up to >1000 μg/mL.

Since the emergence of resistance was only detectable in libraries of PBP-A-L158E and not PBP-A-wt, we hypothesized that improved variants from this library were selected and gradually adapted/fixed by compensatory mutations.

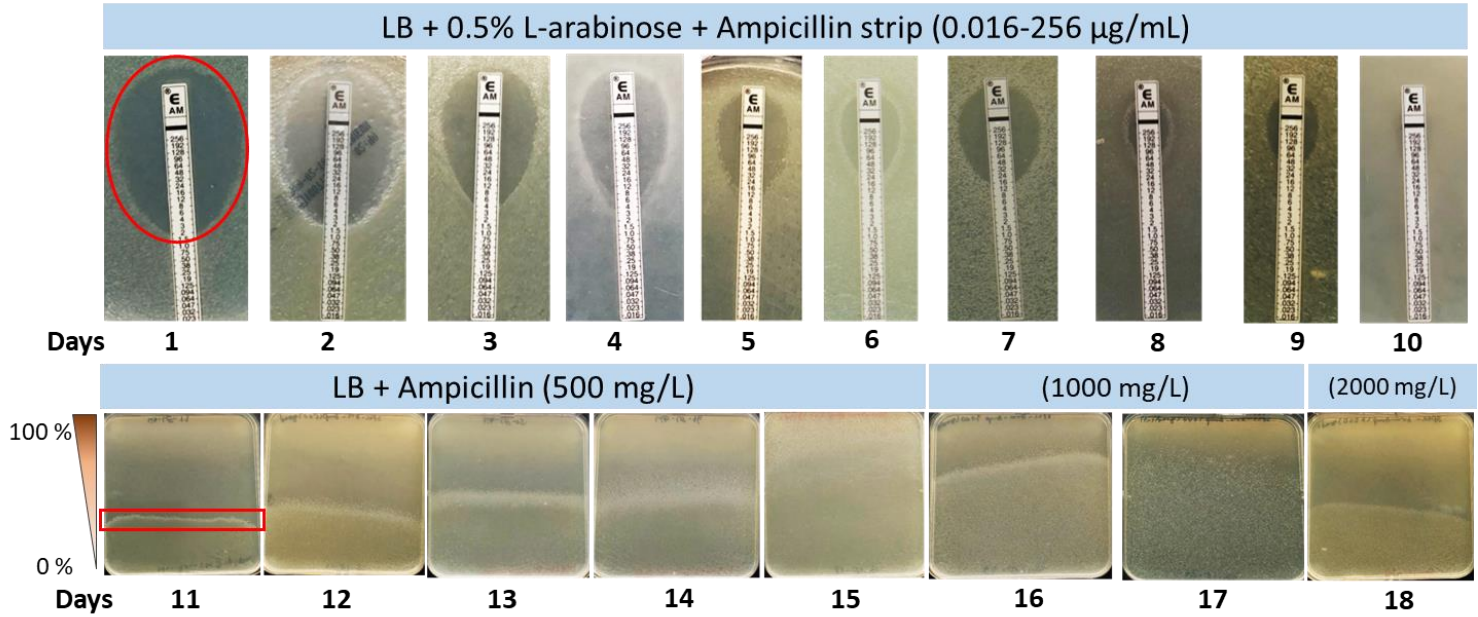


Figure 5-5: Evolution of *E. coli*/pBAD43_PBP-A-L158E under sublethal concentration of ampicillin. Drifting experiment started by selection of PBP-A-L158E library, expressed under low-copy pBAD43 plasmid in *E. coli* Top10, on ampicillin gradient strip (0.016-256 µg/mL). The initial MIC (day 1) showed no increase in the MIC (~2 µg/ml). By drifting the stressed cells under sublethal concentration of ampicillin (inhibition zone around the ampicillin strip, shown in red boxes), for 5 days, an increase on MIC was observed. Therefore, the experiment was continued for further days to evaluate the maximum increase in MIC. After 18 days of drift, we observed gradual increase in MIC and obtained cells with MIC of >1500 µg/ml.

Resistance of the E. coli cells to ampicillin is not provided by pBAD43_PBP-A-L158E plasmid

To check whether the resistant phenotype, is related to PBP-A expression from pBAD43 plasmid or not, firstly, the existence of the plasmid was confirmed by resistance to Spectinomycin as well as plasmid miniprep, and presence of *pbp-a-l158e* gene was confirmed by amplification of the PBP-A gene from the plasmid. Secondly, the extracted plasmid(s) was(were) back-transformed into wild-type *E. coli* Top10 cells, then the MIC was evaluated. Back-transformation of extracted plasmid(s), did not effect the MIC of wild-type *E. coli* Top10 cells (MIC of ~2 μg/ml). This indicated that the plasmid is not the source of resistance.

Resistance of the E. coli cells to ampicillin is not related to presence of the plasmid(s)

To examine the hypothesis that the resistance might be due to adaptation caused by epistatic interactions between plasmid and genome (plasmid seems to be necessary for phenotype emergence), pBAD43 plasmid was cured from the resistant cells using pFree plasmid, which provide a universal plasmid-curing system using CRISPR-Cas9 targeting approach (Lauritsen, Porse, Sommer, & Nørholm, 2017). The curing of pBAD43 plasmid from cells was confirmed by sensitivity to Spectinomycin. The MIC of plasmid-cured cells were similar to cells containing the pBAD43_PBP-A plasmid, which confirms that resistance of the *E. coli* cells to ampicillin is not dependent on presence of this plasmid. We also did PCR on the resistant cells with internal primers binding to PBP-A gene but did not obtained any amplification. However, as the PBP-A has a fused HisTag, we performed western blotting to check any possible expression of PBP-A from chromosome (e.g., via recombination of the gene or plasmid into the genome) we did not detect PBP-A in any fraction of the cell lysate.

A β-lactamase activity in periplasmic extracts of the resistant clones

Resistance to β-lactam antibiotics may occur through various mechanisms, predominantly by production of β-lactamases, the most common mechanism in Gram-negative bacteria, or by alteration in affinity of PBPs against β-lactam antibiotics (Worthington & Melander, 2013). To examine which mechanism is

the case in our resistant clones, we subjected the periplasmic and cytoplasmic lysates of the resistant clone to activity assay against penicillin. B-lactamases are mainly expressed in periplasm of Gram-negative bacteria, hence if we observe such activity it would point towards a β-lactamase as the origin of the resistance phenotype. Lysates of untransformed *E. coli* Top10 cells were used as negative controls and the results of these assays are presented in Figure 5-6.

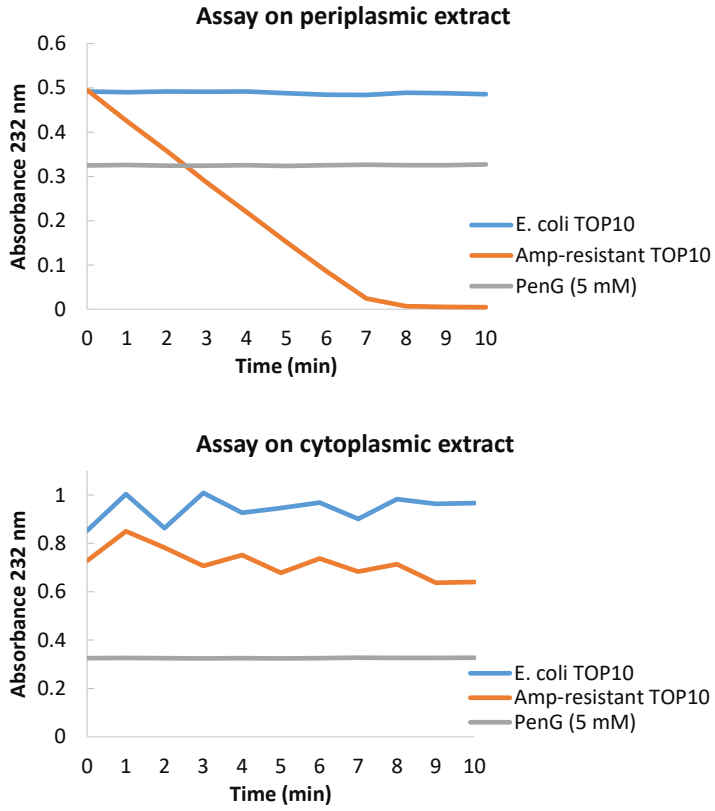


Figure 5-6: Hydrolysis of penicillin detected in periplasmic lysates of the ampicillin-resistant *E. coli* clone.

We observed a strong penicillin hydrolysis activity in periplasmic extract of the resistant clone compare to wild-type *E. coli* Top10. The rate of the hydrolysis suggests a very low K_M (estimated about 6μM), which is 5 fold less than K_M of TEM-1 β-lactamase. Therefore, we hypothesized that either (i) a PBP of *E. coli* has evolved during the serial passage, or (ii) the PBP-A from pBAD43_PBP-A plasmid might have inserted into the chromosome, which also lost the HisTag

(not detectable in western blot). Therefore, we decided to sequence the whole genome of the resistant clone as well as the starting wild-type strain of *E. coli* TOP10 and compare the genotypes in order to discover the origin of the resistant phenotype. Bellow we present the mutations detected from WGS of the ampicillin resistant clone, obtained from serial passage of *E. coli* cells under sublethal concentration of ampicillin.

Whole genome sequencing of resistant E. coli Top10 strain

The whole genome of the resistant strain was sequenced and compared to the genome sequence of starting wild-type strain of *E. coli* Top10. Four variations in three regions across the genome were detected (Table 5-3):

a) Two SNPs upstream of *ampC* gene -11 (C→T) and +31 (C→A) were detected that are located in the box -10 (TATA box) and attenuator region, respectively.

b) Deletion of 11 nucleotides from upstream of *marR* gene. In *E. coli*, MarR is involved in activation of multiple genes related to antibiotic resistance and oxidative stress response. .

c) A missense mutation (Q536L) in *ftsI* gene, known as PBP3 D,D-transpeptidase in *E. coli*.

Table 5-3: List of variants detected in genome of resistant clone of *E. coli* Top10. Four variations in three regions of the genome of resistant clone were detected with 100% frequency. These variants were not detected in sequencing data of the starting *E. coli* wild-type strain.

Region	Type	Reference	Allele	Length	Count	Coverage	Freq
67123	SNV	A	T	1	294	294	100
1706749.. ..1706759	Deletion	TACTTG CCAGG	-	11	33	33	100
4477358	SNV	G	T	1	1023	1023	100
4477399	SNV	G	A	1	1035	1035	100

5.1.3 Drifting PBP-A libraries without ampicillin selection

As fitness cost is suggested to be the main obstacle in directed evolution of PBP-A to β-lactamase, we aimed for mapping the fitness landscape of PBP-A protein. For this, we subjected the libraries of PBP-A-wt and PBP-A-L158E to prolonged

drifting by serial passage under no selection pressure (drift) for PBP-A activity, but only selection for organismal fitness of the *E. coli* cells. Depending on effect of individual mutants on the bacterial growth, selection of beneficial mutations and purge of deleterious mutations will happen over time (see schematic workflow in Figure 5-7). For the non-selective drift, we allowed the libraries of PBP-A under (a) single- and (b) low-copy expression to grow for ~50 generations (5 days) with daily dilutions of 1:1000. To avoid bottlenecking along the drift, we started the drift by inoculating the pooled ep-library and serial passaging (every 24 hours) of at least 100 times more cells ($\geq 1.2\times10^8$ cells) than the initial libraries, which was $\sim 1\times10^6$ individual clones for chromosomal, and $\sim 2\times10^6$ clones for plasmid libraries (Kram et al., 2017). The drift was carried out under two conditions for each library: C_1, basal (non-induced or leaky) and C_2, induced PBP-A expression. For induced condition either L-rhamnose or L-arabinose was added to the growth medium for chromosomal or plasmid libraries, respectively. As described in following sections, we monitored the expression level of PBP-A proteins in the libraries in each day of the drift by western blotting as well as the abundance of variants in the start (t_0) and end (t_5) of the drift by analysis of NGS deep sequencing data.

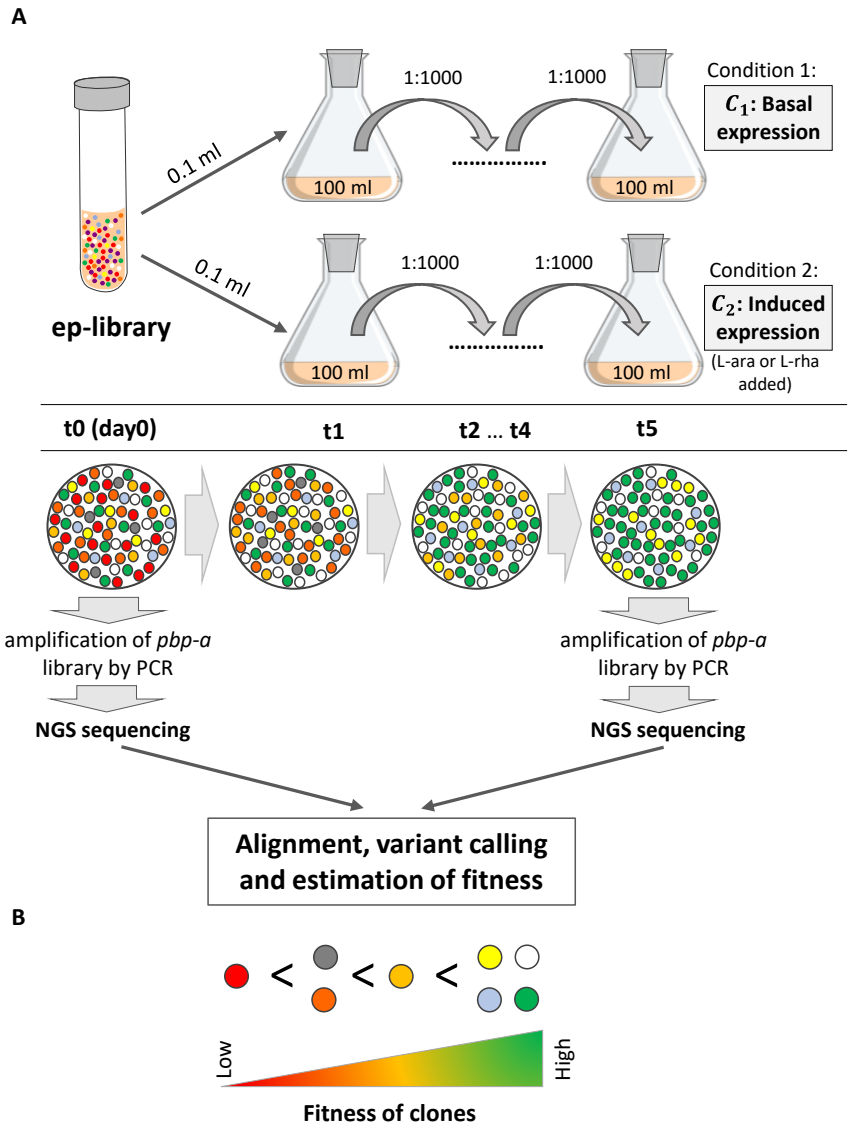


Figure 5-7: Schematic workflow of the prolonged drift of the *pbp-A* library. **A)** The error-prone library (ep-library) of PBP-A was subjected to prolonged drift under basal (non-induced) and induced (adding L-rhamnose or L-arabinose to the growth media for chromosomal or plasmid libraries, respectively) conditions. To avoid bottlenecking along the drift, we started by inoculating and serial passaging (every 24 hours) at least 100 times more cells (0.1 ml of OD ≥1.5, $\geq 1.2\times10^8$) than the initial library diversities (0.5×10^6 to 2×10^6 cells). *pbp-A* gene was amplified at initial library (t_0) and the last day of drift (t_5) and deep sequenced by NGS. The NGS data were analyzed to estimate fitness of individual mutations. **B)** Color scheme comparing the fitness of clones in the library. Mutants of PBP-A that are neutral or introduce the minimum fitness cost to the host *E. coli* cells will survive longer, while mutants with stronger fitness costs will be purged over

time. In the scheme, assume the relative fitness of clones as follow: red<(orange, gray)<gold<(yellow, light-blue, white, green); where red color represents the lowest fitness (higher fitness cost). Therefore, in the end of drift clones with higher fitness (yellow, light-blue, white and green) will win the bulk competition and will be more redundant among sequences in NGS data.

Fate of PBP-A protein expression during drift under no selection

To evaluate the expression level of periplasmic PBP-A proteins during the drift of *pbp-A* libraries, we performed western blot analysis. For that, 0.1 ml of samples, which have been harvested at each day of the drift experiments, was inoculated into 10 ml of media and induced for *pbp-A* expression at OD_{600} ~0.2, grown until OD600 ~0.7 and subjected to periplasmic protein extraction, followed by western blot analysis. Figure 5-8 represents the periplasmic expression profile of *pbp-a* libraries at different time points of the prolonged genetic drift of (a) single-copy (chromosomal) and (b) low-copy (plasmid) libraries of *pbp-a-wt* and *pbp-a-L158E* were analyzed.

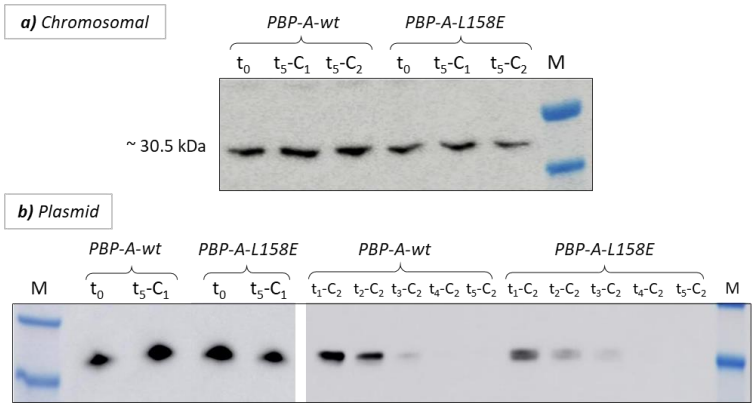


Figure 5-8: Effect of genetic drift on expression of PBP-As. Genetic libraries of PBP-A-wt and PBP-A-L158E under single- and low-copy constructs were subjected to genetic drift for ~50 generations (5 days) without any selection pressure, but the growth of *E. coli* cells. The expression level of periplasmic PBP-A proteins in starting libraries (t_0) is compared with expression level of the same libraries after drifting for 5 days under basal ($t_5 - C_1$) or inducted ($t_5 - C_2$) expression, and for intermediate days, if necessary. For all samples, 0.1 ml of samples from drift experiments was inoculated into 10 ml of media and induced at OD_{600} ~0.2 for *pbp-A* expression, grown until OD_{600} ~0.7 and subjected to periplasmic protein extraction. **A)** Drift under single-copy expression of *pbp-A* from *rhaA* locus in genome of *E. coli*. Expression of both libraries (wt and L158E) appears to be relatively constant during 5 days of drift under both basal and induction conditions. **B)** Drift under low-copy expression of *pbp-A* from pBAD43 plasmid. Expression level in both libraries (wt and L158E) stays constant during 5 days of drift under basal condition (lanes $t_5 - C_1$), but the expression is gradually lost after drifting under induced expression (lanes $t_1 - C_2$ to $t_5 - C_2$).

Loss of PBP-A expression under induced condition in drift under low-copy plasmid: Our results show that expression level of both libraries of *pbp-a* (wt and L158E) has remained constant during 5 days of drift under both basal and induction conditions for the drift under single-copy expression as well as the basal condition for the drift unde low-copy expression. Surprisingly, the expression of both libraries was lost after 5 days of drift under low-copy and induced expression. The periplasmic expression profiles of PBP-A at successive days in this drift (low-copy, induced) show that induction of *pbp-a* library, even in the low-copy pBAD43 plasmid (~ 5 copies/cell), lead to gradual loss of the expression, which can tell us that the fittest cells in the bulk competition were those who does not express PBP-A in their periplasm. For these samples, we also analyzed the cytoplasmic lysate as well as the inclusion bodies by western blot, but no *pbp-a* was detected (results not presented).

NGS sequencing of samples collected from drift under no selection

For NGS sequencing, 2 ml of pooled cells (OD_{600} ≥1.5, corresponds to ≥1.2×10^9 cells) of the initial ep-libraries (t_0) as well as the last day (t_5) of the drifts, were subjected to genomic or plasmid DNA extraction for single- and low-copy libraries, respectively. For chromosomal and plasmid libraries, we amplified a window of 1094 bp and 1070 bp, respectively, which contains extra bases (70–115 bp) upstream and downstream of *DsbA_pbp-a_His* to keep the target gene far enough from extremities of the amplicons (the sequence of full amplicons are available in Appendix 9-25 and Appendix 9-26. The amplicons were purified using Q5® High-Fidelity DNA polymerase, quantified and sent for Ion Torrent or Illumina sequencing for chromosomal or plasmid born libraries, respectively. To avoid bottleneck effect on diversity of libraries during the amplification of *pbp-a* genes, 100 ng of the extracted *E. coli/rhaA::pbp-a* genome for single-copy libraries, and 20 ng of the extracted pBAD43_*pbp-a* plasmids for low-copy libraries were used as PCR templates. (i.e., ~2×10^7 molecules of double stranded DNA of genome, and ~1.3×10^{10} molecules of plasmid).

5.2.1.1 Deep sequencing and mutational scanning of drifted libraries

A bioinformatics pipeline was developed in our lab, which starts by processing and quality control of the reads Figure 5-9. The quality control of reads was performed by assigning Phred quality score (Q) of 20, which define a base call accuracy of 99% (Ewing & Green, 1998). Based on that, the reads were trimmed to remove the extremities with lower quality ($Q < 20$) using Mott algorithm (*Richard Mott personal communication*). After the quality control, the sequences were aligned to the template DNA *(DsbA_pbp-a_6×His)* using Smith-Waterman algorithm based on the 'local mode'. Sequences with poor alignment of bellow 30 bp were discarded. Variant calling was carried out to obtain counts of mutations containing substitutions and InDels. We first evaluated the frequency and distribution of insertions and deletions (InDels), and since these reads will cause frameshift and/or truncation in protein, we decided to avoid InDels in further analysis.

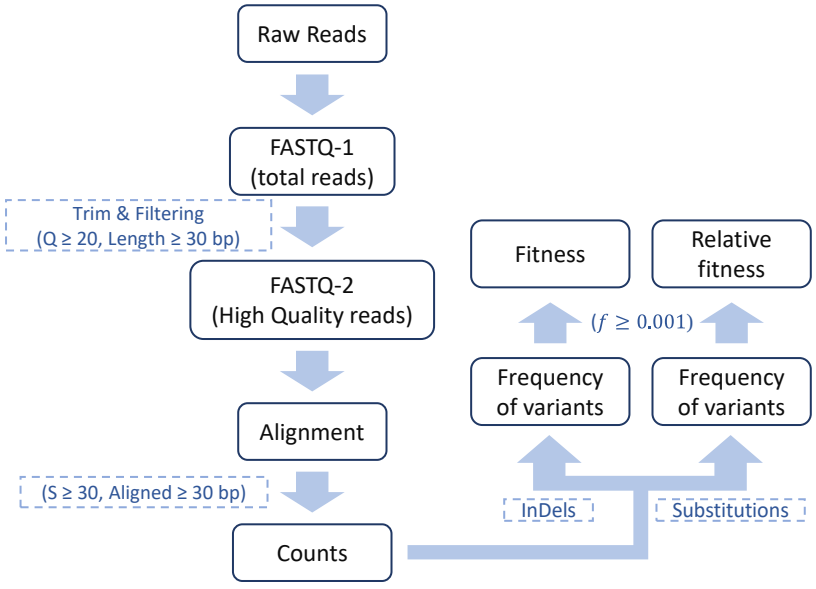


Figure 5-9: Workflow of NGS data analysis. The total reads of NGS data (FASTQ-1) are filtered based on Phred quality score $(Q \geq 20)$, sequences shorted than 30 bp are discarded and the low quality extremities are trimmed to retain high quality bases in the reads (FASTQ-2). Alignment of FASTQ-2 reads were performed Smith-Waterman algorithm and reads with a minimum alignment score (S) of 30 $(S \geq 30)$ with at least 30 bp of aligned sequence were retained for variant calling. Frequency of variants was obtained by calculating counts of each variant relative the total counts of that position.

A threshold of 0.1% ($f \geq 0.001$) was applied to avoid variants with lower frequencies than background frequency, then relative fitness values are calculated by equation.

Initial assessment of NGS data

We initially assessed the NGS data for each library to have a global view of the data. In Table 5-4 we present total counts as wells as average length of the reads, short and low quality reads that were removed, retained reads and mutational load of each library.

Table 5-4: General parameters of the NGS data analysis for different *pbp-A* libraries in the drift experiment.

			NGS raw data		After trimming and filtering				
Copy number	Sample(s)	Condition	Total of reads	Total number of bases	Total of reads	Number of bases	Number of bases with high quality (Q≥20)	Average length of reads (base pairs)	NGS Sequencing
(*a*) Chromosomal (single-copy)	PBP-A-wt	t_0	2.58E+05	3.76E+07	2.58E+05	3.40E+07	3.32E+07	132 bp	Ion Torrent
		$t_5 - C_1$	3.22E+05	4.65E+07	3.22E+05	4.21E+07	4.12E+07	131 bp	
		$t_5 - C_2$	4.62E+05	7.74E+07	4.62E+05	7.00E+07	6.82E+07	151 bp	
	PBP-A-L158E	t_0	3.35E+05	5.57E+07	3.35E+05	5.02E+07	4.89E+07	150 bp	
		$t_5 - C_1$	3.45E+05	4.43E+07	3.45E+05	4.03E+07	3.95E+07	117 bp	
		$t_5 - C_2$	4.15E+05	5.62E+07	4.15E+05	5.05E+07	4.93E+07	122 bp	
(*b*) Plasmid (low-copy)	PBP-A-wt	t_0	4.49E+06	6.64E+08	4.49E+06	6.28E+08	6.22E+08	140 bp	Illumina

		$t_5 - C_1$	2.82E+06	4.17E+08	2.82E+06	3.94E+08	3.91E+08	140 bp	
		$t_5 - C_2$	4.09E+06	6.06E+08	4.09E+06	5.73E+08	5.67E+08	140 bp	
	PBP-A-L158E	t_0	4.05E+06	6.03E+08	4.05E+06	5.71E+08	5.65E+08	141 bp	
		$t_5 - C_1$	4.05E+06	6.01E+08	4.05E+06	5.68E+08	5.62E+08	140 bp	
		$t_5 - C_2$	4.90E+06	7.29E+08	4.90E+06	6.90E+08	6.83E+08	141 bp	

Coverage of the sequence space

All the possible missense single point mutations (ie., all amino acids accessible by single nucleotide substitutions) of the 271 codons *pbp-a-wt* and *pbp-a-L158E* genes are 1618 and 1619 substitutions, respectively, thus easily accessible for our ep-libraries, which have a minimum size of 0.5×10^6. The total possible double point mutations of 271 codons are 2583 and 2582 for WT and L158E, respectively, while for triple mutation 948 mutations are possible. Compare to single nucleotide mutations, double and triple mutations require large libraries to fully cover the corresponding mutational space. However, since the Mutazyme II DNA polymerase is introducing single point mutations probability of two or three mutations happening in the same codons become lower. The coverage of single (1nt), double (2nt) and triple nucleotide (3nt), as well as the type of generated codons are available in Table 5-5.

Table 5-5: The theoretically possible and observed mutational space of *pbp-A* wt and L158E libraries. a and b represent chromosomal and plasmid libraries, respectively. Condition 1 (C_1) and condition 2 (C_2) represent basal and induced expression, respectivey.

Number of mutations in PBP-A proteins											
			Nonsynonymous mutations						Synonymous Mutations		
			Missense			Nonsense					
			1nt	2nt	3nt	1nt	2nt	3nt	1nt	2nt	3nt
WT	**All possible**		**1618**	**2583**	**948**	**49**	**169**	**53**	**259**	**57**	**13**
	a	t_0	1573	200	2	49	13	0	259	0	0
		$t_5 - C_1$	1588	253	4	49	15	0	259	0	0
		$t_5 - C_2$	1605	366	6	49	30	0	259	0	0
	b	t_0	1618	1567	10	49	124	0	259	0	0
		$t_5 - C_1$	1618	1310	6	49	120	2	259	0	0
		$t_5 - C_2$	1618	688	6	49	42	0	259	0	0
LE	**All possible**		**1619**	**2582**	**948**	**50**	**168**	**53**	**259**	**56**	**13**
	a	t_0	1612	478	4	50	34	0	259	0	0
		$t_5 - C_1$	1616	365	3	50	28	0	259	0	0
		$t_5 - C_2$	1617	435	3	50	37	0	259	0	0
	b	t_0	1619	1550	10	50	132	0	259	0	0
		$t_5 - C_1$	1619	1367	11	50	120	0	259	0	0
		$t_5 - C_2$	1618	816	7	50	59	0	259	0	0

Deep sequencing results show that, 1572 out of 1618 (97%) mutations for *pbp-a-wt* library and 1611 out of 1619 (99%) for *pbp-a-L158E* library are present in the initial chromosomal libraries, while we have all (100%) of the possible mutants present in plasmid libraries (Table 5-5). Detection of less mutations in chromosomal libraries is expected as the size of these libraries (~1×10^6 individual clones) are smaller than plasmid libraries (2×10^6 clones), and/or could also be due to less reads of Ion torrent sequencing compare to Illumina sequencing (Table 5-4).

The significant decrease of 2nt and 3nt nonsynonymous substitutions is most probably due to loss of the active variants and effect of InDels (see next section), where the diversity of library is decreased upon enrichment of InDels, thus less accessible mutations such as 1nt and 2nt are less probable to maintain.

Frequency and fitness of InDels

In production of ep-libraries by Mutazyme II DNA polymerase, there is a spectrum up to 0.7% and 4.8% to generate insertion and deletion mutations,

respectively. However, InDels are also possible to emerge during assembly of epPCR amplicon to the expression genetic construct. Table 5-6 represents percentage of InDels observed at initial libraries (t_0) of PBP-A.

Table 5-6: Percentage of InDels in initial libraries (t_0) of PBP-As. Reads containing InDels are divided by the depth (total reads) of the NGS data.

	Reads containing InDels	
	PBP-A-wt	PBP-A-L158E
Chromosomal	11.7 %	14.5 %
Plasmid	4.4 %	4.0 %

As can be seen in Table 5-6, there is a negative correlation between InDels and NGS depth, meaning that the more deep the sequencing is the less chance to count InDels. To evaluate fate of InDels during the drift of PBP-A libraries, we compare their fitness values, which are obtained from the frequency of variants during the drift in comparison with the initial library. For site r (nucleotide position of the InDels) if frequency of $InDel$ at initial library (t_0) is $f^0_{InDel,r}$, frequency of that $InDel$ at timepoint t in the drift is $f^t_{InDel,r}$. The fitness of $InDel$ (W_{InDel}) at site r can be calculated as:

$$W_{InDel,r} = \frac{f^t_{InDel,r}}{f^0_{InDel,r}} \qquad (5.1)$$

A threshold of equal or more than 0.1% $(f \geq 0.1\%)$ was applied to avoid variants lower than background frequencies, and the fitness values are calculated accordingly. InDels are retained, purged or fixed during the drift experiment. As the distribution of fitness effect of InDels in both basal and induced conditions in chromosomal libraries (conditions $a - C_1$ and $a - C_2$) as well as basal condition in plasmid libraries ($b - C_1$), for both PBP-A-wt (WT) and L158E (LE), is centered around neutral (mean ($\overline{X}$) = 1), we used the variability in the fitness effects of InDels to categorize their effects. For that, we chose a threshold of two standard deviations over the mean ($\overline{X}$+2SD) of the fitness values of InDels to assign the beneficial fitness effect, a threshold of two standard deviations under the mean ($\overline{X}$-2SD) to assign deleterious fitness values (Table

5-7). The values of W_{InDel} in condition $\mathrm{b} - \mathrm{C}_2$ of both WT and L158E libraries were not used due to their highly variable distribution.

Table 5-7: Assignment of thresholds for categorization of the fitness effect of InDel variants in the drift experiments. Fitness of InDels were calculated using equation 5.1, from which mean ($\overline{\mathrm{X}}$) and standard deviation (SD) were calculated.

Fitness values for total InDels variants					
		(a) Chromosomal		(b) Plasmid	
		Basal $(a - C_1)$	**Induction** $(a - C_2)$	**Basal** $(b - C_1)$	**Mean**
WT	$\overline{\mathrm{X}}$	1.14	1.1	0.85	1.0
	SD	0.23	0.3	0.15	
	$\overline{\mathrm{X}}$-2SD	0.68	0.5	0.55	0.58
	$\overline{\mathrm{X}}$-SD	0.91	0.8	0.7	0.80
	$\overline{\mathrm{X}}$+2SD	1.6	1.7	1.15	1.48
L158E	$\overline{\mathrm{X}}$	0.98	1.16	0.9	1.0
	SD	0.26	0.47	0.14	
	$\overline{\mathrm{X}}$-2SD	0.46	0.22	0.62	0.43
	$\overline{\mathrm{X}}$-SD	0.72	0.69	0.76	0.72
	$\overline{\mathrm{X}}$+2SD	1.5	2.1	1.18	1.59
Thresholds assignment					
Deleterious			$W < 0.5$		
Neutral			$0.5 \leq W \leq 1.5$		
Beneficial			$W > 1.5$		

Therefore, if the fitness (W) of an $InDel$ is between 0.75 and 1.0 ($0.5 \leq W \leq 1.5$), it is neutral and therefore retained. While, if the value of fitness is less than 0.5 ($W < 0.5$), the $InDel$ is purged, and if the value of relative fitness is greater than 1.5 ($W > 1.5$), the $InDel$ is enriched. Based on that, we analyzed distribution of fitness effects of InDels on each library at basal and induced conditions and the results re presented in Figure 5-10. We then analyzed distribution of fitness effects of InDels on each library under basal and induced conditions. For this, the mutations were binned by unit interval of 0.25, ranging from 0.25 to more than 2. The bins were categorized into four classes based on the thresholds calculated previously (Table 5-7). The results are presented in Figure 5-10.

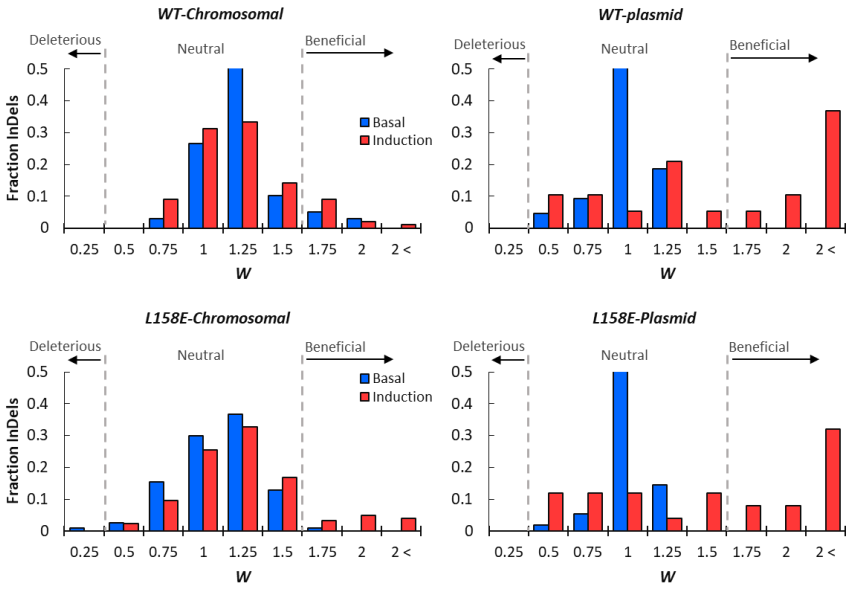


Figure 5-10: Distribution of fitness effect of InDels after 5 days of drift of libraries.

InDels mostly remained neutral during the drift under the basal expression of PBP-A in both chromosomal and plasmid context. This was expected since (i) there is no selection pressure for activity of PBP-A and (ii) the expression level is very low to cause misfolding due to inappropriate protein expression. However, under induction conditions in chromosomal libraries, the InDels become slightly beneficial for fitness. The beneficial effect of InDels on fitness of *E. coli* cells, significantly detectable under induced expression in plasmid libraries. This is in agreement with our observations, where we observed loss of expression of PBP-A proteins (Figure 5-8). This suggest that, during the drift under higher expression (induced condition on plasmid context), the highest fitness for *E. coli* cells was achieved when the PBP-A protein was totally lost, most probably by frameshift in the open reading frame (ORF) of PBP-A.

To evaluate the distribution of InDels across the ORF of *DsbA-pbp-a-His*, and to address the question that whether the InDels are generated by Mutazyme II DNA polymerase or generated during assembly of genetic constructs, the positional fitness values $W_{InDel,r}$ were plotted alongside the ORF. Figure 5-11

represents the distribution of InDels' fitness across the 894 bp ORF of *DsbA-pbp-a-His*, in different conditions of the drift.

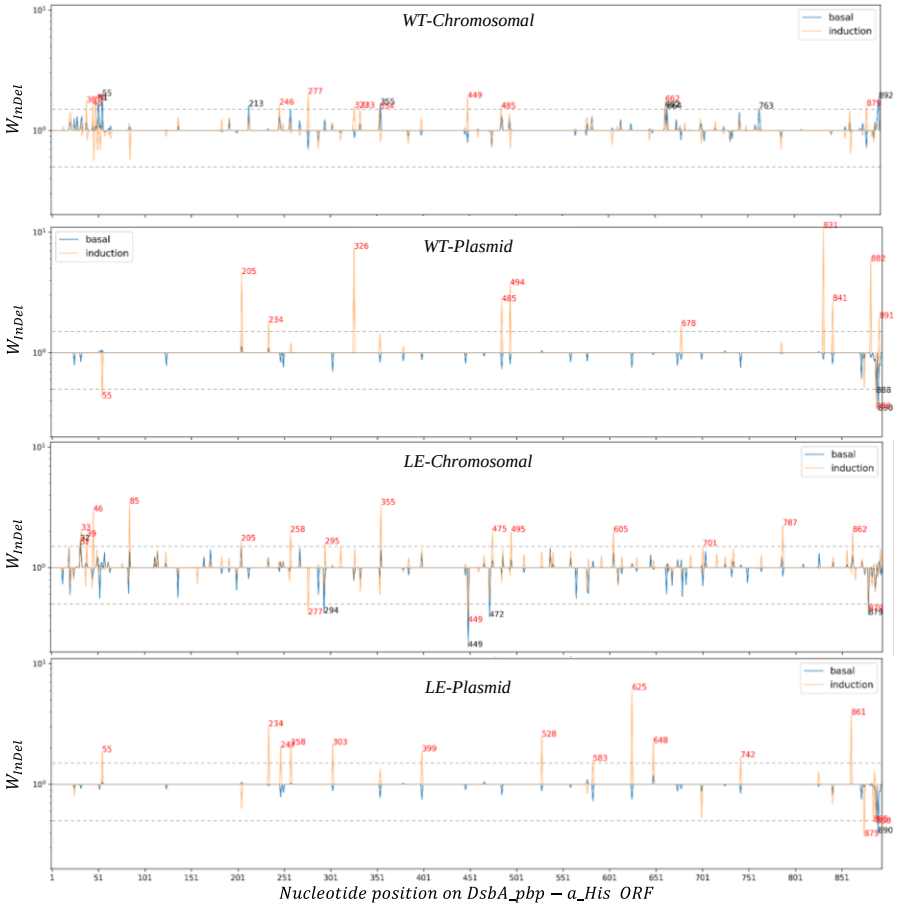


Figure 5-11: Distribution of fitness due to InDels (W_{InDel}) across the ORF of *pbp-A* WT and LE after the drift.

The distribution of W_{InDel} values show that, InDels are spread across the PBP-A ORF, which suggest their generation by the Mutazyme II DNA polymerase. However, the higher density of InDels at the beginning (position 40~60 bp) and the end (positions 880~894) of the *pbp-A* gene, where the overlapping region between ep-library and the expression construct exists, suggests that these InDels are generated during assembly of the epPCR amplicon into the expression construct.

Under basal expression from plasmid ($\boldsymbol{b - C_1}$) and less in basal expression from chromosome ($\boldsymbol{a - C_1}$), InDels seems to be slightly deleterious (W<1), which may suggest that in these conditions, where fitness cost is minimized, the active variants do not impose high fitness cost and will mostly be retained during the drift, thus outgrowing the variants with InDels. Or it could be that frameshift of the plasmid and chromosomal DNA is costly under these neutral conditions, hence causing the purge of variants with InDels.

Frequency and fitness of nonsynonymous substitutions

To analyze substitutions, we removed the reads containing indels. However, our sequencing approaches (Ion torrent and Illumina) do not cover full-length sequencing and, therefore, reads without InDels from variants with InDel will not be discarded and may affect analysis.

We initially estimated average number of mutations per gene at different conditions through the drift experiment. Average number of mutations/gene was obtained using equation bellow:

$$x^t = \frac{N_{s\,(or\,ns)}^t}{N^t} \times length\ of\ the\ gene \qquad (5.2)$$

where x^t is the average number of mutations per gene, N_s^t or N_{ns}^t is the number of reads with nonsynonymous (missense plus nonsense) or synonymous mutations, respectively, at timepoint t, and N^t is the total number of reads at time point t. Total number of nonsynonymous and synonymous mutations are calculated, and from that, we obtained the ratio of nonsynonymous over synonymous mutations per gene that is presented in Figure 5-12.

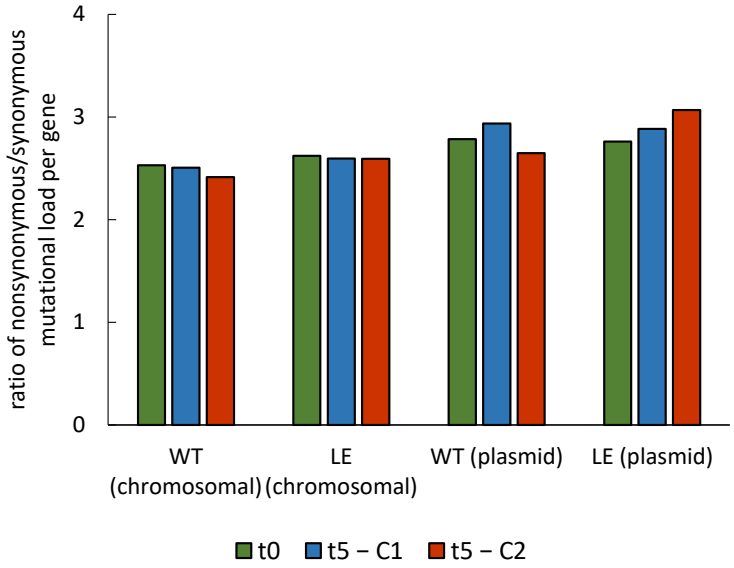


Figure 5-12: The ratio of nonsynonymous over synonymous mutations per gene across the timepoints of the serial passage in different conditions.

The ratio of nonsynonymous/synonymous mutations in all drifts are slightly decreasing for induced condition (t5-c2) compared to initial library (t0), which indicates a purifying selection at the protein level. However, for the libraries of PBP-A-L158E in plasmid an opposite behaviour is observed, meaning that nonsynomymous mutations are in average more beneficial than synonymous ones.

To evaluate fate of individual variants of PBP-A, we compare their relative fitness values, which are obtained from the redundancy of variants during the drift relative to the redundancy of each variant in the initial library and the redundancy of corresponding wild-type residue. For that, we calculate change in frequency of mutants along the drift, which reflects the effects of non-selective drifting on each mutation or, as defined here, their relative fitness (W_{rel}) effects (Rockah-Shmuel et al., 2015). For a residue at position r, total number of reads are N_r , where counts with wild-type residue x are n_x and counts with mutant amino acid m are n_m , frequency of x and m are calculated by equation 5.3 and 5.4, respectively:

$$f_{x,r} = \frac{n_{x,r}}{N_r} \qquad (5.3)$$

$$f_{m,r} = \frac{n_{m,r}}{N_r} \qquad (5.4)$$

To avoid variants lower than background frequencies we applied a threshold of 0.1%. Using frequencies of each missense or nonsense mutation relative to the wild-type amino acid, we can obtain the W_{rel} as follow: if frequency of x at initial library (t_0) is $f_{x,r}^0$, frequency of x at timepoint t in the drift is $f_{x,r}^t$, and similarly, if the frequency of m at initial library is $f_{m,r}^0$, frequency of m at timepoint t in the drift is $f_{m,r}^t$. The relative fitness (W_{rel}) for m at site r can be calculated as:

$$W_{rel} = \frac{\left(\frac{f_{m,r}^t}{f_{x,r}^t}\right)}{\left(\frac{f_{m,r}^0}{f_{x,r}^0}\right)} \qquad (5.5)$$

Mutations are retained, purged or enriched during the drift experiment. As the distribution of relative fitness missense mutations in both basal and induced conditions in chromosomal libraries as well as basal condition in plasmid libraries, for both WT and LE, is centered around neutral (mean = 1), we used the variability in the fitness effects of these variants to categorize their effects. For that, we chose a threshold of two standard deviations over the mean ($\overline{X}$+2SD) of the relative fitness values (W_{rel}) to assign the beneficial fitness effect for greater values, while a threshold of two standard deviations under the mean ($\overline{X}$-2SD) to assign deleterious fitness effect of smaller values (Table **5-7**). Due to the high variability of W_{rel} for missense mutations in condition $\mathrm{b} - \mathrm{C}_2$ of both WT and LE libraries, these were not used for calculations of the thresholds. The average W_{rel} values and standard deviation (SD) for synonymous mutations obtained in different libraries are presented in Table 5-8.

Table 5-8: Assignment of thresholds for categorization of the relative fitness effect of nonsynonymous variants in the drift experiments. Relative fitness (W_{rel}) of the variants were calculated using equation 5.5, from which mean ($\overline{X}$) and standard deviation (SD) were calculated.

W_{rel} values of total nonsynonymouse mutations					
		(a) Chromosomal		(b) Plasmid	
		Basal $(a - C_1)$	Induction $(a - C_2)$	Basal $(b - C_1)$	Mean
WT	$\overline{X}$	1.1	0.97	0.87	1.0
	SD	0.2	0.19	0.21	
	$\overline{X}$-2SD	0.69	0.593	0.438	0.57
	$\overline{X}$-SD	0.9	0.78	0.66	0.78
	$\overline{X}$+2SD	1.5	1.35	1.29	1.38
L158E	$\overline{X}$	1.05	0.99	0.94	1.0
	SD	0.18	0.16	0.12	
	$\overline{X}$-2SD	0.69	0.67	0.7	0.69
	$\overline{X}$-SD	0.51	0.51	0.58	0.53
	$\overline{X}$+2SD	1.41	1.31	1.18	1.30
Thresholds assignment					
Deleterious		$W_{rel} < 0.6$			
Neutral		$0.6 \leq W_{rel} \leq 1.3$			
Beneficial		$W_{rel} > 1.3$			

Therefore, if the value of relative fitness of a missense or nonsense mutation is between 0.7 and 1.3 $(0.6 \leq W_{rel} \leq 1.3)$, mutation ($m$) is neutral and therefore retained during the drift. While, if the W_{rel} is lower than 0.6 $(W_{rel} < 0.6)$, the m is purged, and if the W_{rel} is greater than 1.3 $(W_{rel} > 1.3)$, the m is enriched. We then analyzed distribution of fitness effects of mutations on each library under basal and induced conditions. For this, the mutations were binned by unit interval of 0.1, and the bins were categorized into three classes based on the calculated thresholds in Table 5-8. Figure 5-13 represents the global distribution of nonsynonymous and nonsense mutations based on their relative fitness (W_{rel}) values.

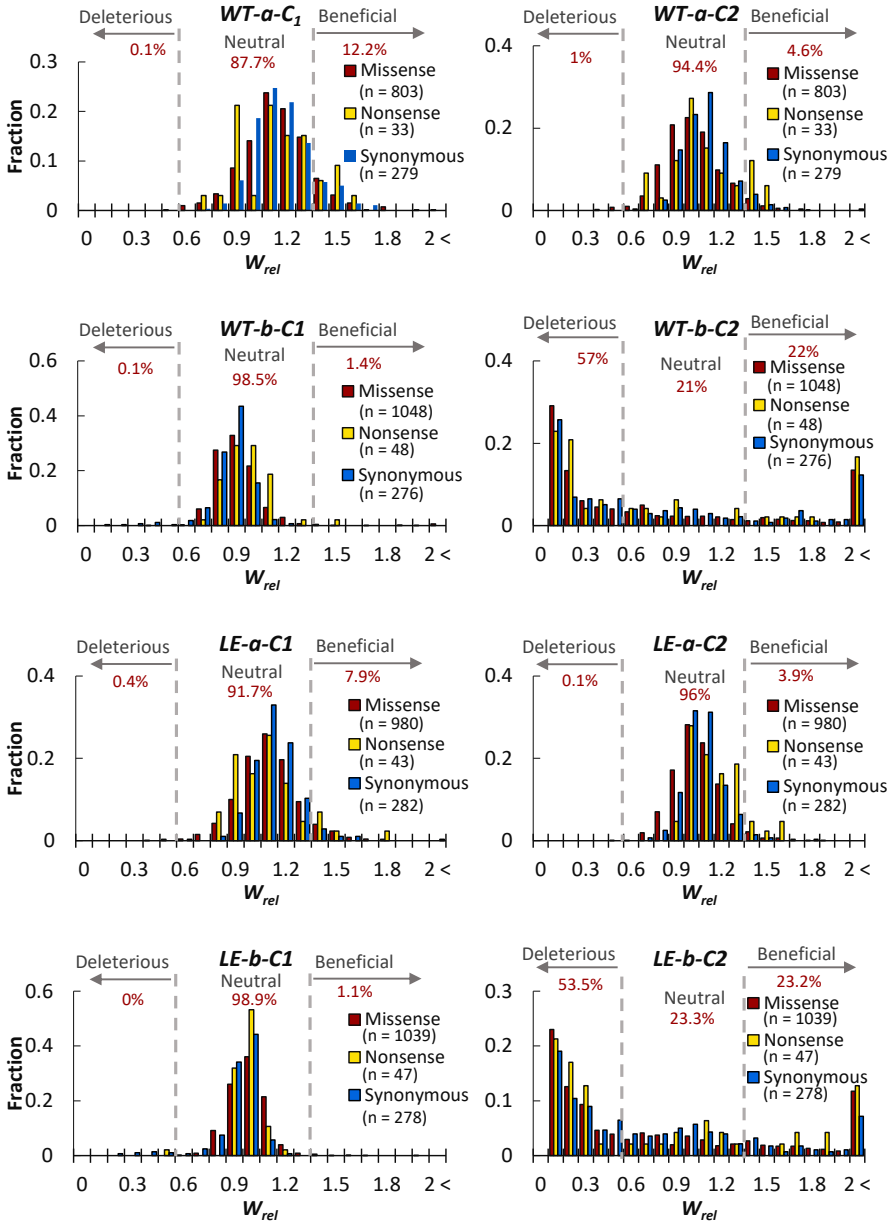


Figure 5-13: The distribution of fitness effects of missense, nonsense and synonymous mutations in drift of PBP-A-wt (WT, top) and PBP-A-L158E (LE, bottom) under basal and induced expression in single-copy chromosomal context, a-C_1 and a-C_2, respectively, as well as basal and induced expression in low-copy plasmid context, b-C_1 and b-C_2, respectively. Values of relative fitness (W_{rel}) are separated in bins of 0.1. Based on thresholds calculated in Table **5-8**, values of W_{rel} including and between 0.7 and 1.3 ($0.6 \leq W_{rel} \leq 1.3$) are considered as neutral, while W_{rel} less than 0.6 ($W_{rel} < 0.6$) as deleterious and W_{rel} more than 1.3 ($W_{rel} > 1.3$) as beneficial. Percentage of missense mutations in

each category is presented on the charts in red color. *n* represents number of total mutations analyzed in each chart.

Under both basal and induced expression in chromosomal context (conditions a-C_1 and a-C_2), as well as basal condition in plasmid context (b-C_1), and for both WT and LE, missense and nonsense mutations are predominantly neutral (Figure 5-13). However, under chromosomal context (a-C_1 and a-C_2), a small fraction of mutations are associated to a relatively high fitness ($W_{rel} > 1.3$), hence located in the beneficial category. Majority of these beneficial mutations are nonsense, thus producing truncated proteins.

In plasmid-induction ($b - C_2$) condition, majority of missense and nonsense mutations are purged due to deleterious effect (fitness cost) while a relatively low amount is fixed. This is in agreement with our observation on loss of PBP-A expression due to selection of InDels, leading to purge of the full length protein variants. However, the relatively low amount of missense and nonsense mutations that are fixed under $b - C_2$ condition may belong to the remaining segments of variants with InDels that has been selected in this condition. Since the sequencing covers only up to 150 bp reads, although we discarded reads with InDels in analysis of missense and nonsense mutations, the residual part of the InDel variant will remain among the non-InDels reads. Therefore, observations in this condition has been masked by the InDels need to be rejected to not cause misleading. To overcome such constraints, approaches such as full-length NGS sequencing and/or bioinformatically traceable DNA barcodes can be applied in the prolonged drift experiments.

Drifting PBP-A libraries under lower expression conditions (*a-C_1, a-C_2* and *b-C_1*) is predominantly neutral, while in higher expression (induced expression from plasmid) nonsynonymous and nonsense mutations are predominantly purged due to selection of the InDels and/or due to deleterious effects, which is not concludable as the InDels are masking the data. However, small subsets of mutations are outside the defined thresholds that are slightly enriched in all both conditions of the single-copy expression. To further evaluate the possible origin of these effects, in next section we systematically analyze the distribution of

fitness for amino acids in the active site, surface, and hydrophobic core residues of PBP-A. However, to verify the origin of the deleterious and beneficial effect of mutants on fitness of *E. coli* cells, extensive characterization of the mutants need to be carried out.

Distribution of fitness for nonsynonymous substitutions on the PBP-A protein

To evaluate any possible correlation between distribution of fitness based on types of nonsynonymous mutations and types of residues in the PBP-A protein, we evaluate the distribution first on the full-length PBP-A protein, and second on different region of the protein. Figure 5-14 illustrates the distribution of relative fitness values for all missense substitutions across the PBP-A-wt and L158E proteins as well as distribution based on type of amino acids (e.g., acidic, basic, polar and nonpolar) under different conditions of the neutral drift.

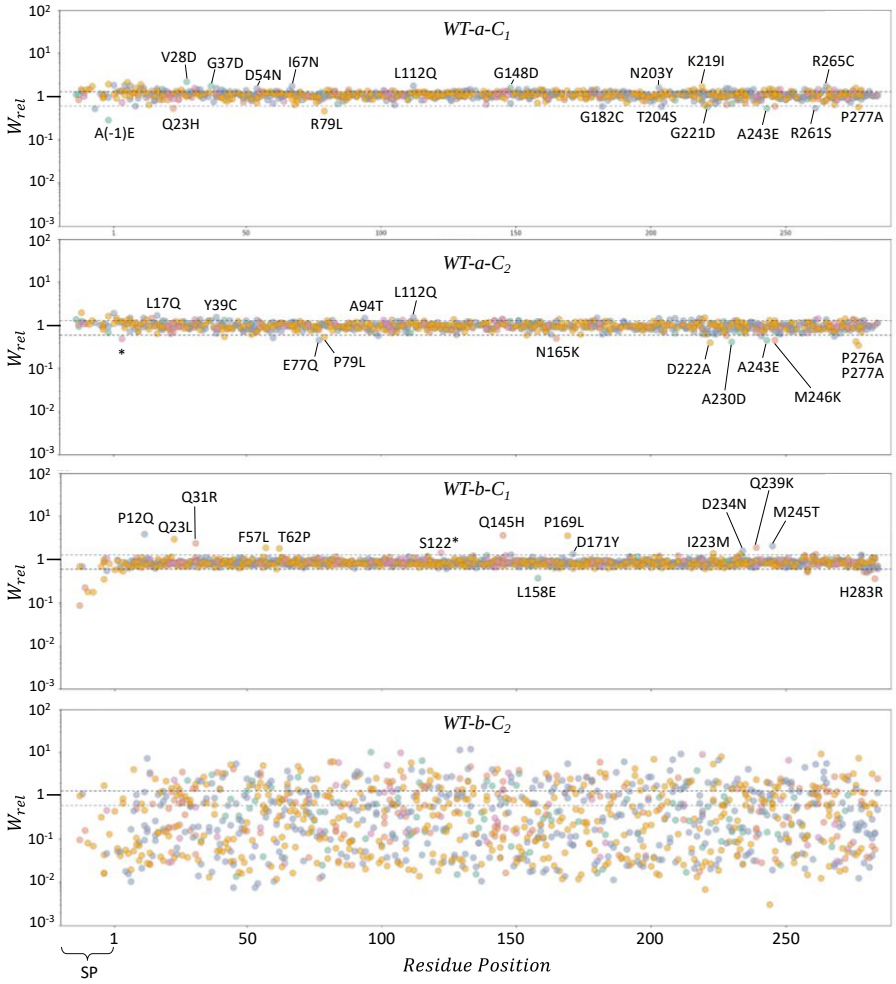


Figure 5-14: Distribution of relative fitness (W_{rel}) of nonsynonymous substitutions on the PBP-A-wt (WT) protein under different conditions of the drift. N-terminal DsbA signal peptide, which was not subjected to epPCR is indicated as SP. Dashed lines represent the thresholds for beneficial (W_{rel} >1.3) and deleterious (W_{rel} < 0.6) fitness values. Color code on the graph represent the types of substituted amino acids. Distribution of W_{rel} for nonsynonymous substitutions on the PBP-A-L158E protein is available in Appendix 9-27.

As it is clearly observable in Figure 5-14, the fitness effect of nonsynonymous is equally distributed among different types of substitutions as well as along the PBP-A proteins (i.e., there is no region on the protein to be significantly affected). However, from very low (*a-C1*) to low (*a-C2*), and then to slightly higher (and *b-C1*) expression conditions (from top to bottom in Figure 5-14), there

seems to be a slight purge of total mutations towards (i.e., the mean W_{rel} is decreasing). This suggests that the higher the expression (even slightly) the stronger the effect of the drift, and it also indicates that if the drift is continued for a longer period, the effect might emerge and provide more informative data. Next, we discuss the distribution of missense mutations in conditions *a-*C_1 *, a-*C_2 and *b-*C_1 based on three types of residues: (i) active site, (ii) surface (exposed), and (iii) hydrophobic core (buried). The residues in the active site motifs include, S61 and K64 in the SxxK motif, E96 anchoring the active site cavity, S122, D123 and N124 in the SDN motif, L158 in the omega-loop, K219, T220, G221 in the KTG motif and D222 adjacent to the KTG motif. To define exposed or buried residues in PBP-A, we used RaptorX Property server (S. Wang, Li, Liu, & Xu, 2016) based on their relative solvent accessibility (RSA) values. RSA for each residue will be considered as buried if $RSA < 25\%$, or exposed if $RSA > 75\%$. represent fitness effects of missense substitutions on residues of active site, surface and the hydrophobic core of PBP-A. Figure 5-15 and Figure 5-16 represent the distribution of fitness effects of missense and nonsense mutations on residues of active site, surface and hydrophobic core of PBP-A enzyme. Since we observed fitness-improving effect of changing the pI of PBP-A, we analyzed the fitness distribution of missense substitutions on surface of the enzyme, which are generating positively- (Lysine or Arginine) or negatively-charged (Glutamic Acid or Aspartic Acid) amino acids. Figure 5-17 represents distribution of missense substitution for generation of the positively- and negatively-charged amino acids.

Table 5-9: Fitness effect of missense substitutions on residues of active site, surface and hydrophobic core. Percentage of deleterious, neutral and beneficial mutations are presented. *Del, deleterious; neu, neutral; ben, beneficial.*

		Percentage of mutations observed								
		$a - C_1$			$a - C_2$			$b - C_1$		
		Del	*Neu*	*Ben*	*Del*	*Neu*	*Ben*	*Del*	*Neu*	*Ben*
WT	*Active site*	0.0 %	91.9 %	8.1%	2.7 %	91.9 %	5.4 %	2.3 %	97.7 %	0%
	Exposed	0.3 %	67.1 %	11.2 %	1.0 %	73.2 %	4.5 %	0%	98.8 %	1.2 %
	Buried	0.0 %	87.9 %	12.1 %	1.3 %	94.9 %	3.8 %	0%	98.5 %	1.5 %
LE	*Active site*	2.3 %	88.4 %	9.3%	0.0 %	93.0 %	7.0 %	0%	100 %	0%
	Exposed	0.3 %	92.4 %	7.3%	0.0 %	95.7 %	4.3 %	0%	98.8 %	1.2 %
	Buried	0.3 %	93.4 %	6.3%	0.3 %	96.8 %	2.9 %	0%	98.7 %	1.3 %

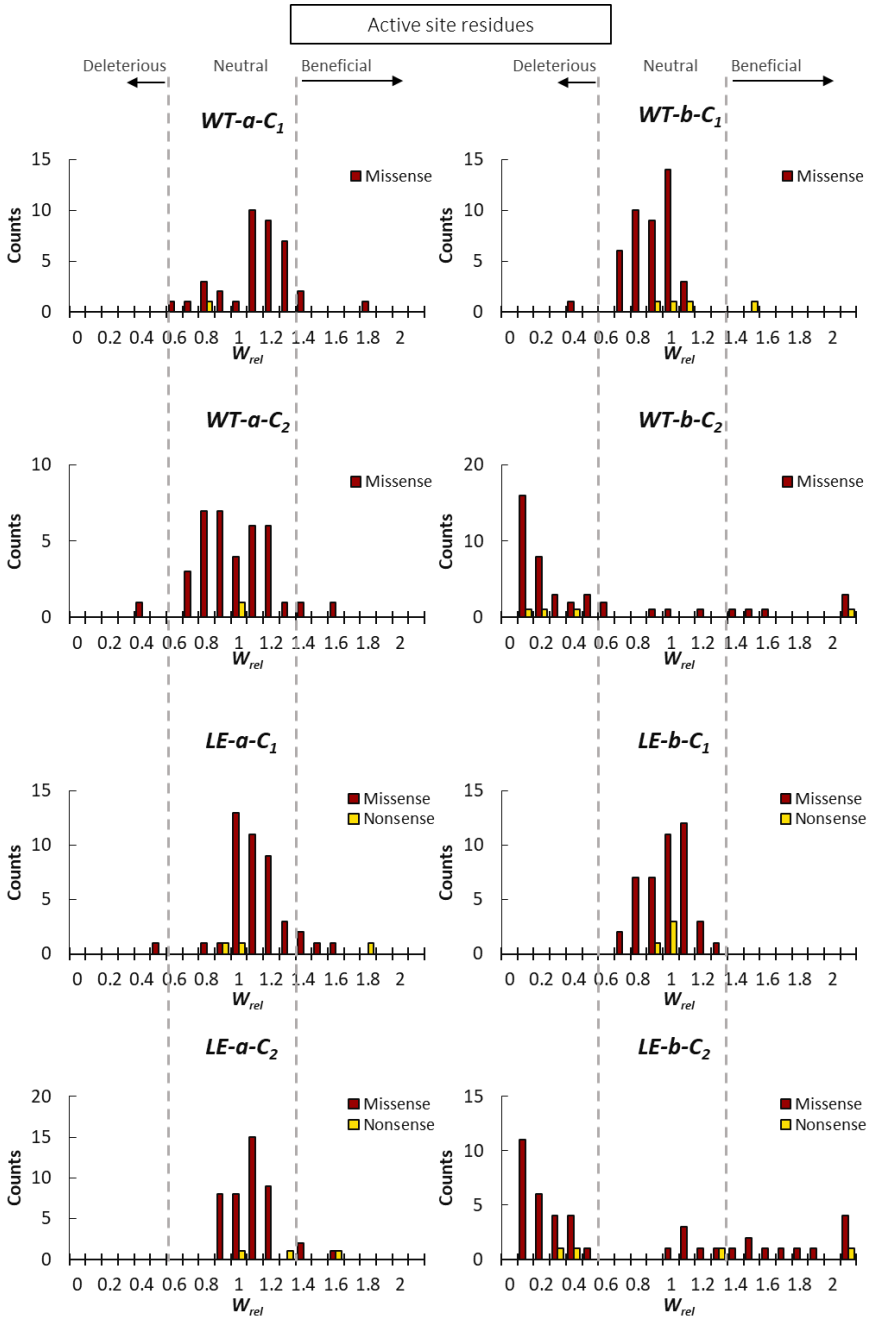


Figure 5-15: Distribution of fitness effects of nonsynonymous substitutions in residues involved in the active site of PBP-A-wt and L158E.

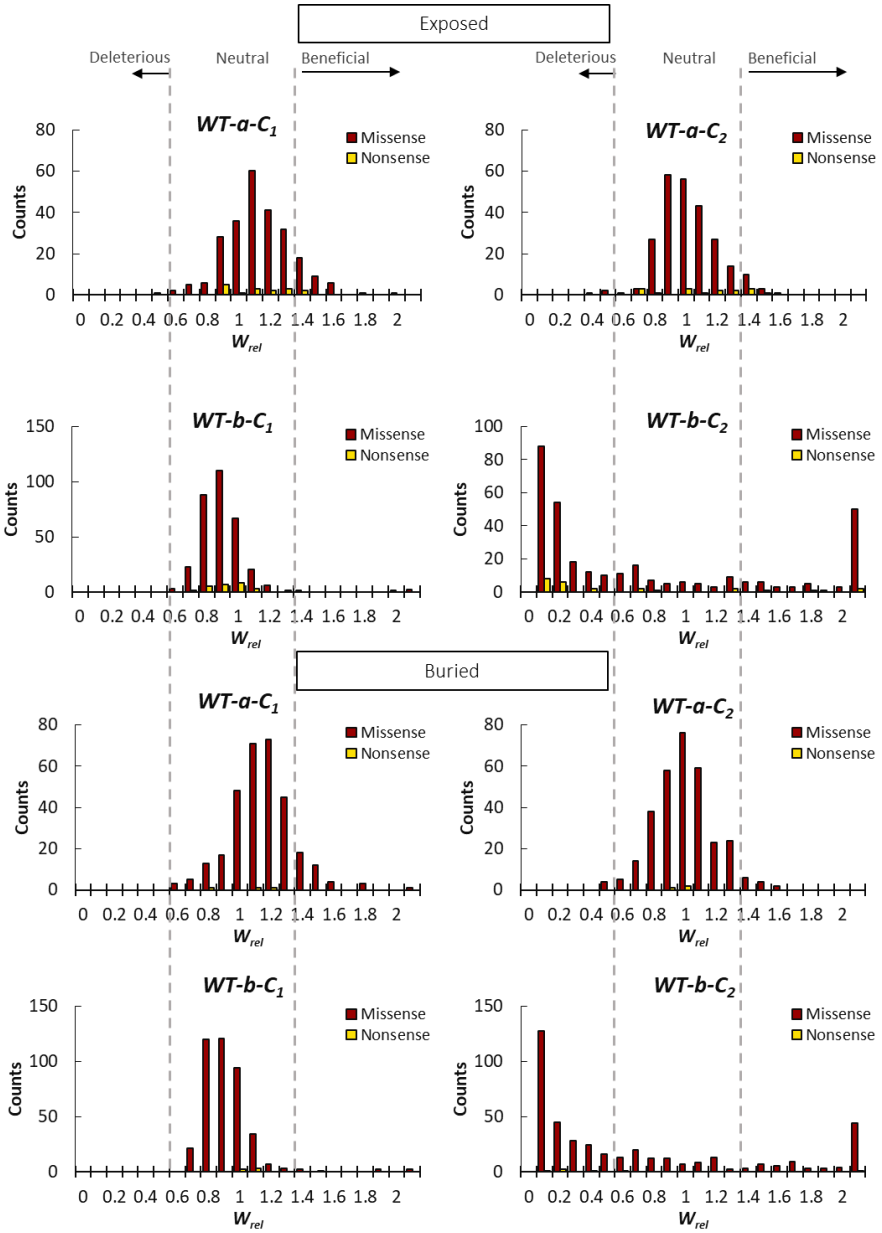


Figure 5-16: Distribution of fitness effects of nonsynonymous substitutions in exposed and buried residues of PBP-A protein.

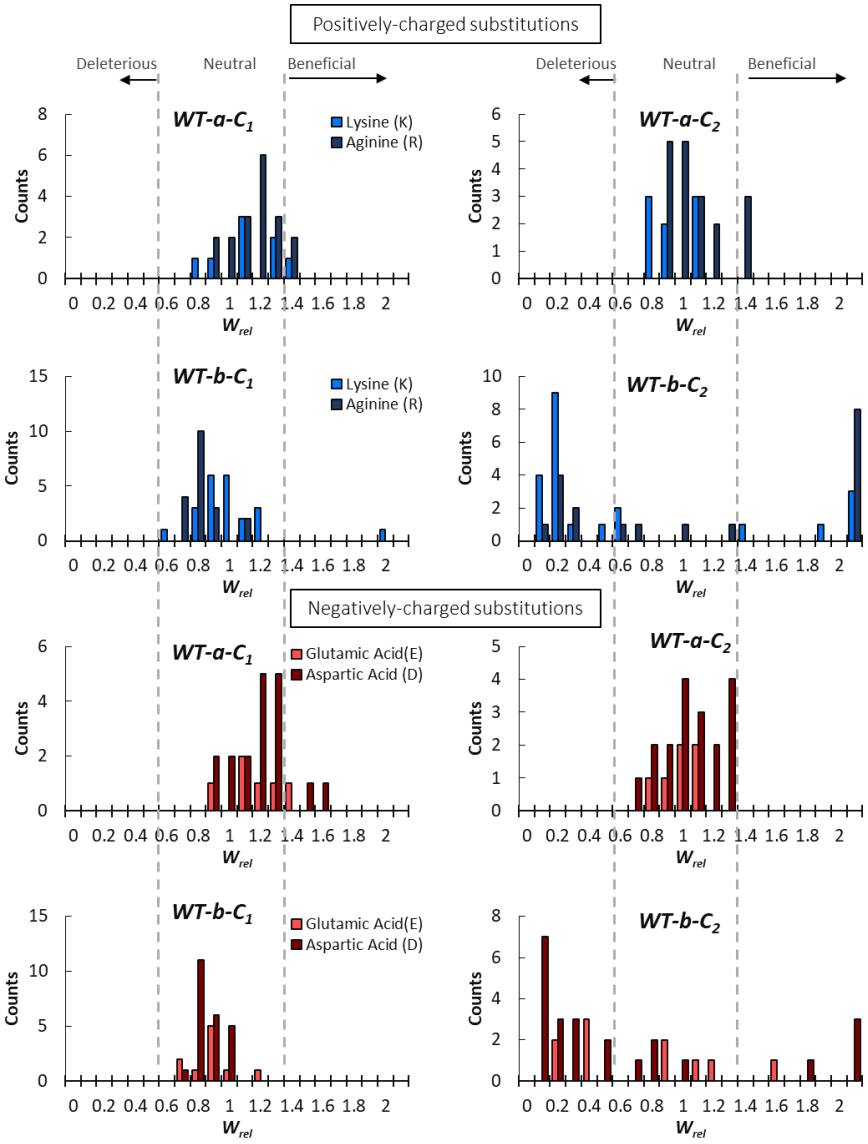


Figure 5-17: Distribution of fitness effects of missense substitutions generating positively- or negatively-charged amino acids on surface residues of PBP-A protein.

Predominantly, missense substitutions on residues located at active site, surface or hydrophobic core of PBP-A, are neutral, slightly beneficial and rarely deleterious. Similarly no significant effect of substitutions towards positively- or negatively-charged amino acids on surface of protein on the fitness is observed. This is probably because the expression level in these conditions is very low (single-copy), and thus all types of mutations are allowed to retain during the

drift. Slight enrichment of mutations in the active site and hydrophobic core may explain their role in reducing fitness cost of PBP-A by compromising the activity. However, we checked for distribution of beneficial missense mutations between the conserved motifs in the active site of PBP-A to see which motifs are more mutable and which are robust against mutations, and the results of this analysis are presented in Figure 5-18 and Figure 5-19.

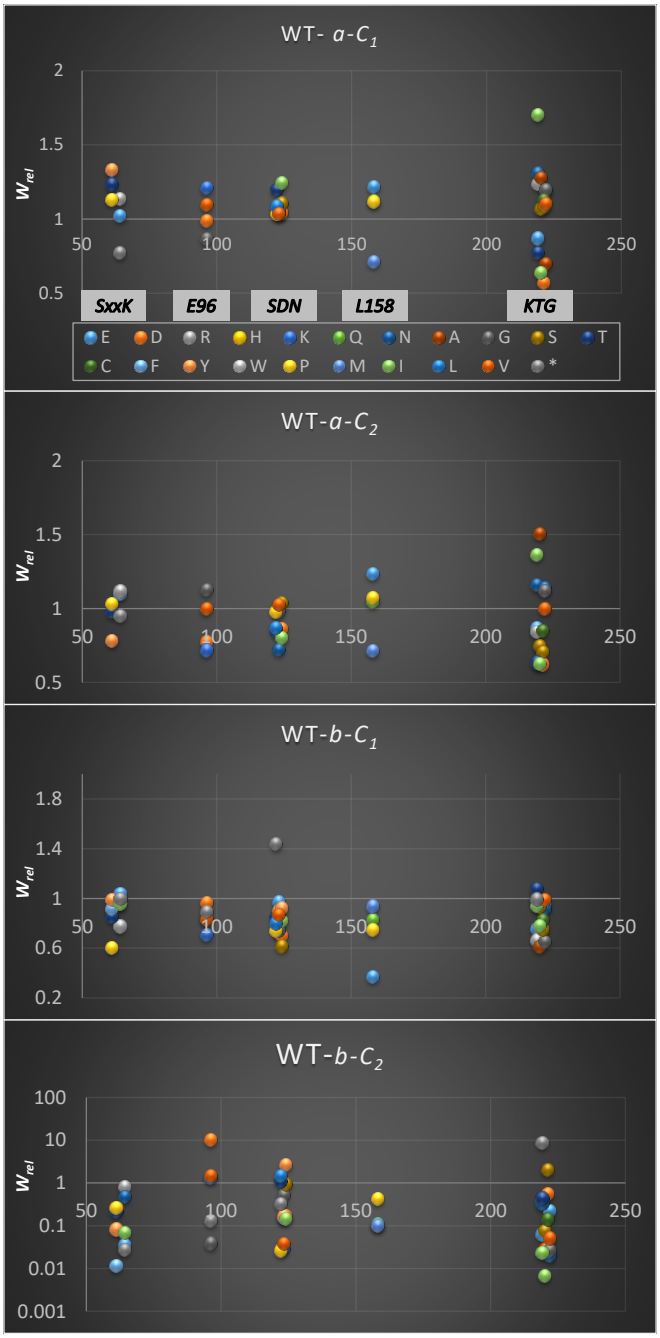


Figure 5-18: Distribution of nonsynonymous substitutions among the active site residues of PBP-A-wt, including S61 and K64 in the SxxK motif, E96 anchoring the active site cavity, S122, D123 and N124 in the SDN motif, L158 in the omega-loop, K219, T220, G221 in the KTG motif and D222 adjacent to the KTG motif. Color code represent the 20 amino acids. Stop codon is shown as *.

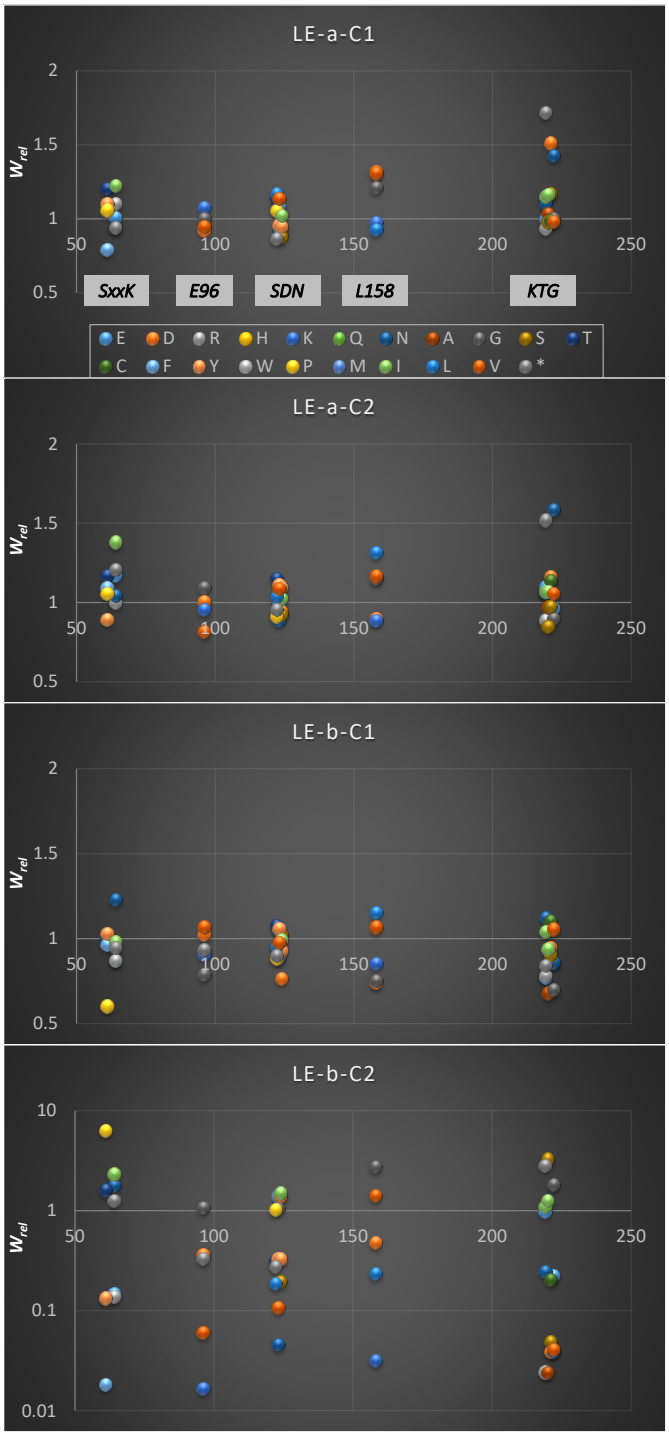


Figure 5-19: Distribution of nonsynonymous substitutions among the active site residues of PBP-A-L158E. Annotations are same as in Figure 5-18.

The results in Figure 5-18 shows that by increase in the expression level (from a-C_1 to b-C_1), majority of the mutations in the active site of PBP-A-wt are purged along the drift, which is due to their destabilizing effect. However there are mutations that are enriched during the drift, which are presented later in this section. This effect is less pronounced in the case of PBP-A-L158E (Figure 5-19). Here, we list the beneficial substitutions in the active site residues under low expression conditions (from a-C_1 to b-C_1), and discuss their probable role in the fitness improvement. In Table **5-10** we present the list of mutations with beneficial relative fitness ($W_{rel} > 1.3$).

Table 5-10: Beneficial mutations in the active site of PBP-A observed in the drift under single-copy expression. Mutation with relative fitness values higher than 1.3 ($W_{rel} > 1.3$) in the conserved active site motif of PBP-A are presented.

	Beneficial mutation in the active site ($W_{rel} > 1.3$)	
	$a - C_1$	$a - C_2$
WT	S61Y (1.33), K219I (1.7), K219N (1.31)	K219I (1.37), T220A (1.5)
LE	E158D (1.31), E158V (1.31), G221D (1.51), D222N (1.4)	K64I (1.4), E158L (1.31), D222N (1.58)

Enriched mutations in the catalytically conserved motifs of PBP-A are mainly on substitution of K219 in the KTG motif for PBP-A-wt, and E158 as well as D222 for PBP-A-L158E. Enrichment of such missense substitutions in the active site of PBP-A suggests that these changes in active site residues are tolerated and are beneficial for the fitness, meaning that either inactive protein (modified active site) has a higher fitness in *E. coli,* or these mutations may only reduce the activity against peptidoglycan of *E. coli,* while keeping the general PBP characteristics. Mutations on Lys219 (to isoleucine or asparagine) and Asp222 (to asparagine) may play a role in recognition of peptidoglycan substrates, thus decrease in deleterious activity. If so, this would become interesting to pre-install these mutations in the scaffold of the enzyme as a new starting point for directed evolution. However, further characterization of these mutants are essential to draw clear conclusions. Changing Glu158 in the drift of PBP-A-L158E libraries

seems to be beneficial, suggesting that this mutant, which has a higher hydrolytic activity, is costly for fitness of the cell.

5.3 Discussion

Failure in directed evolution

The results on directed evolution confirms the counter-intuitive nature of the evolutionary trajectory from PBP-A to a β-lactamase in *E. coli*. Surprisingly, under fitness improving conditions such as low expression of libraries, selection under basic pH and osmo-protective condition, and libraries of disulfide-bonded PBP-A, we did not observe any variant of PBP-A with significant improvement on β-lactamase activity as resistant phenotype in *E. coli*. Even subjecting different variants of *pbp-a* as starting point of directed evolution, from the wild-type (*pbp-a-wt)* and Leu158Glu *(pbp-a-L158E)* to variants with the disulfide bridge *(pbp-a-wt-SS* and *pbp-a-L158E-SS)* as well as variants with a longer Ω-loop *(pbp-a_TEM-1-Ω-loop* and *pbp-a_TEM-1-Ω-loop-KTG_fit)*, did not led to success. However, a combined approach has not been attempted, e.g., directed evolution of PBP-A by combining few or all the fitness-improving criteria.

Nonetheless, these libraries were mainly selected under induced expression of PBP-A, which might not be the best condition for selection of plasmid-borne libraries since we learnt from the drift experiments that induction of plasmid libraries leads to loss of the active variants. In the single-copy chromosomal condition, where the fitness cost in significantly minimized, the reason that no variants was selected could be either due to low expression of PBP-A in the periplasm, hence unable to confer resistance against ampicillin. Although we confirmed that expression of the TEM-1 β-lactamase from the same locus (*rhaA*) in *E. coli* genome had the same MIC of its expression from plasmid, TEM-1 is a very efficient enzyme and a fraction of it in the periplasm might be sufficient to confer resistance, while best variants in PBP-A library might be just slightly improved for β-lactamase, thus require higher fraction in the periplasm.

The most promising results were obtained from the directed evolution of disulfide-bonded PBP-As, where an initial increase of resistance against

ampicillin was detectable, although not stable for further isolation. We hypothesize that the unstable resistant phenotype is because after hydrolysis of the ampicillin, the PBP-A enzymes, with improved hydrolytic activity, became active on the PG and caused lysis of the host *E. coli* cells. However, such increase in MIC have not been observe under directed evolution of other variants of PBP-A, suggesting that the disulfide bond has a role in evolvability. Since the disulfide bond found to be improving the fitness cost, by reducing its DD-peptidase activity, without compromising the reactivity against penicillin (chapter 4), this increase in the resistance suggests that the loss of original activity(ies) (peptidase activity(ies) on muropeptides/PG) is beneficial for generation of β-lactamase activity. Nonetheless, combining directed evolution of disulfide-bonded PBP-A with other fitness-improving factors may lead to a more stable resistant phenotype and increase the evolvability.

Whole genome sequencing of serially passaged ampicillin resistant E. coli

Although the ampicillin resistant phenotype, was not linked to the pBAD43_PPB-A plasmid, the emergence of resistant clones in stressed-passage of *E. coli*, under sublethal concentration of ampicillin, only in cultures that are expressing variants of PBP-A-L158E is still unexplainable. However, it could be due to a special stress induced by this protein.

SNPs upstream of ampC gene

Two SNPs upstream of *ampC* gene were detected at positions -11 (C→T) and +31 (C→A), which are located on box -10 and the attenuator region of the gene's promoter region, respectively (Figure 5-20). In *E. coli*, unlike most other bacteria in Enterobacteriaceae family, the AmpC expresses a non-inducible phenotype, which produce a very low level of AmpC β-lactamase enzyme unable of conferring resistance phenotype (Bajaj, Singh, & Virdi, 2016; Jacoby, 2009; Bengtåke Jaurin, Grundström, Edlund, & Normark, 1981). Many studies reported that various mutations in the promoter/attenuator region of the *ampC* gene causes constitutive overexpression of AmpC in *E. coli* (Caroff, Espaze, Bérard, Richet, & Reynaud, 1999; Caroff, Espaze, Gautreau, Richet, & Reynaud,

2000; Forward et al., 2001; E. C. Nelson & Elisha, 1999; Peter-Getzlaff et al., 2011; Tracz et al., 2005, 2007). Particularly, the mutations detected in our study are located inside the TATA box (box -10) and attenuator region in the *ampC* promoter, which increase their potential role in regulation and overexpression of AmpC β-lactamase in the ampicillin resistant strain. Changing the -10 region from TACAAT to the consensus sequence of TATAAT is reported to result in a seven times stronger promoter of *ampC* in *E. coli* (B Jaurin, Grundström, & Normark, 1982). Interestingly, the exact mutation of (C→T) in Pribnow box of *ampC* promoter wes reported to increase resistance to β-Lactams in *E. coli* (Corvec, Caroff, Espaze, Marraillac, & Reynaud, 2002). Moreover, a mutation (C→G) at the similar position of +31 in attenuator region, beside 4 other mutations, was identified in clinical isolates of *E. coli* with increased resistance to oxyiminocephalosporins (Caroff et al., 1999). These results confirm that the two point substitutions upstream of *ampC* gene are most probably the source of resistant phenotype in the serially passaged clones. However, the role of each mutation in the resistant phenotype remains to be characterized.

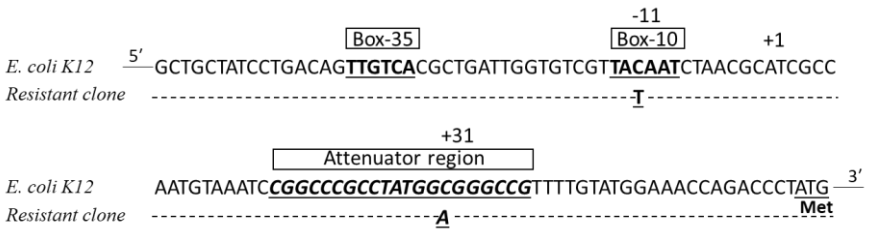


Figure 5-20: Mutations upstream of ampC gene in ampicillin resistant *E. coli* strain. Two SNPs are detected at upstream of ampC gene in *E. coli* after 18 days of serial passage under sublethal concentration of ampicillin up to 1500 μg/ml. The two mutations are detected at positions -11 (C→T) and +31 (C→A) in the box -10 (TATA box) of the promoter and attenuator region of ampC gene, respectively.

*Deletion of 11 nucleotides upstream of mar*R *gene*

In *E. coli*, MarR negatively regulates *marRAB* multiple antibiotic resistance (*mar*) operon that encodes a drug efflux pump, and mutations in proteins that participate in this system lead to a multiple antibiotic resistance phenotype (Anne Grove, 2013). Transcription of the *marRAB* operon is normally auto-repressed by MarR (Cohen, Hächler, & Levy, 1993; Seoane & Levy, 1995). It is widely

known that alteration in control of *marRAB* expression is associated with resistance in bacteria against low-level antibiotics. It has been shown that exposure of *E. coli* cells to external stresses, such as uncoupling agents and redox-cycling compounds, can reverse the MarR repression and trigger response of bacteria (Seoane & Levy, 1995). Another study reported that by exposing *E. coli* cells to low concentrations of tetracycline or chloramphenicol caused multiple-antibiotic-resistant to structurally unrelated antibiotics (George & Levy, 1983).

The deletion of 11 nucleotides upstream of *mar*R gene is located between -10 element (Pribnow box) and the transcription initiation site of *mar*R gene. This variation in length would probably decrease the expression of MarR repressor, thus increasing expression of marA global transcriptional activator protein, which is shown to effect regulation of bacterial drug export systems and cause efflux pump mediated antibiotic resistance, e.g., by inducing expression of AcrAB-TolC MDR efflux complex leading to multidrug resistance in *E. coli* (Chetri et al., 2019; Grkovic, Brown, & Skurray, 2002; Ruiz & Levy, 2010). Since we observed in Chapter 2 that expression of PBP-A weakens the outer membrane of *E. coli* and increase its permeability for antibiotics such as vancomycin and ampicillin, the effect of this deletion on efflux pump can play an epistatic role against the increased (deleterious) influx of ampicillin caused by expression of PBP-A during the experiment. Such regulation of antibiotic influx might provide a beneficial background towards gradual rise of the resistance phenotype. However, the effect of this deletion on expression of MarA gene as well as on susceptibility of cells against antibiotics remains to be characterized.

A missense mutation in ftsI gene

Another mutation observed in the resistant clone, is a missense mutation (Q536L) in *ftsI* gene, known as PBP3 D,D-transpeptidase in *E. coli*, which is known to be involved in cell division and Z-ring formation (Dewar & Dorazi, 2000; Du & Lutkenhaus, 2017). Interestingly, Q536L substitution in PBP3 of *E. coli* is reported to decreasing susceptibility of *E. coli* against LYS228 and aztreonam (ATM) monobactam antibiotics after 20 serial passages in presence of

these antibiotics. It was suggested that Q536L may reduce interaction of PBP3 with LYS228 and ATM (Dean et al., 2018). Analysis of *E. coli* PBP3 structure (PDB: 6I1I) shows that the substitution is in a loop ending to an alpha-helix at the entry of the active site pocket (Figure 5-21) which could affect the shape of active site cavity, substrate (antibiotic) recognition/binding, and/or catalytic dynamics of the enzyme. Escaping ampicillin by PBP3, as one of the targets of this antibiotic, could be a source of decreased susceptibility of the cells against ampicillin. The opposite can also be imagined as if the affinity of PBPs towards ampicillin has increased that can bind and trap more ampicillin in shorter time so that the periplasmic concentration of ampicillin would drop for a fraction of time to buy time for a temporary survival. Such effects on susceptibility might have provided a complementary and favorable condition for other two mutations (upstream of *AmpC* and/or *mar*R genes) to generate a resistance phenotype. However, structural alignment of PBP3 and TEM-1 β-lactamase shows high similarity (Figure 5-21) and the possibility of conversion of PBP3 to a β-lactamase is not far from expectations, especially that we observed a PenG hydrolysis activity with a very low K_M (similar to PBPs rather than β-lactamases) in periplasmic extract of the resistant clone. Unlike active site of TEM-1 β-lactamase (PDB: 1BTL), which is an open cavity, active pocket of PBP3 is roofed by an arch-like electron density made between V344 and Y541, that Y541 belong to the alpha-helix held by the loop than contains Q536L mutation. This arch might be removed or repositioned by introducing a hydrophobic residue (leucine) at the surface of the protein in place of hydrophilic Gln536, which might cause a torsion in the loop for positioning of the side chain of the leucine towards the hydrophobic core. Nevertheless, role of Q536L mutation on activity of PBP3 DD-peptidase against β-lactams, such as ampicillin, remains to be characterized.

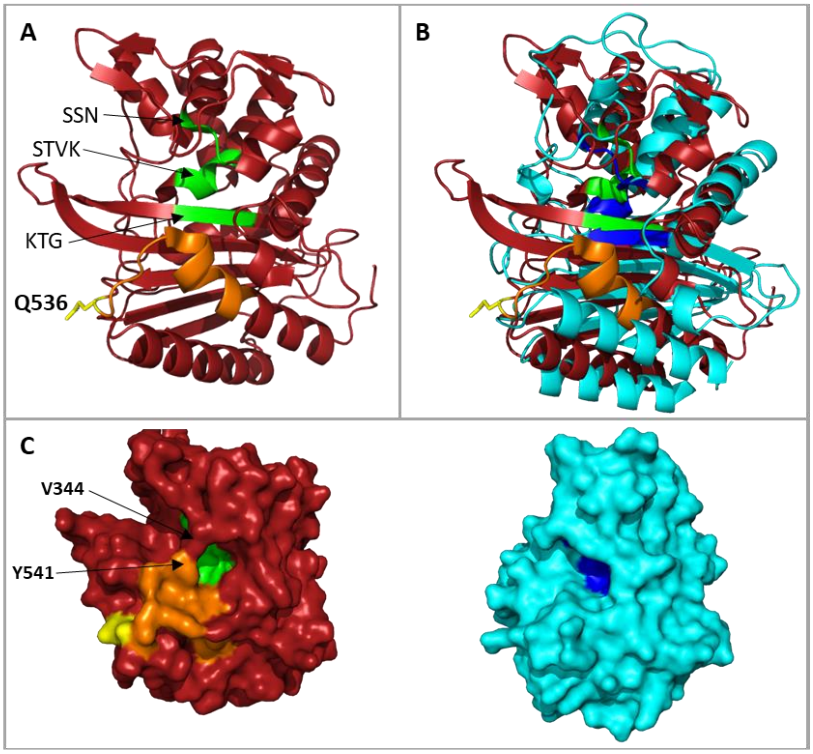


Figure 5-21: Structure of *E. coli* PBP3 and position of Q536L mutation. A) Structure of PBP3 from *E. coli* (PDB: 6I1I) and the three catalytic motifs (SXXK, SXN and KTG, shown in green) conserved among PBPs and SBLs. The Q536 is shown in yellow, while the loop that contain Q536 and its connection to the α-helix that is facing the active site is shown in orange. B) Structural alignment of PBP3 (dark red) and TEM-1 β-lactamase (cyan, PDB: 1BTL). The conserved catalytic motifs are shown in green for PBP3 and in blue for TEM-1. C) Surface view of the PBP3 (left) and TEM-1 (right) structures. Unlike TEM-1, the active site cavity of PBP3 is not fully accessible as an α-helix, which is held by the loop that contains Q536, is shifted towards the cavity and make an arc-like electron density formed between V344 and Y541.

Evolution of ampicillin resistance by mechanisms rather than recruiting an existing PBP-A, that is expressed from plasmid along the serial passage of *E. coli* populations, tells us that either (i) possible active variants of PBP-A are not having any chance to fight against ampicillin probably because of their occupied active site, (ii) there is no active variants in PBP-A libraries capable of hydrolyzing ampicillin, (iii) it requires epistatic mechanisms to compensate the deleterious intermediates along the trajectory. However, the effect of MarR on efflux pumps of *E. coli* could explain a compensatory mechanism against the fitness cost of

PBP-A in *E. coli*, which we found to be alteration in PG profile and increased susceptibility of outer membrane against external agents. Nonetheless, beside the induction of AmpC beta-lactamase, the two compensatory mutation might play an important role in bypassing the fitness cost associated with expression of PBP-A in *E. coli* and restoration of the probable higher influx of antibiotics (due to a damaged cell envelope) to the wild-type level, and thus support the selection of PBP-A against ampicillin. This could be examined by conducting directed evolution experiments of PBP-A in an *E. coli* context that harbor the same deletion upstream of *mar*R gene.

Drift of PBP-A libraries under no selection pressure

In contrast to prolonged drifts under selective pressure (Rockah-Shmuel et al., 2015) that majority of substitutions are not tolerated, our results suggest that under drift without selection pressure the substitutions are predominantly remained neutral during the drift. However, intrinsic fitness costs such as interfering activities and/or protein aggregations, are highly dependent on expression level. For the toxic PBP-A protein in *E. coli,* we observed that fitness costs are tolerated at lower expression levels (conditions *a*-C_1 , *a*-C_2 and *b*-C_1), while at higher expression level (*b*-C_2, induced expression from plasmid) cell would not pay the toll and the most favorable scenario to gain fitness is to lose the expression completely (i.e. the best fit is no protein). This observation suggests that compensatory adaptation of the cell for fitness improvement and/or selection of variants from ep-libraries did not get the chance to play a role in fitness improvement. Such loss of expression is commonly observed in molecular biology when overexpression of toxic proteins is attempted. For instance, attempts for expression of a class A BL, called blaARL, from Gram-positive *Staphylococcus arlettae* in *E. coli* has been failed and resulted in protein loss by survival of clones carrying nonsense mutations in *blaARL* gene (Andreis et al., 2017).

NGS analysis of drifted libraries of PBP-A under single- and low-copy conditions showed that at low expression levels there is no robustness against mutation in

the protein and mutations are evenly acceptable at any position of the protein. The analysis of the active site, the hydrophobic core (buried residues) and the surface as well as analysis by type of missense substitution (e.g. acidic, basic, polar or non-polar) show no trend towards a particular situation. Anyhow, some patterns are starting to emerge and probably if the drift is continued for a longer period, more clear information would be achieved. However, further laborious analysis of these data, together with structural and biochemical interpretations, may provide useful information regarding mutants that could play as good starting points for directed evolution experiments.

Nonetheless, NGS results suggest that during the drift under single-copy expression context, even under induced expression, the diversity of ep-libraries of toxic PBP-A maintains for generations by successfully bypassing the fitness cost by lowering the expression level. Although directed evolution of PBP-A to β-lactamase was not successful, which might be due to low expression against a relatively strong selection that is linked to survival of the cells, chromosomal expression may be an appropriate system, rather than plasmid system, for directed evolution of toxic proteins that do not rely on couterselection (i.e., their selection is not dependent on survival of the clones, e.g., screening by activity assays).

5.4 Conclusions

Surprisingly, directed evolution of PBP-A to β-lactamase was not successful under fitness improving conditions suggesting that either combined approaches are essential or order of implementation of these conditions is important. However, we described a scarless genome-editing method for construction of chromosomally-encoded ep-libraries (up to 10^6 individual variants) in *E. coli*.

The results of drift under sub-lethal concentration of ampicillin, and in presence of PBP-A library suggest that under these conditions the *E. coli* cells can circumvent the antibiotic stress by induction of AmpC β-lactamase expression. However, this might require series of mechanisms that are involved in alteration of the initial susceptibility against ampicillin, such as changes in in the efflux of

antibiotics (mutation on MarR) or decrease in susceptibility against the antibiotics by mutations in the endogenous PBPs of *E. coli*. However, role of these two mutations in the emergence of resistance by AmpC remains to be investigated. This role can also be examined for directed evolution of PBP-A.

Our results highlight the importance of expression level in evolution of toxic genes to not face any deleterious consequences such as loss of gene or loss of expression. Therefore, we suggest that chromosomal libraries could serve as a neutral evolutionary context and as an alternative for plasmid libraries in directed evolution of toxic genes in bacteria.

5.5 Contribution of authors

S.A. greatly contributed in the bioinformatic analyses; The rest of the works including the experiments, development of theory, discussions on data and preparation of manuscript are carried out by G.M.D under supervision of P.S.

5.6 Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programm under the Marie Skłodowska-Curie grant agreement No 722610 (ES-Cat).

5.7 Materials and methods

A desired mutation rate can be achieved by calibrating ep-PCR conditions as recommended by the manufacturer. We optimized the ep-PCR for a low mutation load on PBP-A (0-6 mutations per gene) by adjusting the number of epPCR cycles and the amount of initial template to 20–25 cycles and 100 ng template, respectively. See also the user manual for the "Random Mutagenesis Kit" from Agilent.

5.7.1 Genome engineering and chromosomal library construction

We have already published the detailed protocol for 'Building Scarless Gene Libraries in the Chromosome of Bacteria', which is attached at the end of general methodology chapter and available online: https://doi.org/10.1007/978-1-0716-0720-6_11. Briefly, *php-a* gene comprising *DsbA* signal sequence and 6×His

tag fused at the amino- and carboxy-terminal, respectively, was replaced with *rhaA* gene at the second position of the polycistronic operon of rhamnose in the genome of *E. coli* using CRISPR-Cas9 technology and λ-Red recombineering system (see Figure 5-2).

5.7.2 Drift of pbp-a libraries under no ampicillin selection

The libraries of PBP-A-wt and PBP-A-L158E were subjected to serial passage under no selection pressure for PBP-A activity under (a) single- and (b) low-copy expression for ~50 generations (5 days) at a dilution of 1:1000. To avoid bottlenecking along the drift, we started the drift by inoculating the pooled ep-library and serial passaging (every 24 hours, ~10 generations) of at least 100 times more cells (≥1.2×10^8 cells) than the initial libraries, which was ~1×10^6 individual clones for chromosomal, and ~2×10^6 clones for plasmid libraries (Kram et al., 2017). The drift was carried out under two expression conditions for each library: C_1, basal (non-induced or leaky) and C_2, induced PBP-A expression. For induced condition either L-rhamnose or L-arabinose was added to the growth medium for chromosomal or plasmid libraries, respectively.

5.7.3 Western blot analysis of PBP-A expression in drifted samples

To compare expression level of PBP-A in different time points during the drift, 200 µl of the frozen libraries at each time point were inoculated into 10 ml LB media and grown at 37 °C under 180 rpm agitation. Once samples reached early exponential phase (OD_{600} ~ 0.2), PBP-A expression was induced by addition of 0.5 % (w/v) L-arabinose followed by growing for up to 4 hours. Cultures were normalized to the same OD_{600} prior to harvesting by centrifugation, followed by periplasmic protein extraction via diluting higher density cultures by LB media.

5.7.4 Preparation of DNA for NGS sequencing

For NGS sequencing, 2 ml of pooled cells of the initial ep-libraries (t0) as well as the cultures from last day (t5) of the drifts, were subjected to genomic or plasmid DNA extraction for single- and low-copy libraries, respectively. All the pooled libraries and cultures had OD_{600} of ≥1.5, which corresponds to ≥1.2×10^9 cells. From chromosomal and plasmid libraries, a window of 1094 bp and 1070 bp,

respectively, was amplified with primer pairs of A and B. These amplicons contain extra bases (70−115 bp) upstream and downstream of DsbA_*php-A*_His to not leave the target gene in extremities of amplicon (the sequence of full amplicons are available in Appendix 9-25Appendix 9-26. The amplicons were purified, quantified and sent for Ion Torrent or Illumina sequencing. To avoid bottleneck effect on diversity of libraries during the amplification of *php-a* genes, 100 ng of the extracted *E. coli/rhaA::php-a* genome for single-copy libraries, and 20 ng of the extracted pBAD43_*php-a* plasmids for low-copy libraries were used as PCR templates. (i.e., ~2×10^7 molecules of double stranded DNA of genome, and ~1.3×10^{10} molecules of plasmid). The PCR was performed using Q5® High-Fidelity DNA Polymerase.

Primer pair A (for amplification of chromosomal libraries):

chr_lib_NGSseq_fw: CCACCTTTACCCCTAATCC

chr_lib_NGSseq _rv: ACTGGCGAATGCTGTCGT

Primer pair B (for amplification of chromosomal libraries):

plasmid_lib_NGSseq_fw: CATAGCATTTTTATCCATAAGATTAGCGG

plasmid_lib NGSseq_rv: TCTCATCCGCCAAAACAGCC

5.7.5 Alignment and Mapping to template

The high quality reads, that obtained after trimming and filtering the raw reads, were locally aligned to the *DsbA_php-a_6×His* sequence as template using the Smith-Waterman algorithm with the parameters given in Table 5-11. Alignment score (S) of 30 was applied to avoid poorly aligned sequences followed by discarding reads with aligned region of less than 30 bp.

Table 5-11: Parameters of Smith-Waterman algorithm used for alignment of NGS data.

Parameter	Value
match_score	1
mismatch_score	-2
target_open_gap_score	-5
target_extend_gap_score	-5
target_left_open_gap_score	-5
target_left_extend_gap_score	-5

target_right_open_gap_score	-5
target_right_extend_gap_score	-5
query_open_gap_score	-5
query_extend_gap_score	-5
query_left_open_gap_score	-5
query_left_extend_gap_score	-5
query_right_open_gap_score	-5
query_right_extend_gap_score	-5
Mode	local

5.7.6 Serial passage of *E. coli*/pBAD43_PBP-A-L158E cells stressed under sublethal concentration of ampicillin

Library of PBP-A-L158E and PBP-A-wt were subjected to selection on LB media containing 0.5% L-arabinose (w/v) and ampicillin gradient strip (0.016-256 μg/mL). To benefit from stress response and/or spontaneous mutation of bacteria, the cells on the edge of the inhibition zone are collected, mixed gently and re-subjected to MIC evaluation by ampicillin strips on LB plates containing 0.5% L-arabinose (w/v). Each day the diameter of the zones of bacterial inhibition around the antibiotic strip became smaller for cells expressing PBP-A-L158E library but not PBP-A-wt. Once these cultures become resistant to more ampicillin than the limit of the strip (more than 256 μg/mL), the passage experiment was continued on gradient plates of ampicillin up to 500 mg/L and the top resistant layer were collected gently mixed and re-plated on gradient plates, followed by gradient plates up to 1000 and 2000 mg/L (see Figure 5-5).

5.7.7 Whole genome sequencing of resistant strain of *E. coli* Top10

Culture from serial passage of the 18th day was streaked onto solid medium and an isolated colony from each experiment were subjected for genomic DNA extraction. Genomic DNA of resistant strain and wild-type *E. coli* were purified using QIAGEN genomic DNA extraction kit according to manufacturer's specifications. Samples were sent to GENEWIZ (Leipzig, Germany) for whole genome sequencing (WGS) using Illumina MiSeq (2x150bp) next generation sequencing. Standard data analysis was carried out by GENEWIZ, where they asked to use *E. coli* K-12, also called *DH10B*, (Accession: NC_010473.1) genome

as reference (fasta sequence available at https://www.ncbi.nlm.nih.gov/nuccore/NC_010473.1?report=fasta).
According to the reference map, single nucleotide polymorphisms (SNP) and InDel variants of samples were compared to each other.

Chapter 6

General Discussion and Future Perspectives

The evolution of proteins is mainly regarded as optimization **for** a specific task only, while it is rarely associated with optimization **against** deleterious activities. Therefore, it is important to revise our general perception in this regard. Natural evolution employs different mechanisms to optimize proteins "against" unwanted deleterious activities, either directly on the protein itself or indirectly by involving other protagonists. In direct ways, mutations can be selected to control promiscuity and reduce as much as possible interfering activities. Mutations affecting the active site may prevent the accommodation or turnover of unwanted substrates. This may be easy for some enzymes, while very difficult for others especially when there is a need to discriminate between very similar substrates such as a penicillin and a D-Alanyl-D-Alanine stem peptide. Besides, mutations changing the physicochemical properties of the protein, such as its charge, its size or its diffusion freedom (i.e. membrane anchorage) may also prevent the protein to get access to unwanted substrates by electrostatic repulsion or physical separation. This may be particularly beneficial to avoid polymeric and polyelectrolyte substrate. On the other hand, molecular evolution

may also rely on indirect mechanisms to prevent or minimize deleterious effects. Quite obviously, gene regulation may be used to avoid the expression of a protein under conditions when its activity may be deleterious. A deleterious/interfering effect may eventually be controlled by tuning the expression level and it is unknown how many enzymes would become toxic if constitutively expressed. In case of beta-lactamase, the presence of penicillin itself may eventually reduce deleterious activity by competing with the active site and reducing access to unwanted substrates. In this context, induction of beta-lactamase only in the presence of penicillin may have a protective function. Another way to minimize damages or production of toxic compounds is to evolve repair systems. Indeed, recent literature have highlighted widespread metabolic repair enzymes, indicating that, in many cases, repair may be easier than prevention.

In summary, natural selection that optimize the enzymes towards their specific and necessary function but against any deleterious and interfering activities, may proceed through various intra- and intermolecular mechanisms, leading to very tortuous trajectories and involving changes in multiple properties of a protein. Some of these changes may be considered as secondary in regard to the active site and catalytic mechanism, even though they might be crucial.

The failure in directed evolution attempts, suggests that mechanisms to optimize "against" PBP-A activity are essential to unlock evolvability of this enzyme. LMW-PBPs are considered as ancestors of class A serine beta-lactamase enzymes. PBPs are mainly membrane-anchored enzymes (Figure 6-1). As beta-lactamases are soluble proteins in periplasm, we simply removed the membrane-anchorage domain in PBP-A enzyme to only keep the peptidase domain for the directed evolution towards a beta-lactamase. It is possible however that membrane-anchorage played as a mechanism "against" costly activities of PBPs. On the other hand, beta-lactamases such as TEM-1 have acidic pI and possess a disulfide bond in their structure. Fitness improvements associated with lowering PBP-A's pI and introducing a disulfide bond indicate that such changes in properties of this protein may as well secure the enzyme against deleterious activities. Additionally, the presence of beta-lactam substrate may also play a very

important role by protecting the peptidoglycan "against" the deleterious activity(ies) of PBP-A, as suggested by the loss of resistance phenotype in the library of PBP-A-SS.

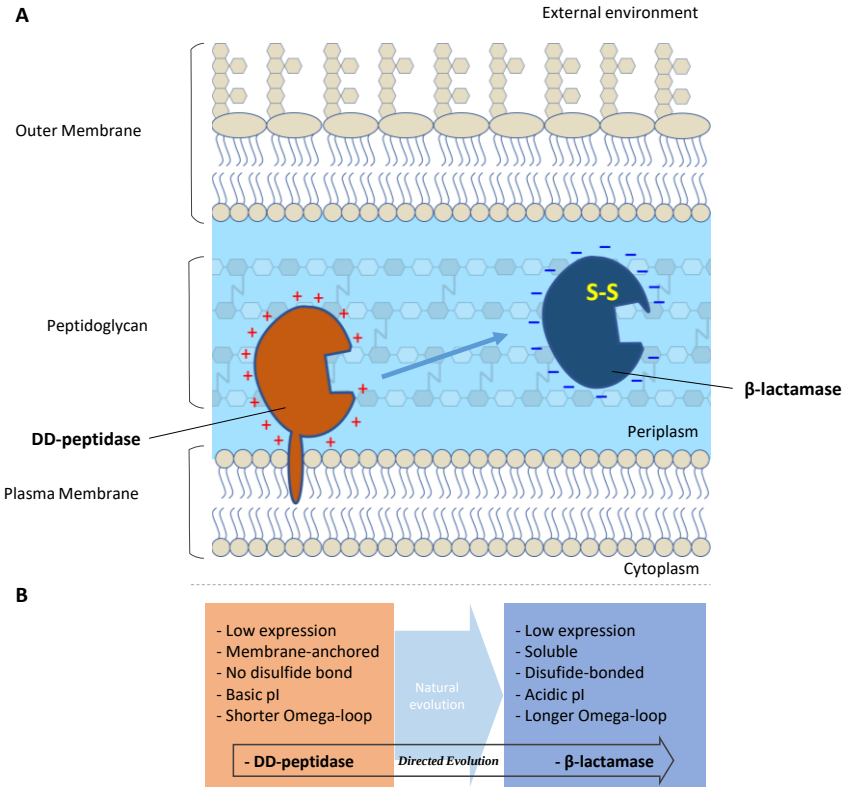


Figure 6-1: Schematic overview of evolutionary emergence of class A beta-lactamases from PBPs. **A)** In periplasm of Gram-negative bacteria, PBPs are mainly membrane anchored, while beta-lactamases are soluble enzymes. Class A beta-lactamases such as TEM-1 possess acidic pI and a disulfide bond while PBP-A, has a basic pI without any disulfide bond in the structural fold. **B)** General comparison of natural evolution versus the focus of directed evolution complain, in emergence of beta-lactamases from PBPs.

Under fitness improving conditions, directed evolution of PBP-A to beta-lactamase was, unexpectedly, not successful suggesting that either (i) combined approaches are essential, (ii) order of implementation of these conditions is important or (iii) mechanism of genetic diversification other than (or combined with) point mutations are required to access the new phenotype. For example, regarding the respective order of beneficial conditions, PBP-A may have to be

maintained as membrane-anchored to gain hydrolytic power against beta-lactams while losing its activity on peptide substrates, and only then be released as soluble.

As a perspective, we propose to put together all the elements that are improving the fitness and attempt the conversion to a beta-lactamase. Site directed mutations would be introduce to lower the pI and introduce the disulfide bond. Beneficial mutations identified in the drift experiment will be introduce only if the reactivity towards penicillin is not impaired. Then, the library will be generated and the selection for ampicillin resistance will be performed in a pH-buffered medium containing osmoprotectants. Eventually, a library of membrane anchored PBP-A would be tested as well but the impact of anchoring a beta-lactamase on ampicillin resistance will have to be evaluated first.

Chapter 7

General methodology

7.1 General methods for Chapters 2-5

7.1.1 Synthetic genes

pbp-a

The *pbp-a* gene is purchased as a gblock (Integrated DNA Technologies (IDT), Inc) based on sequence reported by Urbach et al. (2008) and optimized for *E. coli* codon usage by introducing silent mutations. A signal sequence from DsbA protein (*DsbAss*), which facilitates periplasmic export using the signal recognition particle (SRP), is fused at N-terminal of *pbp-a*. A poly-histidine tag (6 x His) was fused to the C-terminal of the protein for purification of enzymes on nickel-affinity column and its subcellular expression detection by western blotting (Figure 7-1).

pbp-a_TEM-1-Ω-loop

Chimeric PBP-A_TEM-1-Ω-loop is constructed by replacing Ω-loop of PBP-A protein with the Ω-loop of TEM-1 beta lactamase by exchange of 13 amino acids from PBP-A, residues 153 to 165 (VINNPLPDMKGTN, Ω-loop of PBP-A is underlined), with 19 AA from TEM-1, residues 159 to 177

(RLDRWEPELNEAIPNDERD, Ω-loop of TEM-1 is underlined) and the corresponding DNA sequence is purchased as a gblock (IDT, Inc).

pbp-a_TEM-1-Ω-loop-fit

To avoid possible steric clashes between the longer loop and the floor of active site (adjacent to KTG motif) in *pbp-a_TEM-1-Ω-loop*, 6 residues of PBP-A, 222 to 227 (DIGIVV), are replaced with 7 residues from TEM-1 beta lactamase, 235 to 241 (AGERGSR) and the corresponding DNA sequence is purchased as a gblock (IDT).

Figure 7-1: Nucleotide and amino acid sequence of DsbA_*pbp-a*_His construct.

7.1.2 Strains and plasmids

Escherichia coli strain Top10 was purchased from Invitrogen, which has a genotype of F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-.

pBAD43 plasmid (Figure 7-2-A) is a low-copy (~5 copies/cell) expression vector (Thompson et al., 2018), which is stringently replicated under pSC101origin (T. S. Lee et al., 2011). It carries tightly regulated pBAD expression system under araBAD promoter and AraC and catabolite activator protein (CAP)-cAMP regulators from araBAD (arabinose) operon (Guzman et al., 1995). The expression of gene of interest (GOI) is tightly repressed in the presence of glucose and absence of L-arabinose and highly induced in absence of glucose and presence of L-arabinose. pBad43 plasmid was kindly provided by Michael Deghelt from Institut de Duve, UCLouvain, Brussels.

pBad plasmid (Figure **7-2**) is a low copy (15-20 copies/cell) (Boros, Pósfai, & Venetianer, 1984) vector under pMB1 (pBR322) origin and like pBAD43, it carries pBAD expression system.

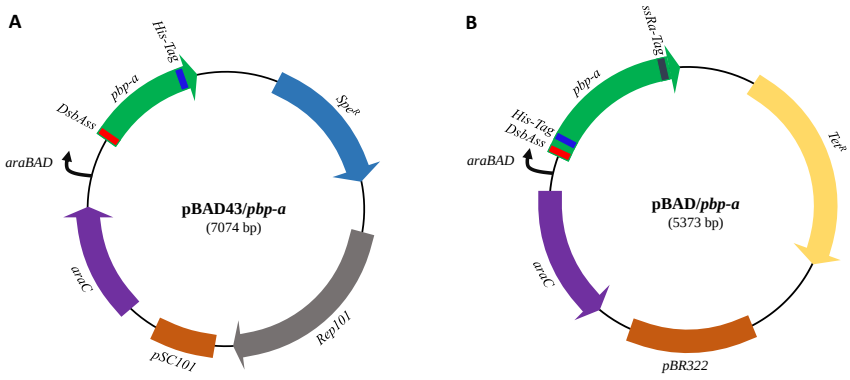


Figure 7-2 : Schematic illustration of pBAD43 and pBAD expression plasmids harboring *pbp-a* gene. The *pbp-a* gene was cloned under araBAD promoter. **(A)** pSC101 replication origin provides a very low copy number (~5 copies per cell) while, **(B)** pBR322 is a medium copy number origin. Selection markers of the vectors are spectinomycin and tetracycline resistant genes for pBAD43 and pBAD, repectively.

7.1.3 Cloning *pbp-a* variants into pBAD43 plasmid

Both of *pbp-a-wt* and *pbp-a-L158E* were amplified from *RhaA* locus from a previously engineered genome of *E. coli* Top10 (Chapter 3) and assembled into pBAD43 plasmid thanks to the overlapping sequences integrated in primers followed by Gibson assembly. Following primer pairs were used for PCR amplification of *pbp-a* gene variants and pBAD43 plasmid backbone. Overlapping sequences, upstream and downstream of *pbp-a* gene, are indicated as underlined and bold, respectively:

Pair 1:

pbp-a fw: ttttagcgtttagcgcatcggcggc

pbp-a rv: **ttggctttcgcccattcaactcagtg**atgatgatgatgg

Pair 2:

pBAD-BackBone-PBPs fw: **cactgagttgaatgggcgaaagccaa**tctagagtcgacctgcaggc

pBAD-BackBone-PBPs rv: cgatgcgctaaacgctaaaactaaaccagccagcgccag

The PCR mix was as follows: 10 µL of 5x Q5 reaction buffer, 0.7 µL of 20 mM dNTP mix, 2.5 µL of 10 µM forward and reverse primers mix, 1 U Q5 High-Fidelity DNA Polymerase (NEB), 50-100 ng of template DNA and nuclease-free H_2O up to 50 µL. PCR thermocycler conditions was set as follows: initial denaturation at 98 °C for 1 min; 30 cycles denaturation at 98 °C for 20 s; annealing 30 s at T = Tm – 5 °C; extension 30 s for 1-kb templates at 72 °C; a final extension for 10 min at 72 °C and hold at 4 °C. If a single band of the PCR products achieved, the products were then purified using Monarch PCR & DNA Cleanup Kit (New England Biolabs), and if not, the expected band was purified from agarose gel using the Monarch Gel Extraction Kit (New England Biolabs) prior to assembly. The purified PCR products were inserted into pBAD43 plasmid backbone by mixing 15 µL of Gibson Master Mix (see details for Gibson assembly in Chapter 8, general methodology), together with 50-100 ng of the linear vector backbone with a 2-3 folds molar excess of the insert, 0.5 µL of DpnI (to digest the remaining original (methylated) plasmid template), and ddH_2O up to 20 µL. The assembly mix is then incubated for 1 h at 50 °C and the assembled fragments (circularized plasmid) can be detected on agarose gel. One µL of the

assembly mix is electroporated into 50 µL of *E. coli* Top10 electrocompetent cells (Invitrogen). Transformants are incubated at 37 °C, for 1 hour before plating all the cells on LB plates containing spectinomycin and glucose (for repression of PBP-A expression) followed by overnight incubation 37 °C. Insertion of *pbp-a* genes in pBAD43 plasmid was verified by colony PCR and confirmed by sequencing of the PCR products from random clones.

7.1.4 In silico analysis tools

Analysis, alignment and all manipulation of DNA and protein sequences were performed either by Benchling or CLC Main Workbench tools.

Tm of each pair of primers was estimated by NEB Tm Calculator (https://tmcalculator.neb. com/#!/main)

7.1.5 DNA manipulation and analysis

Agarose gel electrophoresis

The DNA samples were diluted in 5x loading dye (Thermo Scientific) and then separated by migration on 0,8 % agarose gel (w/v) running under 120 V and 70 mA in TAE 1x buffer (40 mM Tris - 20 mM Acetate – 1 mM EDTA). The gel placed in ethidium bromide bath and incubated for 20-40 (final concentration of ~1.3 mg/ml) and the visualization and imaging under UV light were carried out in Bio-rad Gel Doc XR System (Bio-rad).

DNA quantification

Quantity and quality of DNA samples were estimated using Nanodrop fluorospectrometer (Thermo Scientific) and, if necessary, followed by agarose gel electrophoresis.

7.1.6 Preparation of *E. coli* electro-competent cells

From a single colony on a freshly prepared plate, a pre-culture is started into 5 ml of LB media with appropriate antibiotics in a 50 ml falcon tube and incubated for overnight at 37°C under agitation of 180 rpm. Next day, LB medium supplemented with appropriate antibiotics was inoculated with 1:1000 dilution of the overnight pre-culture and grown at 37°C and 180 rpm agitation until the optical density at 600 nm (OD600) was reached to 0.5 – 0.7. At this exponential growth phase, the culture was cooled down by placing in ice bath for 15 minutes.

All the subsequent steps were carried out in cold and solutions and centrifuge machines were pre-chilled. The cells are harvested by centrifugation at 4000 rpm, 4°C for 10 minutes. Cells were washed 3 times with sterile MilliQ water of ¾ volume of initial culture followed by one time washing with sterile 10% glycerol (v/v) of 1/10 volume of initial culture, after each wash cells are centrifuged at 4000 rpm, 4°C for 8 minutes. The pellet from last centrifugation was suspended in 0.5-1 ml of sterile 10% glycerol and divided into 50 or 100 µl aliquots that were either directly used for transformation or immediately stored by quick freezing in liquid nitrogen and placing at -80°C. Efficiency of electrocompetent cells is assessed by electroporating 10 pg of supercoiled pUC19 plasmid (Invitrogen) followed by serial dilution on LB agar plates containing 100 µg/ml ampicillin. After overnight incubation at 37°C, the colonies are counted and the competence efficiency was estimated as the number of colonies per µg of electroporated DNA.

7.1.7 DNA transformation

All transformations were performed by electroporation using a Bio-Rad Gene Pulser Xcell™ electroporator together with 2 mm electroporation cuvettes.

7.1.8 Site directed mutagenesis

All the site directed mutagenesis are introduced according to QuickLib protocol (Galka et al., 2017). Briefly, target mutations are incorporated into primers and introduced into gene of interest by a single PCR amplification of the gene together with harboring plasmid. The PCR product of linearized plasmid+gene is purified and then circularized using Gibson assembly method, thanks to overlap sequences incorporated into primers, followed by transformation into *E. coli* competent cells.

7.1.9 Library construction

Chromosomal library: Detailed protocol for construction of chromosomal library is available in section 7.2.

Plasmid library: *pbp-a* gene or its variants cloned into pBAD43 plasmid were used as a template for epPCR. GeneMorph II Random Mutagenesis Kit was used to

mutate the *pbp-a* genes and the mutation rate was controlled by adjusting the number of ep-PCR cycles and the amount of initial template as suggested by the user manual of the kit. In order to achieve a low mutation frequency (between 0 and 5 mutations/kb), 100 ng template DNA and 20–25 cycles were applied. The ep-PCR mix was as follows: 5 µL of 10x Mutazyme II reaction buffer, 1 µL of 40 mM dNTP mix (provided in kit), 0.5 µL of primer mix (250 ng/µL of each primer), 2.5 U of Mutazyme II DNA polymerase, 100 ng of template DNA and nuclease-free H2O up to 50 µL. PCR thermocycler conditions was set as follows: initial denaturation at 95 °C for 2 min; 20–25 cycles (depending on desired mutation rate); denaturation at 95 °C for 30 s; annealing 30 s at T = Tm – 5 °C (Tm of each pair of primers was estimated by NEB Tm Calculator (https://tmcalculator.neb. com/#!/main); extension 1 min (≤1-kb targets) or 1 min/kb (>1- kb targets) at 72 °C; a final extension for 10 min at 72 °C and hold at 4 °C. If a single band of the PCR products achieved, the products were then purified using Monarch PCR & DNA Cleanup Kit (New England Biolabs), and if not, the expected band was purified from agarose gel using the Monarch Gel Extraction Kit (New England Biolabs) or an equivalent kit prior to transformation. The purified ep-PCR products were inserted into pBAD43 plasmid backbone by mixing 15 µL of Gibson Master Mix (100 mM Tris–HCl pH 7.5, 10 mM $MgCl_2$, 0.2 mM of each four dNTPs, 1 mM NAD+, 15% (w/v) PEG-8000, T5 exonuclease (2 U/mL), Phusion DNA polymerase (33 U/mL), and Taq DNA ligase (1666 U/mL) with 50-100 ng of vector with a 2-3 folds molar excess of the insert, 0.5 µL of DpnI (to digest the original (methylated) plasmid template), and ddH2O up to 20 µL. The assembly mix is then incubated for 1 h at 50 °C and the assembled fragments (circularized plasmid) can be detected on agarose gel. One µL of the assembly mix is electroporated into 50 µL of *E. coli* Top10 electrocompetent cells (Invitrogen). Depending on desired size of library, number of transformations can vary: for each transformation up to 5 x 10^4 individual clones can be achieved. Transformants are incubated at 37 °C, for 1 hour before plating all the cells on LB plates containing spectinomycin and glucose (for repression of expression) followed by overnight incubation 37

°C. Library's size were estimated by either counting individual transformants or counting the number of CFU/ml. Insertion of *php-a* genes in empty pBAD43 plasmid was verified by colony PCR (primers pair A) and the mutation rate was estimated by sequencing of the PCR products from random clones. Libraries were recovered and pooled by scraping colonies from the plates with LB-media and, for long-term storage, divided into aliquots of 100-250 µl and stored at -80 °C.

Following primer pairs were used in ep-PCR on *php-a* genes and amplification of plasmid backbone. Overlapping sequences, upstream and downstream of *php-a* gene, are indicated as underlined and bold, respectively:

ep-lib_ *php-a*-fw: ttttagcgtttagcgcatcggcggc

ep-lib_php-a-rv: **ttggctttcgcccattcaactcagtg**atgatgatgatgg

pBAD-BackBone-PBPs fw: **cactgagttgaatgggcgaaagccaa**tctagagtcgacctgcaggc

pBAD-BackBone-PBPs rv: cgatgcgctaaacgctaaaactaaaccagccagcgccag

7.1.10 *In-vivo* selection

Libraries were both, directly subjected to *in-vivo* selection on LB agar plates containing 0.5% (w/v) L-rhanmose (for chromosomal libraries) or 0.5% (w/v) L-arabinose (for plasmid libraries) and gradient of ampicillin (Figure 7-3) and, in parallel, prior to selection on gradient plates, 200 µl of the pooled libraries were inoculated into 20 ml of LB media followed by induction of PBP-A expression at OD_{600} 0.3-0.4 and continuing incubation for 1.5 hours at 37 °C under agitation. To compare the minimum inhibitory concentration (MIC) of top variants in library, colonies that were growing at higher resistant on gradient plate were picked, grown and induced in liquid media followed by inoculating on LB agar plates containing L-arabinose, applying a MIC Test Strip of ampicillin (bioMérieux) on top and incubating overnight at 37 °C. For successive rounds of directed evolution, a layer of highest resistant colonies were scraped and pooled from the gradient plates and the template for ep-PCR is prepared either by amplifying the *php-a* gene from *rhaA* locus from chromosomal libraries using

primers outside the gene (primer pair B) or by pBAD43_*pbp-a*(s) plasmid extraction from plasmid libraries.

7.1.11 Preparation of gradient plates

Gradient plate of ampicillin (0 to 100% of a desired concentration) can be prepared thanks for the diffusible property of ampicillin and hardness of LB agar. In sterile environment, a sterile (squared) petri dish was put in a slope position simply by placing the plate on its lid (or something else) in such a way that provided a constant slope for all over the area (Figure 7-3-A). Thirty ml of warm LB agar containing desired concentration of antibiotic and inducer (L-rhamnose or L-arabinose) was poured to form a triangle from lowest to highest edge of the plate and left at room temperature to solidify. Then the plate was placed horizontally and same volume of warm LB medium without antibiotic but having inducer, was poured to cover all the surface area (Figure 7-3-B). After solidifying, the gradient plate was left at least 6-8 hours (typically overnight) to let the ampicillin to equilibrate its diffusion.

For selection, 200 µl the libraries was placed at lowest gradient part of the plate and streaked all over the plate, this step is important as we do not know the precise concentration on each position in the gradient plate as well as the MIC of variants in our library, and, therefore, placing the cells at higher concentration of ampicillin will lead to death of non-resistant bacteria in the library before streaking all over the gradient plate.

Figure 7-3 : Preparation of gradient antibiotic plate. A: A sterile petri dish is placed slopped and media containing antibiotic is poured to solidify. B: the plate is placed horizontally and media without antibiotic is added.

7.1.12 MIC evaluation

E-TEST M.I.C test strips (bioMérieux) consists of a predefined exponential gradient of antibiotics on a plastic strip that diffuses into the agar media. The ampicillin strips provide an accurate gradient from 0.015 µg/ml to 256 µg/ml of

ampicillin in agar plates containing, if necessary, additional antibiotics, inducer or repressor. From a single colony (or aliquots of a library) a bacterial culture started in liquid media under appropriate conditions at 37°C, 180 rpm of agitation. At OD_{600} ~0.5-0.6, cells were inoculated all over the agar plate using a sterile cotton swab and then the strip was applied on the preinoculated agar and incubated overnight at 37 °C. The MIC was defined from the concentration scale on the strip, where the ellipse zone intersected the strip.

7.1.13 Fitness competition assay

To compare fitness between different clones, growing competition for a few generations in LB-media under basal and induced condition were carried out. From a single colony on a recently inoculated plate, pre-cultures for each strain are started into 5 ml of LB media containing appropriate antibiotics in a 50 ml falcon tube and incubated for overnight at 37°C under agitation of 180 rpm. A fresh culture is started for each strain under repression conditions (0.5% glucose) and growth to reach exponential phase (OD_{600} ~ 0.5-0.6). To start each competition, the cultures for both competitors are normalized for the same OD_{600} and mixed at a ratio of 1:1 and 100 µl of the mix was inoculated into 100 ml LB medium containing appropriate antibiotic and inducer or repressor. The culture was incubated at 37 °C under 180 rpm agitation and after 24 hours, 100 µl of the cultures we inoculated into fresh media with similar conditions. Competition assay was conducted for 4 days by repeating the inoculation after every 24 hours. From the remaining volume of cultures on each day, plasmid was extracted and used as template for PCR to amplify the region containing the gene of interest, followed by sequencing. Winner of the each competitions are determined by analyzing sequencing data.

7.1.14 Periplasmic extraction

E. coli cultures (10 ml) were harvested at 4400xg for 10 min, supernatants were discarded and cells were resuspended in 0.33 mL of sucrose 20% (w/v) in 20 mM tris-HCl, pH 8, and 33 µl of EDTA 0.1 M, pH 8 was added. Samples was incubated at room temperature for 10 min followed by 10 min centrifugation of

cells at 6 000 rpm. Supernatants were discarded and the pellets subjected to osmotic shock by resuspending in 0.5 mL of 5 mM $MgSO_4$. The mix was incubated at 4 °C for 10 min followed by 10 min centrifugation at 10 000 xg. The supernatants were recovered as periplasmic extracts. The pellets could be continued to cytoplasmic extraction.

7.1.15 Cytoplasmic extraction

Pellet from 10 ml bacterial culture or the remaining pellet from the periplasmic extraction was resuspended in 500 µl of lysis buffer (50 mM Tris, 100 mM NaCl, 1 mM EDTA, pH8) and 10 µl of lysozyme (10 mg/ml) was added to the suspension and incubated at room temperature for 30 minutes. Before sonication, 20 µl DNAse (1 mg/ml) were added to the samples and incubated at room temperature for 15 minutes. For cell lysis, sonication was applied 3-5 times for 20 sec at 40% amplitude to the samples on an ice bath. Samples were centrifuged at 4 °C and 12 000 xg for 10 min and the supernatant collected as the soluble cytoplasmic fraction.

7.1.16 Complete cell lysis

Pellet of 10 ml bacterial culture was resuspended in 500 µl of lysis buffer containing 10% glycerol, 0.1% Triton X-100, 100ug/ml lysozyme, 1 mM EDTA and 3 U DNAse. The cell suspension was incubated at room temperature for 30 min and sonicated 5 times for 20 seconds until the samples are no longer viscous and a clear lysate was obtained. After centrifugation at 4 °C and 12 000 xg for 10 min, the supernatant collected as the soluble cell lysate.

7.1.17 Insoluble fraction

The pellet remaining from the cytoplasmic extraction was considered as insoluble fraction and it was suspended in 400 µl of PBS 1X pH 7 for further analysis.

7.1.18 SDS-PAGE

Samples from different fractions were suspended in a sample buffer (50 mM Tris base, 10% glycerol, 0.01% bromophenol blue, 2% mercaptoethanol, 2% SDS and H2O at pH 6.8) and boiled for 10 minutes at 95 °C. The samples were then loaded onto a 8% or 10% precast gel (Eurogentec, Belgium) immersed in MOPS

running buffer, and subjected to a current of 130V for 50-80 minutes. The gels were then revealed by soaking in InstantBlue Coomassie Protein Stain (Westburg) for 30 minutes, followed by washing with H_2O.

7.1.19 Western Blotting

Protein samples are run on polyacrylamide gel (SDS PAGE) and transferred onto a 0.2 μm nitrocellulose membrane using the Trans-Blot® Turbo™ Transfer System at 2.5 constant A and up to 25 V for 5 minutes. Then the membrane is incubated in blocking solution (5% non-fat dry milk, 0.2% Tween, PBS 1x buffer pH 7.2) for 1 h at room temperature under gentle agitation to prevent non-specific antibody binding. The membrane is incubated for 1 h in primary antibody (5mg/mL of rabbit anti-His (Sigma) diluted in blocking solution). The membrane is washed 3 times with the blocking solution for 10 minutes under agitation to remove unbound primary antibody. The tray is changed to decrease the background effect of primary antibodies and the membrane is incubated for 1 h in secondary antibody (Peroxidase AffiniPure Donkey Anti-rabbit IgG (H + L)), Jackson ImmunoResearch, UK). Before proceeding to detection, the membrane is washed 2 times with the blocking solution for 10 minutes under agitation followed by one time washing with PBS 1x buffer. For chemiluminescence detection of signals, the membrane is soaked in substrates mixture of substrate A and substrate B (Pierce ™ ECL Plus Western Blotting Substrate, Thermofisher) with a ratio of 40: 1 as suggested by supplier. The signals are detected on Amersham™ Imager 600 (GE Healthcare).

7.1.20 Purification of PBP-A proteins

The total soluble lysate of *E. coli* cells expressing different PBP-As were loaded onto a Ni-Sepharose column. After several Wash by washing buffer (HEPES 20 mM, Imidazole 20mM, 150 mM NaCl, pH 8) several times the column volume, a stepwise gradient of imidazole (50 to 500 mM in 50 mM steps) was applied to elute the variants.

7.1.21 Beta-lactamase activity assay

The assay for beta-lactamase activity was conducted for periplasmic, cytoplasmic or insoluble fractions at 25 °C in 1x PBS, pH 7. The 20 µl of lysate samples were added to 480 µl of PenG (0,5 mM) in PBS buffer (1X, pH 7) in a quartz cuvette. To measure the hydrolysis rate of benzylpenicillin (penicillin G or PenG), the absorbance at 232 nm was spectrophotometerically monitored over time on (Δ = -800 $M^{-1}.cm^{-1}$).

7.1.22 Bacterial growth measurements

Pre-cultures are started by inoculation of cells from single colony into 5 ml of media containing appropriate antibiotics and glucose (for repression of leaky recombinant protein expression) and incubated overnight at 37 °C under agitation of 180 rpm. To minimize bias related to initial population, prior to starting the growth measurement experiment, OD_{600} of different cultures were normalized by dilution of more dense cultures. From OD_{600}-normalized cells, 1:100 dilution is made into fresh LB media with appropriate conditions in a sterile 96 well cell culture plate (Greiner Bio-One). Plates are incubated at 37 °C, 180 rpm in either a Tecan (Tecan Group Ltd) or Hidex Sense (Hidex) multiwell microplate reader and OD_{600} was measured every 15 minutes for 16-20 hours. Specific growth rate (μ_{max}) of each culture was estimated by measuring the slope of the growth curve in exponential phase and the maximum OD_{600} (OD_{max}) was obtained from the highest value of absorbance reached during the growth.

7.1.23 Serial dilution

Serial dilutions is used to compare fitness of different bacterial strains under different conditions on agar plates based on their colony forming units per volume (cfu/ml). Similar to previous section pre-cultures are prepared. Cultures of different bacterial strains were induced for expression of PBP-As by adding L-arabinose (0.5% w/v) at OD600 ~ 0.2-0.3 and grown to late exponential phase (OD600nm ~ 0.6-0.7). Prior to start the dilution process, cultures were normalized to the same OD_{600}, by diluting the higher density cultures. Then cultures were diluted in series of sequential 10 fold dilutions using sterile LB

media and were spotted on LB agar supplemented with/without appropriate antibiotics and inducer for PBP-A expression. Plates were incubated overnight at 37 °C and analyzed next day by estimating cfu/ml.

7.2 Building Scarless Gene Libraries in the Chromosome of Bacteria

Gol Mohammad Dorrazehi, Sebastian Worms, Jason Baby Chirakadavil, Johann Mignolet, Pascal Hols, Patrice Soumillion

Reprinted/adapted by permission from Springer Nature: Building Scarless Gene Libraries in the Chromosome of Bacteria, Gol Mohammad Dorrazehi, et al [License Number : 5176420084558 5] (2020).

7.2.1 Principle

Since its first reported use for genomic editing in 2013, CRISPR-Cas9 has established itself as the prime technology for lab-scale genome editing (Cong et al., 2013; Ran et al., 2013). In bacteria, CRISPR-Cas9 can be coupled with the λ-Red recombination system to efficiently introduce foreign DNA fragments in the chromosome (Ao et al., 2018; Jiang et al., 2015; Hong Zhang, Cheng, Liu, Zhao, & Wang, 2017). When used independently, the λ-Red system usually has relatively low efficiency, necessitating the additional use of a selection marker to select for positive colonies. By using CRISPR-Cas9 to cleave the target chromosomal site, insertion of the foreign DNA through λ-Red-mediated homologous recombination is the only way for the bacteria to repair the double strand break and to survive. Therefore, we avoid the need to transform a selection marker along with our gene of interest (Chayot, Montagne, Mazel, & Ricchetti, 2010).

For our libraries, we adopted the genome modification protocol from Jiang et al. (2015), which uses a two-plasmid system: pCas, encoding the λ-Red recombination system under the arabinose-inducible P_{araBAD} promoter as well as Cas9 under its constitutive promoter, and pTarget which encodes the sgRNA targeting the genome (Figure 7-2) (Jiang et al., 2015). In addition, pCas includes a kanamycin resistance marker and a sgRNA that targets the pTarget plasmid under the IPTG-inducible P_{trc} promoter affording the curing of the plasmid after genome modification. pCas itself has the temperature-sensitive replication origin *repA101*(Ts) for self-curing.

The library building relies on several steps (Figure 7-4). First, the chromosomal locus of interest is determined and a Cas9 cutting site is chosen. A sequence encoding the chosen sgRNA is introduced in the pTarget. If the library is to be cloned into pTarget (see below), two homology regions (minimum 500 bp each) corresponding to the sequence on each side of the chosen cutting site and required for the λ-Red recombination are also cloned into pTarget. The activity of the sgRNA can then be tested by a toxicity assay in *E. coli* (see section 7.2.3). In parallel, a DNA library (Lib) is created by any classical technique (error-prone PCR, incorporation of degenerate oligonucleotides, DNA shuffling…) and flanked by two homology regions (U and D for up and down) for chromosomal integration. For creating small size libraries (10^3-10^4 transformants), a linear DNA library U-Lib-D can be used directly for co-transformation with pTarget. For larger libraries (10^5-10^6 transformants), the library must be integrated into the pTarget plasmid, in between the previously added U and D homology arms. Once all parts are assembled, the actual genome modification can take place. Cells containing the pCas plasmid are grown and the expression of the λ-Red recombination enzymes is induced by adding arabinose. The cells are then made competent and co-transformed with the linear DNA library and pTarget plasmid or transformed with the library inserted into pTarget. Cas9 is activated by the sgRNA encoded on pTarget and generates the site-specific double-stranded break. Homologous recombination with U-Lib-D DNA is facilitated by λ-Red system while cells that failed to recombine cannot survive.

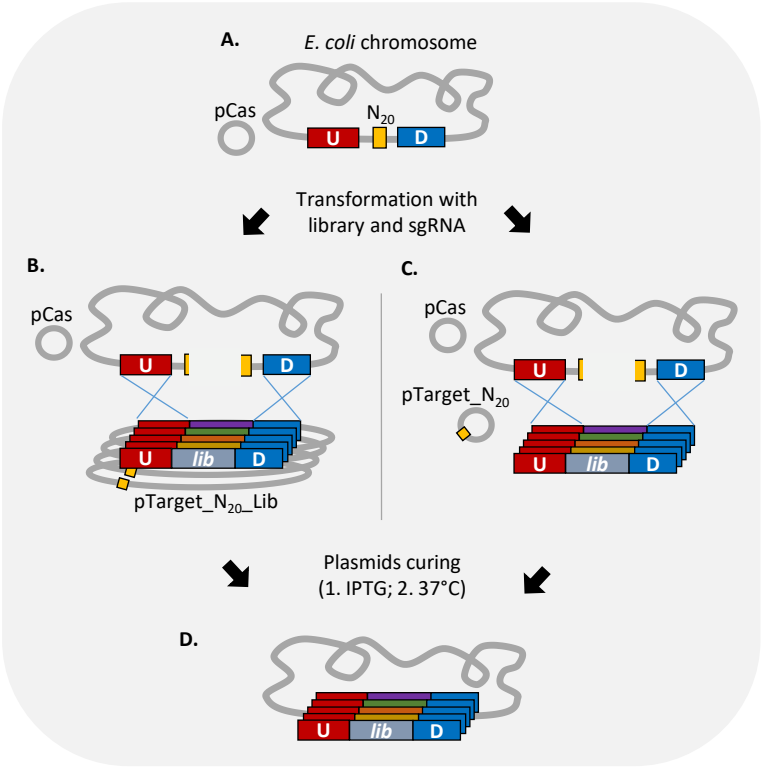


Figure 7-4: Strategy for the generation of gene libraries in the chromosome of *E. coli*. **A)** A suitable nonessential locus is chosen. Within that locus, a Cas9 site (PAM + N_{20}) is selected with the help of scoring algorithms, and 500 bp homology arms (U and D) as close as possible to the Cas9 site (less than 200 bp is recommended). The pCas plasmid is introduced into the strain. **B)** For reaching high diversities (10^5–10^6 individual transformants), the library is cloned into the pTarget, flanked by the U and D homology arms, together with the sequence encoding the sgRNA. The plasmid library is then introduced into the pCas containing strain, triggering chromosome cleavage. Homologous recombination, facilitated by the λ-Red proteins, is the only way to survive. **C)** If high diversity is not required (10^3–10^4 individual transformants), a cotransformation with the sgRNA encoding pTarget and the U- and D-flanked library as a linear DNA fragment is a simpler protocol. **D)** After transformation, the pTarget_N_{20} induces (IPTG) the expression of an antipTarget sgRNA encoded into the pCas, and the thermosensitive pCas itself is subsequently cured by growing the library at 37°C.

Following verification of a correct insertion, cells can be grown on IPTG to induce the expression of a pCas-borne sgRNA targeting the pTarget plasmid, resulting in its curing. Finally, growing the cells at 37°C will remove the pCas plasmid due to its temperature-sensitive origin of replication. The loss of antibiotic resistance can be checked after curing.

As an example, the following protocols are described for a library of *phpA* genes (894 bp, encoding a periplasmic penicillin binding protein) created by error-prone PCR and inserted into the rhamnose operon of *E. coli* in such a way that the *phpA* gene is inserted at the second position of the polycistronic operon, with its ATG at the same position as that of the wt *rhaA* gene (Figure 5-2). The *phpA* gene comprises a sequence encoding the amino-terminal export signal (SS) of the DsbA protein for periplasmic localization and a sequence encoding a carboxy-terminal 6xHis tag for downstream purification. These external sequences are not targeted by the error-prone PCR.

7.2.2 Selection of sgRNAs and cloning into pTarget

The Cas9 nuclease requires a guide RNA (sgRNA) that comprises a sequence recognized by Cas9 and a 20-nucleotide sequence (N_{20}) complementary to a region of the genome next to a protospacer-adjacent motif (PAM) featuring a NGG sequence. The Cas9 endonuclease cuts three base pairs upstream of the PAM sequence. The choice of an N_{20} is of paramount importance to the library creation process as the efficiency of the Cas9 nuclease depends on the sgRNA (Doench et al., 2016). Several algorithms have been developed to help scientists select a suitable N_{20} site. Many of them are available online as user-friendly interfaces. We used the Design CRISPR Guides tool on the Benchling platform (http://www.benchling.com), which calculates on- and off-target scores based on algorithms developed by Doench, Fusi et al. 2016 and Hsu et al. 2013 respectively (Doench et al., 2016; Hsu et al., 2013). The first 200-300 bp of the *rhaA* gene of *E. coli* were selected as input to find the highest-scoring N_{20}. Scores above 60 and 50 are considered to correspond to appropriate guides for on-target and off-target effects respectively.

After selection of an appropriate N_{20}, it is introduced into the pTarget vector by amplifying the whole plasmid with a long primer containing a floating tail with the N_{20} and a shorter reverse primer complementary to the end of the floating tail. After amplification, 100 ng of purified PCR product (linear plasmid) can be self-ligated by adding 15 µL of Gibson Master Mix, 10 U of DpnI and $_{dd}H_2O$ to

20 µL. After incubation for 1 hour at 50 C, 1 µL of the Gibson reaction product can be electroporated in *E. coli* TOP10 cells (Galka et al., 2017). In the example used in this chapter, the primers used to insert the $N_{20\text{-}rhaA}$ (GCCTGTGGCCTGAATCCCCC), targeting the *rhaA* locus, are found in the primer table as primer pair A.

7.2.3 pTarget validation

To validate the sgRNA, *E. coli* TOP10 cells containing the pCas plasmid are transformed either with the modified pTarget or with a pTarget expressing a sgRNA that does not target the *E. coli* chromosome and the ratio of transformants is determined. Since double strand breaks are lethal to *E. coli*, a plasmid encoding an adequate sgRNA should not be able to transform *E. coli* that produce Cas9. We use a pTarget_ALR plasmid (N_{20} : GCTGTCTCTAAACTTAGGGC) as negative control.

1. Start a preculture of *E. coli* TOP10 containing pCas in 5 mL of LB-kanamycin and incubate it O/N at 30 C.
2. Inoculate 10 mL of LB-kanamycin with 100 µL of pre-culture. Grow at 30°C until an OD_{600} of 0.4-0.6 is reached.
3. Centrifuge down the cells for 7 minutes at 3,000 × *g*. Discard the supernatant and resuspend the cells in 5 mL of cold $_{dd}H_2O$.
4. Repeat step 3.
5. Centrifuge the cells for 7 minutes at 3,000 × *g*. Resuspend the cells in 1 mL of cold $_{dd}H_2O$. Transfer to a microcentrifuge tube.
6. Centrifuge the cells in a benchtop centrifuge at 10,000 × *g* for 10 seconds. Discard the supernatant and resuspend the cells in 100 µL of cold $_{dd}H_2O$.
7. To 50 µL of the competent cells, add 20 ng of either the pTarget to be tested or pTarget_ALR. Transfer to a 2 mm electroporation cuvette.
8. Electroporate the cells at 2.5 kV. Resuspend the cells in 950 µL of SOC medium. Allow the cells to recover at 30 C for 45 minutes.

9. Plate 10 and 100 μL of each transformation on LB-Kan-Tet plates and incubate overnight at 30 C.
10. Compare the colony count on both plates. If we see a thousand-fold difference in the colony count on both plates, it is considered as adequate.

7.2.4 Donor DNA cassette design

In order to favor recombination, relatively long homology regions of 500 bp are used. In order to avoid the presence of a promoter in the donor DNA cassette (*see* **Note 1**), we inserted the library in the second gene (*rhaA)* of the rhamnose operon through recombination. The upstream homology arm (U) was design by taking the 500 bp immediately upstream of the *rhaA* start codon. The downstream homology arm (D) was designed by taking 500 bp downstream of the selected N_{20}. The homology regions must be as close as possible from the Cas9 cutting site. Because of library design constraints, in our example (Figure 5-2), U and D regions are 165 bp upstream and downstream the N_{20} sequence. This design allowed the building of a million clones library. Higher diversities may be reached by reducing these distances whenever possible.

7.2.5 Library construction

As an example, an error-prone library (Lib) is constructed by running an ep-PCR (protocol in section 7.2.11) using primer pair F and the *pbpA* gene as template. Both primers used in ep-PCR should harbor homologous sequences to the U and D arms to allow the subsequent assembly of the ep-library into the donor DNA cassette.

Linear cassette assembly to generate small libraries

The homologous arms, U and D, are produced by Q5-mediated PCR from the genome with primer pair B1 for the U arm and pair D1 for the D arm. Since the signal sequence is not included in the library, it must be assembled with the U fragment. An SS-*pbpA* fragment is prepared by PCR using primer pair C and further fused with U by Gibson assembly (*see* **Note** 2). Then, this assembled fragment of U-SS-*pbpA* is used as the template to amplify U-SS without *pbpA* by

the help of primer B1 forward and G reverse. DNA fragments of U-SS and D are harboring overhangs complementary to the extremities of the library.

Following purification of the DNA, the library and the two homologous arms should be assembled in a way to minimize the introduction of bias in the donor DNA library. Two different methods described below can be used that give similar results.

Primerless overlap extension PCR (OE-PCR):

In primer-less overlap extension PCR, the internal overlapping sequence of U-SS, D and Lib fragments are used as primers until full U-Lib-D fragments are assembled.

A Q5-mediated PCR is performed in a volume of 25 µL in the absence of primers using 500 fmol of each fragment as template. In our experience, 15 cycles of extension are enough to fully assemble the cassette. The U-Lib-D fragment should be detectable on an agarose gel, ensuring that a large number of individual molecules have been assembled.

In order to obtain high concentrations of linear donor DNA needed for transformation, the U-Lib-D products of the OE-PCR reaction are amplified in a standard Q5-mediated PCR using external primers (forward primer from pair B1 and reverse from pair D1). The expected band is purified from agarose gel using the Monarch Gel Extraction Kit or an equivalent kit prior to transformation.

Gibson assembly

Instead of the OE-PCR, the U-Lib-D can be assembled by mixing 15 µL of Gibson Master Mix (*see* **Note** 2) with 500 fmol of each of the fragments, 0.5 µL of DpnI and $_{dd}H_2O$ to 20 µL. The DNA is then incubated for 1 hour at 50°C. The U-Lib-D fragment should be detectable on agarose gel before further amplifying it by PCR as described in the previous section.

7.2.6 Cassette assembly in pTarget to generate large libraries

Integrating the donor cassette into the pTarget plasmid gave us more efficient genome integration, leading to a 10- to 100-fold increase in library size compared

to the linear donor DNA method. However, this is at the cost of a more laborious cassette construction process. In order to clone the donor DNA into the plasmid, the homologous arms (U and D) are amplified using primer pairs B2 and D2, which includes flanking regions homologous to pTarget for Gibson assembly.

1. Amplify the pTarget-N_{20} backbone using primer pair E. Purify it using the Monarch PCR & DNA Cleanup kit or equivalent.
2. Mix equimolar amounts (500 fmol) of the U, SS-*pbpA* (PCR product of primer pair C) and D fragments together with plasmid backbone in 15 μL of Gibson Master Mix (*see* **Note** 2), 0.5 μL of DpnI and $_{dd}H_2O$ to 20 μL, and incubate 1 hour at 50°C.
3. Transform 1 μL of the assembly mix into *E. coli* TOP10 competent cells.
4. Extract the plasmid harboring the donor DNA cassette (pTarget-N_{20}-U-SS-*pbpA*-D) by plasmid miniprep.
5. Amplify the plasmid/cassette template obtained in step 4 with primer pair G to amplify the plasmid backbone without *pbpA* in order to assemble it with the library between U-SS and D.
6. Mix equimolar amounts (500 fmol) of the Lib fragment and plasmid backbone with 15 μL of Gibson Master Mix, 0.5 μL of DpnI and $_{dd}H_2O$ to 20 μL, and incubate 1 hour at 50 C.
7. Clean the Gibson product using the Monarch PCR & DNA Cleanup kit or equivalent. Elute the DNA in 7 μl of $_{dd}H_2O$ and use 4 μL of the DNA mix for electroporating *E. coli* TOP10 competent cells.
8. Verify random colonies by colony PCR for the insertion of U-Lib-D into pTarget using primer pair. The size of the library is estimated by calculating the number of transformants obtained at step 7, and the expected library size is achieved by scaling up.
9. Scrape all colonies from plates with LB medium using a spreader.
10. Pool the bacteria and extract the plasmid library using a QIAGEN Plasmid Maxi Kit or equivalent maxi prep kit to obtain the maximum amounts of plasmids for the genome modification step.

7.2.7 Preparation of *E. coli* TOP10 electro-competent cells harboring pCas

1. Plate *E. coli* cells harboring pCas on LB agar containing kanamycin (50 mg/L) and incubate O/N at 30°C (*see* **Note 3**).
2. Among the grown colonies, pick a small one (this is very important, *see* **Note 4**) and inoculate in 10 mL of LB containing 50 mg/L of kanamycin (LB-Kan) in a 15 mL centrifuge tube and incubate overnight at 30 °C without agitation (*see* **Note 5**).
3. Start a culture by inoculating 40 mL of LB-Kan with 400 μL of pre-culture. Incubate in a 500 mL flask at 30°C and 180 rpm until an OD_{600} between 0.4 and 0.6 is reached.
4. Add arabinose (10 mM final concentration) to the culture to induce the λ-Red recombination proteins. Incubate for 15 minutes at 30°C and 180 rpm.
5. Chill the culture in a water-ice bath for 30 min.
6. Transfer the culture to a pre-chilled 50 mL centrifuge tube and centrifuge at 4,600 × *g* and 4 °C for 7 minutes.
7. Add 30 mL of ice-cold $_{dd}H_2O$ to the pellet. Swirl gently to resuspend the cells.
8. Centrifuge again at 4,600 × *g* and 4 °C for 7 minutes.
9. Resuspend the pellet in 1 mL of chilled $_{dd}H_2O$. Transfer it using a 1,000 μL sterile tip to a pre-chilled sterile 1.5 mL microcentrifuge tube.
10. Centrifuge down at 10,000 × *g* and 4 °C for 4 seconds in a benchtop centrifuge.
11. Carefully pipette out the supernatant and repeat this step once more.
12. Resuspend cells in 200 μL of chilled $_{dd}H_2O$. Keep on ice until transformation.

7.2.8 Transformation

1. Add either 100 ng of pTarget-N_{20}-U-Lib-D or 100 ng of pTarget-N_{20} and 400 ng of linear donor U-Lib-D cassette to 50 μL of competent cells.
2. Electroporate the cells in a 2 mm cuvette and recover them in 950 μL of SOC medium.
3. Incubate them for 45 minutes at 30°C before plating all the cells on LB-Kan-Tet plates. Incubate overnight at 30 °C.
4. Library size is estimated by counting the number of individual transformants.

7.2.9 Verification of genome modification

Random colonies are verified by colony PCR using primers outside the recombination arms (primer pair H). The validation of positive clones as well as their mutation rate (~ 4-6 mutation/gene in our case) is checked by sequencing of the colony PCR product using internal primer pair I. Library size is corrected if a fraction of clones does not contain the gene of interest. If the library size is not sufficient, the whole process should be repeated, upscaling the reaction as needed.

7.2.10 Plasmid curing

In order to cure pTarget, the colonies harboring our library are scraped from the plates with LB medium using a spreader and 0.1 mL is inoculated into 10 mL of LB medium supplemented with 50 mg/L of kanamycin and 0.5 mM IPTG. After incubation at 30°C overnight, 0.1 mL of the culture is diluted in 10 mL of LB supplemented with 50 mg/L of streptomycin (*see* **Note 6**) and incubated at 37 °C overnight to cure the pCas plasmid. The curing is confirmed by the inability of cells to grow on LB-tetracyclin (12.5 mg/L) or LB-kanamycin (50 mg/L).

7.2.11 PCR protocols

These three protocols are used in this chapter.

Q5 PCR: 10 µL of 5 x Q5 Reaction Buffer, 0.7 µL of 20 mM dNTPs, 2.5 µL of 10 µM Forward and reverse primers, less than 100 ng of template DNA, 1 U Q5 High-Fidelity DNA Polymerase (NEB) and nuclease free H_2O up to 50 µL.
PCR thermocycler conditions: initial denaturation at 98°C for 30 seconds; 25-30 cycles containing: denaturation at 98°C for 20 seconds; annealing at temperature calculated for each pair of primers on https://tmcalculator.neb.com/#!/main (*see* **Note 7**) for 30 s; 20–30 seconds/kb extension at 72°C; a final extension for 5 minute at 72°C and hold at 4°C.

Error-prone PCR (ep-PCR) using GeneMorph II Random Mutagenesis Kit: 5 µl of 10× Mutazyme II reaction buffer, 1 µl of 40 mM dNTP mix (provided in kit), 0.5 µl of primer mix (250 ng/µl of each primer), 2.5 U of Mutazyme II DNA polymerase, 100 ng of template DNA (*see* **Note 9**), and nuclease free H_2O up to 50 µL.
PCR thermocycler conditions: initial denaturation at 95°C for 2 minutes; 20-25 cycles (depending on desired mutation rate, *see* **Note 8**): denaturation at 95°C for 30 seconds; annealing 30 seconds at $T = T_m - 5°C$ (Tm calculated for each pair of primers on https://tmcalculator.neb.com/#!/main); extension 1 minute (≤1-kb targets) or 1 minute/kb (>1-kb targets) at 72°C; a final extension for 10 minutes at 72°C and hold at 4°C.

Optimization of mutation rate: To generate low mutation rate (between 0 and 5 mutation/kb), instructions of the GeneMorph II Random Mutagenesis kit are followed according to manufacturer. According to the kit's protocols, mutation rates can be lowered by using higher template DNA amounts and by decreasing the number of cycles of ep-PCR reaction. Therefore, for a 1kb fragment, an initial target amount of 100 ng is required for a run with 15-20 cycles of PCR, generating a low mutation rate between 0 and 5 mutation/kb (*see* **Note 9)**.

Colony PCR: 5 µL of 5 × GoTaq Flexi Buffer, 4 µL of 25 mM $MgCl_2$, 0.7 µL of 20 mM dNTPs, 1.5 µL of 10 µM Forward and reverse primers, 1.25 U GoTaq DNA Polymerase (Promega) and nuclease free H_2O up to 50 µL. Use a toothpick

to stab a single colony and rub it inside the PCR tube to detach some cells as template.

PCR thermocycler conditions: initial denaturation at 95°C for 4 minutes; 25 cycles: denaturation at 95°C for 30 seconds; annealing at Ta (calculated for each pair of primers on https://tmcalculator.neb.com/#!/main) for 30 seconds; 1 minute/kb extension at 72°C; a final extension for 7 minutes at 72°C and hold at 4°C.

7.2.12 Notes

1. The expression of a toxic protein from the donor DNA cassette prior to homologous recombination could prevent insertion or generate biases in the library. To minimize this bias, a strategy is to avoid putting promoter sequence in the donor DNA. However, the recombination requires that the cassette includes the region immediately upstream of the library, where the promoter is typically located. A solution is to insert our gene of interest in a locus with a known expression pattern, but far enough from the promoter so a to have room for the upstream homology arm between the promoter and the gene. For example, the library can be inserted as the second gene in a polycistronic operon controlled by a single promoter with an upstream homology arm corresponding to the first wild-type gene of the operon. In this chapter, the example in *E. coli* is the cloning of a *pbpA* library in the rhamnose operon controlled by the P_{rha} promoter (Figure 5-2). It is also important to avoid the use of the P_{lac} and P_{araBAD} promoters for the library since these promoters are already used in the CRISPR-Cas9/□ -Red system.
2. We tried commercial Gibson Mix such as the NEBuilder® HiFi DNA Assembly Master Mix from New England Biolabs, but observed an efficiency similar to our own Gibson Master Mix (Galka et al., 2017). Given the much lower cost of our Gibson mix per transformation, we usually use it.
3. Since pCas has a thermosensitive replication origin, it is important to grow cells containing pCas at 30°C instead of 37°C.

4. Colonies of *E. coli* harboring pCas are heterogeneous in size. This is due to a slight toxicity of Cas9 endonuclease. It is therefore very important to start experiments with a small colony since the large ones do not express active Cas9.
5. Starting the pre-culture of *E. coli* in 10 mL of LB medium in a non-shaking 15 mL centrifuge tubes leads to a shorter lag phase when the culture is inoculated the following day. Hypothesis is that the low oxygen availability reduces cell death (Hall, Acar, Nandipati, & Barlow, 2014).
6. In order to prevent possible contamination, streptomycin is used after the curing steps since *E. coli* TOP10 cells harbors a chromosomal resistance gene.
7. The NEB Biocalculator (https://nebiocalculator.neb.com/) is a handy tool to calculate several parameters of DNA oligonucleotides and fragments such as molarity, molecular weights, etc...
8. When amplifying a library by PCR, it is important to analyze the PCR reaction every 2 or 3 cycles so that the minimal number of cycles that are necessary for obtaining a maximal PCR yield can be estimated. This can be used to estimate the number of DNA molecules that have been amplified. This number must be at least 100-fold higher than the expected final diversity of the library. Avoid any unnecessary PCR cycles that does not improve the yield since this can generate biases in the library.
9. The mutation rate can be controlled by adjusting the number of epPCR cycles and the amount of initial template. 100 ng template and 20-25 cycles afford a low mutation frequency (between 0 and 5 mutations/kb). See also the user manual for the "Random Mutagenesis Kit" from Agilent.

7.2.13 Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 722610 (ES-CAT). This material reflects only the author's views and the Union is not liable for any use that may be made of the information

contained therein. G.M.D, S.W. and J.B.C. contributed equally to this work, P.H. and P.S. edited the manuscript.

7.2.14 Materials

Common kits and reagents

The GoTaq® Flexi DNA Polymerase from Promega and the Q5® High-Fidelity DNA Polymerase from New England Biolabs were used in this chapter. Monarch PCR & DNA Cleanup Kit (New England Biolabs) was used for gel extraction and purification of PCR products. Midi and Maxi prep kits were used for plasmid purification. All other enzymes were purchased from Promega or New England Biolabs unless explicitly mentioned.

Gibson Master Mix: 100 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 0.2 mM of each four dNTPs, 1mM NAD^+, 15% (w/v) PEG-8000, T5 exonuclease (2 U/mL), Phusion DNA polymerase (33 U/ mL), Taq DNA ligase (1666 U/mL) (Galka et al., 2017).

Growth Media: *Lysogeny Broth* (LB): 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl. When appropriate, we supplemented the growth media with kanamycin (50 mg/L), tetracycline (12.5 mg/L) or streptomycin (50 mg/L).

SOC Medium*:* 2% (w/v) tryptone, 0.5% **(w/v)** yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 10 mM $MgCl_2$, 20 mM glucose.

Strains and plasmids

E. coli strain TOP10 was purchased from Invitrogen. pCas and pTargetF plasmids were ordered from Addgene (plasmids #62225 and #62226). The pTarget plasmid with a tetracycline resistance gene was created by amplifying the TetR cassette with primer pair J2 and the pTargetF plasmid with primer pair J1 and ligating both PCR products with the Gibson Master Mix.

Electroporation

All transformations were performed using a Bio-Rad Gene Pulser Xcell™ electroporator together with 2 mm electroporation cuvettes.

List of oligonucleotides

All primers were purchased from Eurogentec (Seraing, Belgium) and are listed in Table 7-1**.** The sequences in bold and black are the 3'-end which hybridize to the target, while the red sequences are the 5' overhang regions necessary for fragment assembly. The sequence in italic and blue for primer A fw is the N_{20}-rha complementary to the *rhaA* locus of the *E. coli* genome (see section 7.2.2). Annealing temperatures (Ta) were predicted taking only the hybridizing sequences of the primers using the online tool available at http://tmcalculator.neb.com/.

Table 7-1 : List of primers used in 'Building Scarless Gene Libraries in the Chromosome of Bacteria'.

PCR/Target to amplify	**Primer Pair**		**Sequence (5'→3')**
N_{20} modification on pTarget	A	Fw	AGCTAGCTCAGTCCTAGGTATAATACTAGT*GCCTGTGGCCTGAATCCCCC***GTTTTAGAGCTAGAAATAGCAAGTTAAAA**
		Rv	**ACTAGTATTATACCTAGGACTGAGCTAGCT**
U arm	B1	Fw	**AGGCACTTCCGGCTTGCC**
		Rv	AAATCTTTTTCAT**GCGCAAAGCTCCTTTGTC**
	B2	Fw	CTCGAGTAGGGATAACAGGG**AGGCACTTCCGGCTTGCC**
		Rv	AAATCTTTTTCAT**GCGCAAAGCTCCTTTGTC**
SS-*pbpA*-6xHis	C	Fw	TTTGCGC**ATGAAAAAGATTTGGCTGGCG**
		Rv	CCCATTCAAC**TCAGTGATGATGATGATGG**
D arm	D1	Fw	TCATCACTGA**GTTGAATGGGCGAAAGCCAATC**
		Rv	**AGCAGCAACTGCGGCACA**

	D2	Fw	TCATCACTGA**GTTGAATGGGCGAAAGCCAATC**
		Rv	GCAGAAGCTTAGATCTATTA**AGCAGCAACTGCGGCACA**
pTarget-N_{20}	E	Fw	**TAATAGATCTAAGCTTCTGCAG**
		Rv	**CCCTGTTATCCCTACTCG**
PBP-A gene (ep-PCR)	F	Fw	**TTTTAGCGTTTAGCGCATCGGCGGCG**
		Rv	GGCTTTCGCCCATTCAAC**TCAGTGATGATGATGATGGTG**
pTarget-N_{20}-U-SS-*pbpA*-D	G	Fw	**GTTGAATGGGCGAAAGCC**
		Rv	CGATGCGCTAAACGCT**AAAACTAAACCAGCCAGCGCCAG**
Modified locus in genome	H	Fw	**TTCAGCGAGTGCTTCAGGA**
		Rv	**AGATCGTGACGCACAATCTC**
PCR product of primer pair H	I	Fw	**CCACCTTTACCCCTAATCC**
		Rv	**ACTGGCGAATGCTGTCGT**
Resistance change on pTargetF	J1	Fw	GCCGGGCCACCTCGACCTG**AGATGCCGCTCGCC**
		Rv	GCGCATTGTTAGATTT**CATCAGTCGATCATAGCACG**
	J2	Fw	**ATGAAATCTAACAATGCGC**
		Rv	**TCAGGTCGAGGTGGC**

8. References

Adachi, H., Ohta, T., & Matsuzawa, H. (1991). Site-Directed Mutants, At Position 166, Of RTEM-1 Beta-Lactamase That Form A Stable Acyl-Enzyme Intermediate With Penicillin. *The Journal Of Biological Chemistry*, *266*(5), 3186–3191.

Aertsen, A., & Michiels, C. W. (2005). Diversify Or Die: Generation Of Diversity In Response To Stress. *Critical Reviews In Microbiology*, *31*(2), 69–78. Https://Doi.Org/10.1080/10408410590921718

Agarwal, P. K. (2019). A Biophysical Perspective On Enzyme Catalysis. *Biochemistry*, *58*(6), 438–449. Https://Doi.Org/10.1021/Acs.Biochem.8b01004

Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. M., Roodveldt, C., & Tawfik, D. S. (2005). The "Evolvability" Of Promiscuous Protein Functions. *Nature Genetics*, *37*(1), 73–76. Https://Doi.Org/10.1038/Ng1482

Alberstein, R. G., Guo, A. B., & Kortemme, T. (2022). Design Principles Of Protein Switches. *Current Opinion In Structural Biology*, *72*, 71–78. Https://Doi.Org/Https://Doi.Org/10.1016/J.Sbi.2021.08.004

Alonso-Del Valle, A., León-Sampedro, R., Rodríguez-Beltrán, J., Delafuente, J., Hernández-García, M., Ruiz-Garbajosa, P., … San Millán, A. (2021). Variability Of Plasmid Fitness Effects Contributes To Plasmid Persistence In Bacterial Communities. *Nature Communications*, *12*(1), 2653. Https://Doi.Org/10.1038/S41467-021-22849-Y

Ambler, R. P. (1980). The Structure Of Beta-Lactamases. *Philosophical Transactions Of The Royal Society Of London. Series B, Biological Sciences*, *289*(1036), 321–331. Https://Doi.Org/10.1098/Rstb.1980.0049

Ambler, R. P., Coulson, A. F., Frère, J. M., Ghuysen, J. M., Joris, B., Forsman, M., … Waley, S. G. (1991). A Standard Numbering Scheme For The Class A Beta-Lactamases. *The Biochemical Journal*, *276 (Pt 1*(Pt 1), 269–270. Https://Doi.Org/10.1042/Bj2760269

Andersson, D. I., & Hughes, D. (2010, April). Antibiotic Resistance And Its Cost: Is It Possible To Reverse Resistance? *Nature Reviews Microbiology*, Vol. 8, Pp.

260–271. Https://Doi.Org/10.1038/Nrmicro2319

Andersson, D. I., & Levin, B. R. (1999). The Biological Cost Of Antibiotic Resistance. *Current Opinion In Microbiology*, *2*(5), 489–493. Https://Doi.Org/Https://Doi.Org/10.1016/S1369-5274(99)00005-3

Andreis, S. N., Perreten, V., & Schwendener, S. (2017). Novel B-Lactamase Bla(ARL) In Staphylococcus Arlettae. *Msphere*, *2*(3). Https://Doi.Org/10.1128/Msphere.00117-17

Ao, X., Yao, Y., Li, T., Yang, T. T., Dong, X., Zheng, Z. T., … Guo, Y. (2018). A Multiplex Genome Editing Method For Escherichia Coli Based On CRISPR-Cas12a. *Frontiers In Microbiology*, *9*(OCT), 1–13. Https://Doi.Org/10.3389/Fmicb.2018.02307

Arnold, F. H. (1996). Directed Evolution: Creating Biocatalysts For The Future. *Chemical Engineering Science*, *51*(23), 5091–5102. Https://Doi.Org/Https://Doi.Org/10.1016/S0009-2509(96)00288-6

Arnold, F. H. (1998). Design By Directed Evolution. *Accounts Of Chemical Research*, *31*(3), 125–131. Https://Doi.Org/10.1021/Ar960017f

Arnold, F. H. (2006). Fancy Footwork In The Sequence Space Shuffle. *Nature Biotechnology*, *24*(3), 328–330. Https://Doi.Org/10.1038/Nbt0306-328

Arnold, F. H. (2018). Directed Evolution: Bringing New Chemistry To Life. *Angewandte Chemie (International Ed. In English)*, *57*(16), 4143–4148. Https://Doi.Org/10.1002/Anie.201708408

Arnold, F. H. (2019). Innovation By Evolution: Bringing New Chemistry To Life (Nobel Lecture). *Angewandte Chemie International Edition*, *58*(41), 14420–14426. Https://Doi.Org/Https://Doi.Org/10.1002/Anie.201907729

Arts, I. S., Gennaris, A., & Collet, J.-F. (2015a). Reducing Systems Protecting The Bacterial Cell Envelope From Oxidative Damage. *FEBS Letters*, *589*(14), 1559–1568. Https://Doi.Org/Https://Doi.Org/10.1016/J.Febslet.2015.04.057

Arts, I. S., Gennaris, A., & Collet, J. F. (2015b, June 3). Reducing Systems Protecting The Bacterial Cell Envelope From Oxidative Damage. *FEBS Letters*, Vol. 589, Pp. 1559–1568. Https://Doi.Org/10.1016/J.Febslet.2015.04.057

Astudillo, L., Bernad, S., Derrien, V., Sebban, P., & Miksovska, J. (2010). Probing

The Role Of The Internal Disulfide Bond In Regulating Conformational Dynamics In Neuroglobin. *Biophysical Journal*, *99*(2), L16–L18. Https://Doi.Org/10.1016/J.Bpj.2010.04.033

Axe, D. D. (2004). Estimating The Prevalence Of Protein Sequences Adopting Functional Enzyme Folds. *Journal Of Molecular Biology*, *341*(5), 1295–1315. Https://Doi.Org/10.1016/J.Jmb.2004.06.058

Babtie, A., Tokuriki, N., & Hollfelder, F. (2010a). What Makes An Enzyme Promiscuous? *Current Opinion In Chemical Biology*, *14*(2), 200–207. Https://Doi.Org/Https://Doi.Org/10.1016/J.Cbpa.2009.11.028

Babtie, A., Tokuriki, N., & Hollfelder, F. (2010b, April). What Makes An Enzyme Promiscuous? *Current Opinion In Chemical Biology*, Vol. 14, Pp. 200–207. Https://Doi.Org/10.1016/J.Cbpa.2009.11.028

Badran, A. H., & Liu, D. R. (2015, February 1). In Vivo Continuous Directed Evolution. *Current Opinion In Chemical Biology*, Vol. 24, Pp. 1–10. Https://Doi.Org/10.1016/J.Cbpa.2014.09.040

Bajaj, P., Singh, N. S., & Virdi, J. S. (2016). Escherichia Coli B-Lactamases: What Really Matters. *Frontiers In Microbiology*, *7*, 417. Https://Doi.Org/10.3389/Fmicb.2016.00417

Baker, D. (2019, April 1). What Has De Novo Protein Design Taught Us About Protein Folding And Biophysics? *Protein Science*, Vol. 28, Pp. 678–683. Https://Doi.Org/10.1002/Pro.3588

Banerjee, S., Pieper, U., Kapadia, G., Pannell, L. K., & Herzberg, O. (1998). Role Of The Omega-Loop In The Activity, Substrate Specificity, And Structure Of Class A Beta-Lactamase. *Biochemistry*, *37*(10), 3286–3296. Https://Doi.Org/10.1021/Bi972127f

Baret, J.-C., Miller, O. J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., … Griffiths, A. D. (2009). Fluorescence-Activated Droplet Sorting (FADS): Efficient Microfluidic Cell Sorting Based On Enzymatic Activity. *Lab On A Chip*, *9*(13), 1850–1858. Https://Doi.Org/10.1039/B902504A

Barreteau, H., Kovač, A., Boniface, A., Sova, M., Gobec, S., & Blanot, D. (2008). Cytoplasmic Steps Of Peptidoglycan Biosynthesis. *FEMS Microbiology Reviews*, *32*(2), 168–207. Https://Doi.Org/10.1111/J.1574-6976.2008.00104.X

Ben-David, M., Elias, M., Filippi, J.-J., Duñach, E., Silman, I., Sussman, J. L., &

Tawfik, D. S. (2012). Catalytic Versatility And Backups In Enzyme Active Sites: The Case Of Serum Paraoxonase 1. *Journal Of Molecular Biology*, *418*(3), 181–196. Https://Doi.Org/Https://Doi.Org/10.1016/J.Jmb.2012.02.042

Bloom, J. D., Labthavikul, S. T., Otey, C. R., Arnold, F. H., & Levitt, M. (2006). *Protein Stability Promotes Evolvability*. Retrieved From Www.Pnas.Orgcgidoi10.1073pnas.0510098103

Blumberg, P. M., & Strominger, J. L. (1974). Interaction Of Penicillin With The Bacterial Cell: Penicillin-Binding Proteins And Penicillin-Sensitive Enzymes. *Bacteriological Reviews*, *38*(3), 291–335. Https://Doi.Org/10.1128/Br.38.3.291-335.1974

Bommer, G. T., Van Schaftingen, E., & Veiga-Da-Cunha, M. (2020). Metabolite Repair Enzymes Control Metabolic Damage In Glycolysis. *Trends In Biochemical Sciences*, *45*(3), 228–243. Https://Doi.Org/10.1016/J.Tibs.2019.07.004

Booth, I R. (1985). Regulation Of Cytoplasmic Ph In Bacteria. *Microbiological Reviews*, *49*(4), 359–378. Https://Doi.Org/10.1128/Mr.49.4.359-378.1985

Booth, Ian R, & Louis, P. (1999). Managing Hypoosmotic Stress: Aquaporins And Medianosensitive Channels In Escherichia Coli. *Current Opinion In Microbiology*, *2*(2), 166–169. Https://Doi.Org/Https://Doi.Org/10.1016/S1369-5274(99)80029-0

Boros, I., Pósfai, G., & Venetianer, P. (1984). High-Copy-Number Derivatives Of The Plasmid Cloning Vector Pbr322. *Gene*, *30*(1), 257–260. Https://Doi.Org/Https://Doi.Org/10.1016/0378-1119(84)90130-6

Bouhss, A., Trunkfield, A. E., Bugg, T. D. H., & Mengin-Lecreulx, D. (2008). The Biosynthesis Of Peptidoglycan Lipid-Linked Intermediates. *FEMS Microbiology Reviews*, *32*(2), 208–233. Https://Doi.Org/10.1111/J.1574-6976.2007.00089.X

Boyken, S. E., Chen, Z., Groves, B., Langan, R. A., Oberdorfer, G., Ford, A., … Baker, D. (2016). De Novo Design Of Protein Homo-Oligomers With Modular Hydrogen-Bond Network–Mediated Specificity. *Science*, *352*(6286), 680 LP – 687. Https://Doi.Org/10.1126/Science.Aad8865

Braun, V., & Bosch, V. (1972). Repetitive Sequences In The Murein-Lipoprotein Of The Cell Wall Of Escherichia Coli. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *69*(4), 970–974.

Https://Doi.Org/10.1073/Pnas.69.4.970

Braun, V., & Rehn, K. (1969). Chemical Characterization, Spatial Distribution And Function Of A Lipoprotein (Murein-Lipoprotein) Of The *E. coli* Cell Wall: The Specific Effect Of Trypsin On The Membrane Structure. *European Journal Of Biochemistry*, *10*(3), 426–438. Https://Doi.Org/10.1111/J.1432-1033.1969.Tb00707.X

Braun, Volkmar, & Hantke, K. (2019). *Lipoproteins: Structure, Function, Biosynthesis BT - Bacterial Cell Walls And Membranes* (A. Kuhn, Ed.). Https://Doi.Org/10.1007/978-3-030-18768-2_3

Brown, L., Wolf, J. M., Prados-Rosales, R., & Casadevall, A. (2015). Through The Wall: Extracellular Vesicles In Gram-Positive Bacteria, Mycobacteria And Fungi. *Nature Reviews Microbiology*, *13*(10), 620–630. Https://Doi.Org/10.1038/Nrmicro3480

Brunette, T. J., Parmeggiani, F., Huang, P.-S., Bhabha, G., Ekiert, D. C., Tsutakawa, S. E., … Baker, D. (2015). Exploring The Repeat Protein Universe Through Computational Protein Design. *Nature*, *528*(7583), 580–584. Https://Doi.Org/10.1038/Nature16162

Buchholz, P. C. F., Fademrecht, S., & Pleiss, J. (2017). Percolation In Protein Sequence Space. *Plos One*, *12*(12), E0189646–E0189646. Https://Doi.Org/10.1371/Journal.Pone.0189646

Bukowska-Faniband, E., & Hederstedt, L. (2017). Transpeptidase Activity Of Penicillin-Binding Protein Spovd In Peptidoglycan Synthesis Conditionally Depends On The Disulfide Reductase Stoa. *Molecular Microbiology*, *105*(1), 98–114. Https://Doi.Org/10.1111/Mmi.13689

Bunzel, H. A., Anderson, J. L. R., & Mulholland, A. J. (2021). Designing Better Enzymes: Insights From Directed Evolution. *Current Opinion In Structural Biology*, *67*, 212–218. Https://Doi.Org/Https://Doi.Org/10.1016/J.Sbi.2020.12.015

Bush, K. (2018). Past And Present Perspectives On B-Lactamases. *Antimicrobial Agents And Chemotherapy*, *62*(10). Https://Doi.Org/10.1128/AAC.01076-18

Bush, K., & Jacoby, G. A. (2010, March). Updated Functional Classification Of B-Lactamases. *Antimicrobial Agents And Chemotherapy*, Vol. 54, Pp. 969–976. Https://Doi.Org/10.1128/AAC.01009-09

Bushweller, J. H. (2020). Protein Disulfide Exchange By The Intramembrane Enzymes Dsbb, Dsbd, And Ccda. *Journal Of Molecular Biology*, *432*(18), 5091–5103. Https://Doi.Org/Https://Doi.Org/10.1016/J.Jmb.2020.04.008

Busto, E., Gotor-Fernández, V., & Gotor, V. (2010). Hydrolases: Catalytically Promiscuous Enzymes For Non-Conventional Reactions In Organic Synthesis. *Chemical Society Reviews*, *39*(11), 4504–4523. Https://Doi.Org/10.1039/C003811C

Butera, D., Passam, F., Ju, L., Cook, K. M., Woon, H., Aponte-Santamaría, C., … Hogg, P. J. (2018). *Autoregulation Of Von Willebrand Factor Function By A Disulfide Bond Switch*. Retrieved From Http://Advances.Sciencemag.Org/

Cabeen, M. T., & Jacobs-Wagner, C. (2005). Bacterial Cell Shape. *Nature Reviews Microbiology*, *3*(8), 601–610. Https://Doi.Org/10.1038/Nrmicro1205

Campbell, E. C., Correy, G. J., Mabbitt, P. D., Buckle, A. M., Tokuriki, N., & Jackson, C. J. (2018). Laboratory Evolution Of Protein Conformational Dynamics. *Current Opinion In Structural Biology*, *50*, 49–57. Https://Doi.Org/10.1016/J.Sbi.2017.09.005

Caroff, N., Espaze, E., Bérard, I., Richet, H., & Reynaud, A. (1999). Mutations In The Ampc Promoter Of Escherichia Coli Isolates Resistant To Oxyiminocephalosporins Without Extended Spectrum Beta-Lactamase Production. *FEMS Microbiology Letters*, *173*(2), 459–465. Https://Doi.Org/10.1111/J.1574-6968.1999.Tb13539.X

Caroff, N., Espaze, E., Gautreau, D., Richet, H., & Reynaud, A. (2000). Analysis Of The Effects Of -42 And -32 Ampc Promoter Mutations In Clinical Isolates Of Escherichia Coli Hyperproducing Ampc. *The Journal Of Antimicrobial Chemotherapy*, *45*(6), 783–788. Https://Doi.Org/10.1093/Jac/45.6.783

Chamary, J. V, Parmley, J. L., & Hurst, L. D. (2006). Hearing Silence: Non-Neutral Evolution At Synonymous Sites In Mammals. *Nature Reviews Genetics*, *7*(2), 98–108. Https://Doi.Org/10.1038/Nrg1770

Chattaway, F. W. (1982). Microbial Cell Walls And Membranes: By H J Rogers, H R Perkins And J B Ward. Pp 564. Chapman & Hall, London. 1980 £30. ISBN 0-412-12030-5. *Biochemical Education*, *10*(1), 33. Https://Doi.Org/Https://Doi.Org/10.1016/0307-4412(82)90023-1

Chayot, R., Montagne, B., Mazel, D., & Ricchetti, M. (2010). An End-Joining

Repair Mechanism In Escherichia Coli. *Proceedings Of The National Academy Of Sciences*, *107*(5), 2141–2146. Https://Doi.Org/10.1073/Pnas.0906355107

Chen, K, & Arnold, F. H. (1993). Tuning The Activity Of An Enzyme For Unusual Environments: Sequential Random Mutagenesis Of Subtilisin E For Catalysis In Dimethylformamide. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *90*(12), 5618–5622. Https://Doi.Org/10.1073/Pnas.90.12.5618

Chen, Keqin, & Arnold, F. H. (1991). Enzyme Engineering For Nonaqueous Solvents: Random Mutagenesis To Enhance Activity Of Subtilisin E In Polar Organic Media. *Bio/Technology*, *9*(11), 1073–1077. Https://Doi.Org/10.1038/Nbt1191-1073

Chen, Y., Zhang, W., Shi, Q., Hesek, D., Lee, M., Mobashery, S., & Shoichet, B. K. (2009). Crystal Structures Of Penicillin-Binding Protein 6 From Escherichia Coli. *Journal Of The American Chemical Society*, *131*(40), 14345–14354. Https://Doi.Org/10.1021/Ja903773f

Cherry, J. L. (2010). Highly Expressed And Slowly Evolving Proteins Share Compositional Properties With Thermophilic Proteins. *Molecular Biology And Evolution*, *27*(3), 735–741. Https://Doi.Org/10.1093/Molbev/Msp270

Chetri, S., Bhowmik, D., Paul, D., Pandey, P., Chanda, D. D., Chakravarty, A., … Bhattacharjee, A. (2019). Acrab-Tolc Efflux Pump System Plays A Role In Carbapenem Non-Susceptibility In Escherichia Coli. *BMC Microbiology*, *19*(1), 210. Https://Doi.Org/10.1186/S12866-019-1589-1

Cho, H., Wivagg, C. N., Kapoor, M., Barry, Z., Rohs, P. D. A., Suh, H., … Bernhardt, T. G. (2016). Bacterial Cell Wall Biogenesis Is Mediated By SEDS And PBP Polymerase Families Functioning Semi-Autonomously. *Nature Microbiology*, *1*, 16172. Https://Doi.Org/10.1038/Nmicrobiol.2016.172

Choi, Y. S., Pack, S. P., & Yoo, Y. J. (2005). Development Of A Plasmid Display System Using GAL4 DNA Binding Domain For The In Vitro Screening Of Functional Proteins. *Biotechnology Letters*, *27*(21), 1707–1711. Https://Doi.Org/10.1007/S10529-005-2735-4

Clarke, T. B., Kawai, F., Park, S.-Y., Tame, J. R. H., Dowson, C. G., & Roper, D. I. (2009). Mutational Analysis Of The Substrate Specificity Of Escherichia

Coli Penicillin Binding Protein 4. *Biochemistry*, *48*(12), 2675–2683. Https://Doi.Org/10.1021/Bi801993x

Cohen, S. P., Hächler, H., & Levy, S. B. (1993). Genetic And Functional Analysis Of The Multiple Antibiotic Resistance (Mar) Locus In Escherichia Coli. *Journal Of Bacteriology*, *175*(5), 1484–1492. Https://Doi.Org/10.1128/Jb.175.5.1484-1492.1993

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., … Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819 LP – 823. Https://Doi.Org/10.1126/Science.1231143

Contreras-Llano, L. E., & Tan, C. (2018). High-Throughput Screening Of Biomolecules Using Cell-Free Gene Expression Systems. *Synthetic Biology*, *3*(1). Https://Doi.Org/10.1093/Synbio/Ysy012

Copley, S. D. (2003). Enzymes With Extra Talents: Moonlighting Functions And Catalytic Promiscuity. *Current Opinion In Chemical Biology*, *7*(2), 265–272. Https://Doi.Org/10.1016/S1367-5931(03)00032-2

Copley, S. D. (2020). Evolution Of New Enzymes By Gene Duplication And Divergence. *The FEBS Journal*, *287*(7), 1262–1283. Https://Doi.Org/Https://Doi.Org/10.1111/Febs.15299

Corvec, S., Caroff, N., Espaze, E., Marraillac, J., & Reynaud, A. (2002). -11 Mutation In The Ampc Promoter Increasing Resistance To Beta-Lactams In A Clinical Escherichia Coli Strain. *Antimicrobial Agents And Chemotherapy*, *46*(10), 3265–3267. Https://Doi.Org/10.1128/AAC.46.10.3265-3267.2002

Costa, S., Almeida, A., Castro, A., & Domingues, L. (2014). Fusion Tags For Protein Solubility, Purification And Immunogenicity In Escherichia Coli: The Novel Fh8 System. *Frontiers In Microbiology*, *5*, 63. Https://Doi.Org/10.3389/Fmicb.2014.00063

Crean, R. M., Gardner, J. M., & Kamerlin, S. C. L. (2020). Harnessing Conformational Plasticity To Generate Designer Enzymes. *Journal Of The American Chemical Society*, *142*(26), 11324–11342. Https://Doi.Org/10.1021/Jacs.0c04924

Dai, M., Temirov, J., Pesavento, E., Kiss, C., Velappan, N., Pavlik, P., … Bradbury, A. R. M. (2008). Using T7 Phage Display To Select GFP-Based Binders. *Protein Engineering, Design & Selection : PEDS*, *21*(7), 413–424.

Https://Doi.Org/10.1093/Protein/Gzn016

Daniel, R. A., Drake, S., Buchanan, C. E., Scholle, R., & Errington, J. (1994). The Bacillus Subtilis Spovd Gene Encodes A Mother-Cell-Specific Penicillin-Binding Protein Required For Spore Morphogenesis. *Journal Of Molecular Biology*, *235*(1), 209–220. Https://Doi.Org/10.1016/S0022-2836(05)80027-0

Datz, M., Joris, B., Azab, E. A. M., Galleni, M., Van Beeumen, J., Frère, J.-M., & Martin, H. H. (1994). A Common System Controls The Induction Of Very Different Genes. *European Journal Of Biochemistry*, *226*(1), 149–157. Https://Doi.Org/Https://Doi.Org/10.1111/J.1432-1033.1994.0t149.X

De Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T., & Bukau, B. (2007). Chaperone-Based Procedure To Increase Yields Of Soluble Recombinant Proteins Produced In *E. coli*. *BMC Biotechnology*, *7*(1), 32. Https://Doi.Org/10.1186/1472-6750-7-32

De Visser, J. A. G. M., & Krug, J. (2014). Empirical Fitness Landscapes And The Predictability Of Evolution. *Nature Reviews Genetics*, *15*(7), 480–490. Https://Doi.Org/10.1038/Nrg3744

Dean, C. R., Barkan, D. T., Bermingham, A., Blais, J., Casey, F., Casarez, A., … Wei, J.-R. (2018). Mode Of Action Of The Monobactam LYS228 And Mechanisms Decreasing In Vitro Susceptibility In Escherichia Coli And Klebsiella Pneumoniae. *Antimicrobial Agents And Chemotherapy*, *62*(10). Https://Doi.Org/10.1128/AAC.01200-18

Dejesus, M. A., Gerrick, E. R., Xu, W., Park, S. W., Long, J. E., Boutte, C. C., … Ioerger, T. R. (2017). Comprehensive Essentiality Analysis Of The Mycobacterium Tuberculosis Genome Via Saturating Transposon Mutagenesis. *Mbio*, *8*(1). Https://Doi.Org/10.1128/Mbio.02133-16

Delagrave, S., Murphy, D. J., Pruss, J. L., Maffia, A. M. 3rd, Marrs, B. L., Bylina, E. J., … Youvan, D. C. (2001). Application Of A Very High-Throughput Digital Imaging Screen To Evolve The Enzyme Galactose Oxidase. *Protein Engineering*, *14*(4), 261–267. Https://Doi.Org/10.1093/Protein/14.4.261

Denome, S. A., Elf, P. K., Henderson, T. A., Nelson, D. E., & Young, K. D. (1999). Escherichia Coli Mutants Lacking All Possible Combinations Of Eight Penicillin Binding Proteins: Viability, Characteristics, And Implications For Peptidoglycan Synthesis. *Journal Of Bacteriology*, *181*(13), 3981–3993. Https://Doi.Org/10.1128/JB.181.13.3981-3993.1999

Denoncin, K., & Collet, J.-F. (2013). Disulfide Bond Formation In The Bacterial Periplasm: Major Achievements And Challenges Ahead. *Antioxidants & Redox Signaling*, *19*(1), 63–71. Https://Doi.Org/10.1089/Ars.2012.4864

Depristo, M. A., Weinreich, D. M., & Hartl, D. L. (2005). Missense Meanderings In Sequence Space: A Biophysical View Of Protein Evolution. *Nature Reviews Genetics*, *6*(9), 678–687. Https://Doi.Org/10.1038/Nrg1672

Depuydt, M., Leonard, S. E., Vertommen, D., Denoncin, K., Morsomme, P., Wahni, K., … Collet, J.-F. (2009). A Periplasmic Reducing System Protects Single Cysteine Residues From Oxidation. *Science (New York, N.Y.)*, *326*(5956), 1109–1111. Https://Doi.Org/10.1126/Science.1179557

Dewar, S. J., & Dorazi, R. (2000). Control Of Division Gene Expression In Escherichia Coli. *FEMS Microbiology Letters*, *187*(1), 1–7. Https://Doi.Org/10.1111/J.1574-6968.2000.Tb09127.X

Dodevski, I., Markou, G. C., & Sarkar, C. A. (2015). Conceptual And Methodological Advances In Cell-Free Directed Evolution. *Current Opinion In Structural Biology*, *33*, 1–7. Https://Doi.Org/10.1016/J.Sbi.2015.04.008

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., … Root, D. E. (2016). Optimized Sgrna Design To Maximize Activity And Minimize Off-Target Effects Of CRISPR-Cas9. *Nature Biotechnology*, *34*(2), 184–191. Https://Doi.Org/10.1038/Nbt.3437

Domínguez-García, S., García, C., Quesada, H., & Caballero, A. (2019). Accelerated Inbreeding Depression Suggests Synergistic Epistasis For Deleterious Mutations In Drosophila Melanogaster. *Heredity*, *123*(6), 709–722. Https://Doi.Org/10.1038/S41437-019-0263-6

Dorant, Y., Cayuela, H., Wellband, K., Laporte, M., Rougemont, Q., Mérot, C., … Bernatchez, L. (2020). Copy Number Variants Outperform Snps To Reveal Genotype–Temperature Association In A Marine Species. *Molecular Ecology*, *29*(24), 4765–4782. Https://Doi.Org/Https://Doi.Org/10.1111/Mec.15565

Dörr, T., Moynihan, P. J., & Mayer, C. (2019). Editorial: Bacterial Cell Wall Structure And Dynamics. *Frontiers In Microbiology*, *10*(September), 4–7. Https://Doi.Org/10.3389/Fmicb.2019.02051

Dorrazehi, G. M., Worms, S., Chirakadavil, J. B., Mignolet, J., Hols, P., & Soumillion, P. (2020). *Building Scarless Gene Libraries In The Chromosome Of Bacteria*. 189–211. Https://Doi.Org/10.1007/978-1-0716-0720-6_11

Dower, W. J., & Mattheakis, L. C. (2002). In Vitro Selection As A Powerful Tool For The Applied Evolution Of Proteins And Peptides. *Current Opinion In Chemical Biology*, *6*(3), 390–398. Https://Doi.Org/Https://Doi.Org/10.1016/S1367-5931(02)00332-0

Draghi, J. A., Parsons, T. L., Wagner, G. P., & Plotkin, J. B. (2010). Mutational Robustness Can Facilitate Adaptation. *Nature*, *463*(7279), 353–355. Https://Doi.Org/10.1038/Nature08694

Dramé-Maigné, A., Espada, R., Mccallum, G., Sieskind, R., & Rondelez, Y. (2021). Directed Evolution Of Enzymes Based On In Vitro Programmable Self-Replication. *Biorxiv*, 2021.04.22.440993. Https://Doi.Org/10.1101/2021.04.22.440993

Dramé-Maigné, A., Zadorin, A. S., Golovkova, I., & Rondelez, Y. (2020). Quantifying The Performance Of Micro-Compartmentalized Directed Evolution Protocols. *Life* , Vol. 10. Https://Doi.Org/10.3390/Life10020017

Driessen, A. J. M., Manting, E. H., & Van Der Does, C. (2001). The Structural Basis Of Protein Targeting And Translocation In Bacteria. *Nature Structural Biology*, *8*(6), 492–498. Https://Doi.Org/10.1038/88549

Drummond, D. A., Bloom, J. D., Adami, C., Wilke, C. O., & Arnold, F. H. (2005). Why Highly Expressed Proteins Evolve Slowly. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *102*(40), 14338 LP – 14343. Https://Doi.Org/10.1073/Pnas.0504070102

Dryden, D. T. F., Thomson, A. R., & White, J. H. (2008). How Much Of Protein Sequence Space Has Been Explored By Life On Earth? *Journal Of The Royal Society, Interface*, *5*(25), 953–956. Https://Doi.Org/10.1098/Rsif.2008.0085

Du, S., & Lutkenhaus, J. (2017). Assembly And Activation Of The Escherichia Coli Divisome. *Molecular Microbiology*, *105*(2), 177–187. Https://Doi.Org/10.1111/Mmi.13696

Egan, A. J. F., Errington, J., & Vollmer, W. (2020). Regulation Of Peptidoglycan Synthesis And Remodelling. *Nature Reviews Microbiology*, *18*(8), 446–460. Https://Doi.Org/10.1038/S41579-020-0366-3

Eichenberger, P. (2010). The Red-Ox Status Of A Penicillin-Binding Protein Is An On/Off Switch For Spore Peptidoglycan Synthesis In Bacillus Subtilis. *Molecular Microbiology*, *75*(1), 10–12. Https://Doi.Org/10.1111/J.1365-2958.2009.06963.X

Emami, K., Guyet, A., Kawai, Y., Devi, J., Wu, L. J., Allenby, N., … Errington, J. (2017). Roda As The Missing Glycosyltransferase In Bacillus Subtilis And Antibiotic Discovery For The Peptidoglycan Polymerase Pathway. *Nature Microbiology*, *2*, 16253. Https://Doi.Org/10.1038/Nmicrobiol.2016.253

Emond, S., Petek, M., Kay, E. J., Heames, B., Devenish, S. R. A., Tokuriki, N., & Hollfelder, F. (2020). Accessing Unexplored Regions Of Sequence Space In Directed Enzyme Evolution Via Insertion/Deletion Mutagenesis. *Nature Communications*, *11*(1). Https://Doi.Org/10.1038/S41467-020-17061-3

Esvelt, K. M., Carlson, J. C., & Liu, D. R. (2011). A System For The Continuous Directed Evolution Of Biomolecules. *Nature*, *472*(7344), 499–503. Https://Doi.Org/10.1038/Nature09929

Ewing, B., & Green, P. (1998). Base-Calling Of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Research*, *8*(3), 186–194.

Eyre-Walker, A., & Keightley, P. D. (2007). The Distribution Of Fitness Effects Of New Mutations. *Nature Reviews Genetics*, *8*(8), 610–618. Https://Doi.Org/10.1038/Nrg2146

Ezraty, B., Gennaris, A., Barras, F., & Collet, J.-F. (2017). Oxidative Stress, Protein Damage And Repair In Bacteria. *Nature Reviews Microbiology*, *15*(7), 385–396. Https://Doi.Org/10.1038/Nrmicro.2017.26

Fallah-Araghi, A., Baret, J.-C., Ryckelynck, M., & Griffiths, A. D. (2012). A Completely In Vitro Ultrahigh-Throughput Droplet-Based Microfluidic Screening System For Protein Engineering And Directed Evolution. *Lab On A Chip*, *12*(5), 882–891. Https://Doi.Org/10.1039/C2lc21035e

Fass, D. (2012). Disulfide Bonding In Protein Biophysics. *Annual Review Of Biophysics*, *41*, 63–79. Https://Doi.Org/10.1146/Annurev-Biophys-050511-102321

Fernández, A., Tawfik, D. S., Berkhout, B., Sanders, R., Kloczkowski, A., Sen, T., & Jernigan, B. (2005). Protein Promiscuity: Drug Resistance And Native Functions--HIV-1 Case. *Journal Of Biomolecular Structure & Dynamics*, *22*(6), 615–624. Https://Doi.Org/10.1080/07391102.2005.10531228

Ferrada, E., & Wagner, A. (2008). Protein Robustness Promotes Evolutionary Innovations On Large Evolutionary Time-Scales. *Proceedings Of The Royal Society B: Biological Sciences*, *275*(1643), 1595–1602. Https://Doi.Org/10.1098/Rspb.2007.1617

Feyertag, F., & Alvarez-Ponce, D. (2017). Disulfide Bonds Enable Accelerated Protein Evolution. *Molecular Biology And Evolution*, *34*(8), 1833–1837. Https://Doi.Org/10.1093/Molbev/Msx135

Figueiredo, T. A., Sobral, R. G., Ludovice, A. M., Almeida, J. M. F. De, Bui, N. K., Vollmer, W., … Tomasz, A. (2012). Identification Of Genetic Determinants And Enzymes Involved With The Amidation Of Glutamic Acid Residues In The Peptidoglycan Of Staphylococcus Aureus. *Plos Pathogens*, *8*(1), E1002508. Https://Doi.Org/10.1371/Journal.Ppat.1002508

Firnberg, E., Labonte, J. W., Gray, J. J., & Ostermeier, M. (2014). A Comprehensive, High-Resolution Map Of A Gene's Fitness Landscape. *Molecular Biology And Evolution*, *31*(6), 1581–1592. Https://Doi.Org/10.1093/Molbev/Msu081

Forward, K. R., Willey, B. M., Low, D. E., Mcgeer, A., Kapala, M. A., Kapala, M. M., & Burrows, L. L. (2001). Molecular Mechanisms Of Cefoxitin Resistance In Escherichia Coli From The Toronto Area Hospitals. *Diagnostic Microbiology And Infectious Disease*, *41*(1–2), 57–63. Https://Doi.Org/10.1016/S0732-8893(01)00278-4

Fröhlich, C., Chen, J. Z., Gholipour, S., Erdogan, A. N., & Tokuriki, N. (2021). Evolution Of B-Lactamases And Enzyme Promiscuity. *Protein Engineering, Design And Selection*, *34*. Https://Doi.Org/10.1093/Protein/Gzab013

Fujii, R., Kitaoka, M., & Hayashi, K. (2006). RAISE: A Simple And Novel Method Of Generating Random Insertion And Deletion Mutations. *Nucleic Acids Research*, *34*(4), E30. Https://Doi.Org/10.1093/Nar/Gnj032

Galdadas, I., Lovera, S., Pérez-Hernández, G., Barnes, M. D., Healy, J., Afsharikho, H., … Haider, S. (2018). Defining The Architecture Of KPC-2 Carbapenemase: Identifying Allosteric Networks To Fight Antibiotics Resistance. *Scientific Reports*, *8*(1), 12916. Https://Doi.Org/10.1038/S41598-018-31176-0

Galdadas, I., Qu, S., Oliveira, A. S. F., Olehnovics, E., Mack, A. R., Mojica, M. F., … Haider, S. (2021). Allosteric Communication In Class A B-Lactamases Occurs Via Cooperative Coupling Of Loop Dynamics. *Elife*, *10*, E66567. Https://Doi.Org/10.7554/Elife.66567

Galhardo, R. S., Hastings, P. J., & Rosenberg, S. M. (2007). Mutation As A Stress Response And The Regulation Of Evolvability. *Critical Reviews In*

Biochemistry And Molecular Biology, *42*(5), 399–435. Https://Doi.Org/10.1080/10409230701648502

Galka, P., Jamez, E., Joachim, G., & Soumillion, P. (2017). Quicklib, A Method For Building Fully Synthetic Plasmid Libraries By Seamless Cloning Of Degenerate Oligonucleotides. *PLOS ONE*, *12*(4), E0175146. Retrieved From Https://Doi.Org/10.1371/Journal.Pone.0175146

Galleni, M., Lamotte-Brasseur, J., Rossolini, G. M., Spencer, J., Dideberg, O., Frère, J. M., & Group, M. W. (2001). Standard Numbering Scheme For Class B Beta-Lactamases. *Antimicrobial Agents And Chemotherapy*, *45*(3), 660–663. Https://Doi.Org/10.1128/AAC.45.3.660-663.2001

Garde, S., Chodisetti, P. K., & Reddy, M. (2021). Peptidoglycan: Structure, Synthesis, And Regulation. *Ecosal Plus*, *9*(2). Https://Doi.Org/10.1128/Ecosalplus.ESP-0010-2020

Gargiulo, S., & Soumillion, P. (2021). Directed Evolution For Enzyme Development In Biocatalysis. *Current Opinion In Chemical Biology*, *61*, 107–113. Https://Doi.Org/Https://Doi.Org/10.1016/J.Cbpa.2020.11.006

George, A. M., & Levy, S. B. (1983). Gene In The Major Cotransduction Gap Of The Escherichia Coli K-12 Linkage Map Required For The Expression Of Chromosomal Resistance To Tetracycline And Other Antibiotics. *Journal Of Bacteriology*, *155*(2), 541–548. Https://Doi.Org/10.1128/Jb.155.2.541-548.1983

Georgopapadakou, N. H., & Liu, F. Y. (1980). Penicillin-Binding Proteins In Bacteria. *Antimicrobial Agents And Chemotherapy*, *18*(1), 148–157. Https://Doi.Org/10.1128/AAC.18.1.148

Ghadessy, F. J., Ong, J. L., & Holliger, P. (2001). Directed Evolution Of Polymerase Function By Compartmentalized Self-Replication. *Proceedings Of The National Academy Of Sciences*, *98*(8), 4552 LP – 4557. Https://Doi.Org/10.1073/Pnas.071052198

Ghuysen, J.-M., Charlier, P., Coyette, J., Duez, C., Fonzé, E., Fraipont, C., … Nguyen-Distèche, M. (1996). Penicillin And Beyond: Evolution, Protein Fold, Multimodular Polypeptides, And Multiprotein Complexes. *Microbial Drug Resistance*, *2*(2), 163–175. Https://Doi.Org/10.1089/Mdr.1996.2.163

Gielen, F., Hours, R., Emond, S., Fischlechner, M., Schell, U., & Hollfelder, F. (2016). Ultrahigh-Throughput–Directed Enzyme Evolution By Absorbance-Activated Droplet Sorting (AADS). *Proceedings Of The National*

Academy Of Sciences, *113*(47), E7383 LP-E7389. Https://Doi.Org/10.1073/Pnas.1606927113

Gilbert, H. F. (1990). Molecular And Cellular Aspects Of Thiol-Disulfide Exchange. *Advances In Enzymology And Related Areas Of Molecular Biology*, *63*, 69–172. Https://Doi.Org/10.1002/9780470123096.Ch2

Glauner, B. (1988). Separation And Quantification Of Muropeptides With High-Performance Liquid Chromatography. *Analytical Biochemistry*, *172*(2), 451–464. Https://Doi.Org/10.1016/0003-2697(88)90468-X

Godessart, P., Lannoy, A., Dieu, M., Van Der Verren, S. E., Soumillion, P., Collet, J.-F., … De Bolle, X. (2021). B-Barrels Covalently Link Peptidoglycan And The Outer Membrane In The A-Proteobacterium Brucella Abortus. *Nature Microbiology*, *6*(1), 27–33. Https://Doi.Org/10.1038/S41564-020-00799-3

Goemans, C., Denoncin, K., & Collet, J. F. (2014). Folding Mechanisms Of Periplasmic Proteins. *Biochimica Et Biophysica Acta - Molecular Cell Research*, *1843*(8), 1517–1528. Https://Doi.Org/10.1016/J.Bbamcr.2013.10.014

Goldsmith, M., & Tawfik, D. S. (2017). Enzyme Engineering: Reaching The Maximal Catalytic Efficiency Peak. *Current Opinion In Structural Biology*, *47*, 140–150. Https://Doi.Org/Https://Doi.Org/10.1016/J.Sbi.2017.09.002

González-Leiza, S. M., De Pedro, M. A., & Ayala, J. A. (2011). Amph, A Bifunctional DD-Endopeptidase And DD-Carboxypeptidase Of Escherichia Coli. *Journal Of Bacteriology*, *193*(24), 6887–6894. Https://Doi.Org/10.1128/JB.05764-11

Goymer, P. (2007). Synonymous Mutations Break Their Silence. *Nature Reviews Genetics*, *8*(2), 92. Https://Doi.Org/10.1038/Nrg2056

Granieri, L., Baret, J.-C., Griffiths, A. D., & Merten, C. A. (2010). High-Throughput Screening Of Enzymes By Retroviral Display Using Droplet-Based Microfluidics. *Chemistry & Biology*, *17*(3), 229–235. Https://Doi.Org/Https://Doi.Org/10.1016/J.Chembiol.2010.02.011

Grasso, K. T., & Chatterjee, A. (2021). An Emerging Link Between Robustness And Evolvability In Directed Protein Evolution. *Biochemistry*, *60*(17), 1325–1326. Https://Doi.Org/10.1021/Acs.Biochem.1c00142

Grkovic, S., Brown, M. H., & Skurray, R. A. (2002). Regulation Of Bacterial Drug Export Systems. *Microbiology And Molecular Biology Reviews : MMBR*, *66*(4),

671–701. Https://Doi.Org/10.1128/MMBR.66.4.671-701.2002

Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight Regulation, Modulation, And High-Level Expression By Vectors Containing The Arabinose PBAD Promoter. *Journal Of Bacteriology*, *177*(14), 4121–4130. Https://Doi.Org/10.1128/Jb.177.14.4121-4130.1995

Hall, B. G., Acar, H., Nandipati, A., & Barlow, M. (2014). Growth Rates Made Easy. *Molecular Biology And Evolution*. Https://Doi.Org/10.1093/Molbev/Mst187

Hall, B. G., & Barlow, M. (2004). Evolution Of The Serine Beta-Lactamases: Past, Present And Future. *Drug Resistance Updates : Reviews And Commentaries In Antimicrobial And Anticancer Chemotherapy*, *7*(2), 111–123. Https://Doi.Org/10.1016/J.Drup.2004.02.003

Halperin, S. O., Tou, C. J., Wong, E. B., Modavi, C., Schaffer, D. V, & Dueber, J. E. (2018). CRISPR-Guided DNA Polymerases Enable Diversification Of All Nucleotides In A Tunable Window. *Nature*, *560*(7717), 248–252. Https://Doi.Org/10.1038/S41586-018-0384-8

Hanes, J., Schaffitzel, C., Knappik, A., & Plückthun, A. (2000). Picomolar Affinity Antibodies From A Fully Synthetic Naive Library Selected And Evolved By Ribosome Display. *Nature Biotechnology*, *18*(12), 1287–1292. Https://Doi.Org/10.1038/82407

Hardy, L. W., & Kirsch, J. F. (1984). Diffusion-Limited Component Of Reactions Catalyzed By Bacillus Cereus Beta-Lactamase I. *Biochemistry*, *23*(6), 1275–1282. Https://Doi.Org/10.1021/Bi00301a040

Hatahet, F., Boyd, D., & Beckwith, J. (2014). Disulfide Bond Formation In Prokaryotes: History, Diversity And Design. *Biochimica Et Biophysica Acta*, *1844*(8), 1402–1414. Https://Doi.Org/10.1016/J.Bbapap.2014.02.014

Hatahet, F., & Ruddock, L. W. (2009). Protein Disulfide Isomerase: A Critical Evaluation Of Its Function In Disulfide Bond Formation. *Antioxidants & Redox Signaling*, *11*(11), 2807–2850. Https://Doi.Org/10.1089/Ars.2009.2466

He, M., & Taussig, M. J. (2002). Ribosome Display: Cell-Free Protein Display Technology. *Briefings In Functional Genomics & Proteomics*, *1*(2), 204–212. Https://Doi.Org/10.1093/Bfgp/1.2.204

Henderson, T. A., Templin, M., & Young, K. D. (1995). Identification And

Cloning Of The Gene Encoding Penicillin-Binding Protein 7 Of Escherichia Coli. *Journal Of Bacteriology*, *177*(8), 2074–2079. Https://Doi.Org/10.1128/Jb.177.8.2074-2079.1995

Henderson, T. A., Young, K. D., Denome, S. A., & Elf, P. K. (1997). Ampc And Amph, Proteins Related To The Class C Beta-Lactamases, Bind Penicillin And Contribute To The Normal Morphology Of Escherichia Coli. *Journal Of Bacteriology*, *179*(19), 6112–6121. Https://Doi.Org/10.1128/Jb.179.19.6112-6121.1997

Hogg, P. J. (2003, April 1). Disulfide Bonds As Switches For Protein Function. *Trends In Biochemical Sciences*, Vol. 28, Pp. 210–214. Https://Doi.Org/10.1016/S0968-0004(03)00057-4

Holland-Moritz, D. A., Wismer, M. K., Mann, B. F., Farasat, I., Devine, P., Guetschow, E. D., … Kennedy, R. T. (2020). Mass Activated Droplet Sorting (MADS) Enables High-Throughput Screening Of Enzymatic Reactions At Nanoliter Scale. *Angewandte Chemie International Edition*, *59*(11), 4470–4477. Https://Doi.Org/Https://Doi.Org/10.1002/Anie.201913203

Holstein, J. M., Gylstorff, C., & Hollfelder, F. (2021). Cell-Free Directed Evolution Of A Protease In Microdroplets At Ultrahigh Throughput. *ACS Synthetic Biology*, *10*(2), 252–257. Https://Doi.Org/10.1021/Acssynbio.0c00538

Hosseinzadeh, P., Bhardwaj, G., Mulligan, V. K., Shortridge, M. D., Craven, T. W., Pardo-Avila, F., … Baker, D. (2017). Comprehensive Computational Design Of Ordered Peptide Macrocycles. *Science*, *358*(6369), 1461 LP – 1466. Https://Doi.Org/10.1126/Science.Aap7577

Houlihan, G., Gatti-Lafranconi, P., Kaltenbach, M., Lowe, D., & Hollfelder, F. (2014). An Experimental Framework For Improved Selection Of Binding Proteins Using SNAP Display. *Journal Of Immunological Methods*, *405*, 47–56. Https://Doi.Org/10.1016/J.Jim.2014.01.006

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., … Zhang, F. (2013). DNA Targeting Specificity Of RNA-Guided Cas9 Nucleases. *Nature Biotechnology*, *31*(9), 827–832. Https://Doi.Org/10.1038/Nbt.2647

Hu, T., & Banzhaf, W. (2018). *Neutrality, Robustness, And Evolvability In Genetic Programming BT - Genetic Programming Theory And Practice XIV* (R. Riolo, B.

Worzel, B. Goldman, & B. Tozier, Eds.). Https://Doi.Org/10.1007/978-3-319-97088-2_7

Hughes, R. A., & Ellington, A. D. (2017). Synthetic DNA Synthesis And Assembly: Putting The Synthetic In Synthetic Biology. *Cold Spring Harbor Perspectives In Biology*, *9*(1), A023812. Https://Doi.Org/10.1101/Cshperspect.A023812

Hult, K., & Berglund, P. (2007, May). Enzyme Promiscuity: Mechanism And Applications. *Trends In Biotechnology*, Vol. 25, Pp. 231–238. Https://Doi.Org/10.1016/J.Tibtech.2007.03.002

Ishikawa, H., Kim, S., Kwak, K., Wakasugi, K., & Fayer, M. D. (2007). Disulfide Bond Influence On Protein Structural Dynamics Probed With 2D-IR Vibrational Echo Spectroscopy. *Proceedings Of The National Academy Of Sciences*, *104*(49), 19309 LP – 19314. Https://Doi.Org/10.1073/Pnas.0709760104

Jäckel, C., Kast, P., & Hilvert, D. (2008). Protein Design By Directed Evolution. *Annual Review Of Biophysics*, *37*(1), 153–173. Https://Doi.Org/10.1146/Annurev.Biophys.37.032807.125832

Jacoby, G. A. (2009). Ampc Beta-Lactamases. *Clinical Microbiology Reviews*, *22*(1), 161–182, Table Of Contents. Https://Doi.Org/10.1128/CMR.00036-08

James, L. C., Roversi, P., & Tawfik, D. S. (2003). Antibody Multispecificity Mediated By Conformational Diversity. *Science*, *299*(5611), 1362 LP – 1367. Https://Doi.Org/10.1126/Science.1079731

James, L. C., & Tawfik, D. S. (2003a). Conformational Diversity And Protein Evolution - A 60-Year-Old Hypothesis Revisited. *Trends In Biochemical Sciences*, *28*(7), 361–368. Https://Doi.Org/10.1016/S0968-0004(03)00135-X

James, L. C., & Tawfik, D. S. (2003b). Conformational Diversity And Protein Evolution – A 60-Year-Old Hypothesis Revisited. *Trends In Biochemical Sciences*, *28*(7), 361–368. Https://Doi.Org/Https://Doi.Org/10.1016/S0968-0004(03)00135-X

Jaurin, B, Grundström, T., & Normark, S. (1982). Sequence Elements Determining Ampc Promoter Strength In *E. coli*. *The EMBO Journal*, *1*(7), 875–881. Retrieved From Https://Pubmed.Ncbi.Nlm.Nih.Gov/6329713

Jaurin, Bengtåke, Grundström, T., Edlund, T., & Normark, S. (1981). The *E. coli*

B-Lactamase Attenuator Mediates Growth Rate-Dependent Regulation. *Nature*, *290*(5803), 221–225. Https://Doi.Org/10.1038/290221a0

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., & Yang, S. (2015). Multigene Editing In The Escherichia Coli Genome Via The CRISPR-Cas9 System. *Applied And Environmental Microbiology*, *81*(7), 2506–2514. Https://Doi.Org/10.1128/Aem.04023-14

Joachim, G. (2016). *Directed Evolution Of A Beta-Lactamase With A Shorter Omega-Loop* (Université Catholique De Louvain). Retrieved From Http://Hdl.Handle.Net/2078.1/182851

Joris, B., Ghuysen, J. M., Dive, G., Renard, A., Dideberg, O., Charlier, P., … Al., Et. (1988). The Active-Site-Serine Penicillin-Recognizing Enzymes As Members Of The Streptomyces R61 DD-Peptidase Family. *The Biochemical Journal*, *250*(2), 313–324. Https://Doi.Org/10.1042/Bj2500313

Jürgens, U. J., Drews, G., & Weckesser, J. (1983). Primary Structure Of The Peptidoglycan From The Unicellular Cyanobacterium Synechocystis Sp. Strain PCC 6714. *Journal Of Bacteriology*, *154*(1), 471–478. Https://Doi.Org/10.1128/Jb.154.1.471-478.1983

Kaltenbach, M., Stein, V., & Hollfelder, F. (2011). SNAP Dendrimers: Multivalent Protein Display On Dendrimer-Like DNA For Directed Evolution. *Chembiochem : A European Journal Of Chemical Biology*, *12*(14), 2208–2216. Https://Doi.Org/10.1002/Cbic.201100240

Kamio, Y., & Nikaido, H. (1976). Outer Membrane Of Salmonella Typhimurium: Accessibility Of Phospholipid Head Groups To Phospholipase C And Cyanogen Bromide Activated Dextran In The External Medium. *Biochemistry*, *15*(12), 2561–2570. Https://Doi.Org/10.1021/Bi00657a012

Kellner, R., Malempré, R., Vandenameele, J., Brans, A., Hennen, A.-F., Rochus, N., … Matagne, A. (2021). Protein Formulation Through Automated Screening Of Ph And Buffer Conditions, Using The Robotein® High Throughput Facility. *European Biophysics Journal*, *50*(3), 473–490. Https://Doi.Org/10.1007/S00249-021-01510-Y

Kelly, J. A., Dideberg, O., Charlier, P., Wery, J. P., Libert, M., Moews, P. C., … Et, Al. (1986). On The Origin Of Bacterial Resistance To Penicillin: Comparison Of A Beta-Lactamase And A Penicillin Target. *Science*, *231*(4744), 1429 LP – 1431. Https://Doi.Org/10.1126/Science.3082007

Keshri, V., Chabrière, E., Pinault, L., Colson, P., Diene, S. M., Rolain, J.-M., … Pontarotti, P. (2020). Promiscuous Enzyme Activity As A Driver Of Allo And Iso Convergent Evolution, Lessons From The B-Lactamases. *International Journal Of Molecular Sciences*, *21*(17). Https://Doi.Org/10.3390/Ijms21176260

Khersonsky, O., & Tawfik, D. S. (2010, July 7). Enzyme Promiscuity: A Mechanistic And Evolutionary Perspective. *Annual Review Of Biochemistry*, Vol. 79, Pp. 471–505. Https://Doi.Org/10.1146/Annurev-Biochem-030409-143718

Kimchi-Sarfaty, C., Oh, J. M., Kim, I.-W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V, & Gottesman, M. M. (2007). A Silent Polymorphism In The MDR1 Gene Changes Substrate Specificity. *Science*, *315*(5811), 525 LP – 528. Https://Doi.Org/10.1126/Science.1135308

Kintses, B., Hein, C., Mohamed, M. F., Fischlechner, M., Courtois, F., Lainé, C., & Hollfelder, F. (2012). Picoliter Cell Lysate Assays In Microfluidic Droplet Compartments For Directed Enzyme Evolution. *Chemistry & Biology*, *19*(8), 1001–1009. Https://Doi.Org/10.1016/J.Chembiol.2012.06.009

Kirschner, M., Gerhart, J., & Gerhartt, J. (2010). Perspective Evolvability. *Sciences-New York*, *95*(15), 8420–8427. Retrieved From Www.Pnas.Org.

Klein, J. C., Lajoie, M. J., Schwartz, J. J., Strauch, E.-M., Nelson, J., Baker, D., & Shendure, J. (2016). Multiplex Pairwise Assembly Of Array-Derived DNA Oligonucleotides. *Nucleic Acids Research*, *44*(5), E43–E43. Https://Doi.Org/10.1093/Nar/Gkv1177

Knappik, A., Krebber, C., & Plückthun, A. (1993). The Effect Of Folding Catalysts On The In Vivo Folding Process Of Different Antibody Fragments Expressed In Escherichia Coli. *Bio/Technology*, *11*(1), 77–83. Https://Doi.Org/10.1038/Nbt0193-77

Knies, J. L., Cai, F., & Weinreich, D. M. (2017). Enzyme Efficiency But Not Thermostability Drives Cefotaxime Resistance Evolution In TEM-1 B-Lactamase. *Molecular Biology And Evolution*, *34*(5), 1040–1054. Https://Doi.Org/10.1093/Molbev/Msx053

Knöppel, A., Knopp, M., Albrecht, L. M., Lundin, E., Lustig, U., Näsvall, J., & Andersson, D. I. (2018). Genetic Adaptation To Growth Under Laboratory Conditions In Escherichia Coli And Salmonella Enterica. *Frontiers In Microbiology*, *9*, 756. Https://Doi.Org/10.3389/Fmicb.2018.00756

Koch, A. L. (2000). Penicillin Binding Proteins, Beta-Lactams, And Lactamases: Offensives, Attacks, And Defensive Countermeasures. *Critical Reviews In Microbiology*, *26*(4), 205–220. Https://Doi.Org/10.1080/10408410091154228

Kohanski, M. A., Dwyer, D. J., Hayete, B., Lawrence, C. A., & Collins, J. J. (2007). A Common Mechanism Of Cellular Death Induced By Bactericidal Antibiotics. *Cell*, *130*(5), 797–810. Https://Doi.Org/10.1016/J.Cell.2007.06.049

Korat, B., Mottl, H., & Keck, W. (1991). Penicillin-Binding Protein 4 Of Escherichia Coli: Molecular Cloning Of The Dacb Gene, Controlled Overexpression, And Alterations In Murein Composition. *Molecular Microbiology*, *5*(3), 675–684. Https://Doi.Org/10.1111/J.1365-2958.1991.Tb00739.X

Kram, K. E., Geiger, C., Ismail, W. M., Lee, H., Tang, H., Foster, P. L., & Finkel, S. E. (2017). Adaptation Of Escherichia Coli To Long-Term Serial Passage In Complex Medium: Evidence Of Parallel Evolution. *Msystems*, *2*(2), E00192-16. Https://Doi.Org/10.1128/Msystems.00192-16

Kuchner, O., & Arnold, F. H. (1997). Directed Evolution Of Enzyme Catalysts. *Trends In Biotechnology*, *15*(12), 523–530. Https://Doi.Org/Https://Doi.Org/10.1016/S0167-7799(97)01138-4

Kuru, E., Lambert, C., Rittichier, J., Till, R., Ducret, A., Derouaux, A., … Sockett, R. E. (2017). Fluorescent D-Amino-Acids Reveal Bi-Cellular Cell Wall Modifications Important For Bdellovibrio Bacteriovorus Predation. *Nature Microbiology*, *2*(12), 1648–1657. Https://Doi.Org/10.1038/S41564-017-0029-Y

Labarbe, C. (2006). *Vers L'évolution D'une DD-Peptidase En Beta-Lactamase* (Université Catholique De Louvain). Retrieved From Http://Hdl.Handle.Net/2078.1/5337

Laminet, A. A., & Plückthun, A. (1989). The Precursor Of Beta-Lactamase: Purification, Properties And Folding Kinetics. *The EMBO Journal*, *8*(5), 1469–1477. Retrieved From Https://Pubmed.Ncbi.Nlm.Nih.Gov/2670555

Lamminmäki, U., Paupério, S., Westerlund-Karlsson, A., Karvinen, J., Virtanen, P. L., Lövgren, T., & Saviranta, P. (1999). Expanding The Conformational Diversity By Random Insertions To CDRH2 Results In Improved Anti-

Estradiol Antibodies. *Journal Of Molecular Biology*, *291*(3), 589–602. Https://Doi.Org/10.1006/Jmbi.1999.2981

Landeta, C., Boyd, D., & Beckwith, J. (2018). Disulfide Bond Formation In Prokaryotes. *Nature Microbiology*, *3*(3), 270–280. Https://Doi.Org/10.1038/S41564-017-0106-2

Lauritsen, I., Porse, A., Sommer, M. O. A., & Nørholm, M. H. H. (2017). A Versatile One-Step CRISPR-Cas9 Based Approach To Plasmid-Curing. *Microbial Cell Factories*, *16*(1), 135. Https://Doi.Org/10.1186/S12934-017-0748-Z

Lavallie, E. R., & Mccoy, J. M. (1995). Gene Fusion Expression Systems In Escherichia Coli. *Current Opinion In Biotechnology*, *6*(5), 501–506. Https://Doi.Org/10.1016/0958-1669(95)80083-2

Lee, M. E., Deloache, W. C., Cervantes, B., & Dueber, J. E. (2015). A Highly Characterized Yeast Toolkit For Modular, Multipart Assembly. *ACS Synthetic Biology*, *4*(9), 975–986. Https://Doi.Org/10.1021/Sb500366v

Lee, S. F., & Davey, L. (2017). Disulfide Bonds: A Key Modification In Bacterial Extracytoplasmic Proteins. *Journal Of Dental Research*, *96*(13), 1465–1473. Https://Doi.Org/10.1177/0022034517725059

Lee, S. Y., Choi, J. H., & Xu, Z. (2003). Microbial Cell-Surface Display. *Trends In Biotechnology*, *21*(1), 45–52. Https://Doi.Org/Https://Doi.Org/10.1016/S0167-7799(02)00006-9

Lee, T. S., Krupa, R. A., Zhang, F., Hajimorad, M., Holtz, W. J., Prasad, N., … Keasling, J. D. (2011). Bglbrick Vectors And Datasheets: A Synthetic Biology Platform For Gene Expression. *Journal Of Biological Engineering*, *5*(1), 12. Https://Doi.Org/10.1186/1754-1611-5-12

Lenski, R. E. (1998). Bacterial Evolution And The Cost Of Antibiotic Resistance. *International Microbiology : The Official Journal Of The Spanish Society For Microbiology*, *1*(4), 265–270.

Li, A., Qu, G., Sun, Z., & Reetz, M. T. (2019). Statistical Analysis Of The Benefits Of Focused Saturation Mutagenesis In Directed Evolution Based On Reduced Amino Acid Alphabets. *ACS Catalysis*, *9*(9), 7769–7778. Https://Doi.Org/10.1021/Acscatal.9b02548

Liao, Q., Kulkarni, Y., Sengupta, U., Petrović, D., Mulholland, A. J., Van Der Kamp, M. W., … Kamerlin, S. C. L. (2018). Loop Motion In

Triosephosphate Isomerase Is Not A Simple Open And Shut Case. *Journal Of The American Chemical Society*, *140*(46), 15889–15903. Https://Doi.Org/10.1021/Jacs.8b09378

Liu, Yao, & Breukink, E. (2016). The Membrane Steps Of Bacterial Cell Wall Synthesis As Antibiotic Targets. *Antibiotics (Basel, Switzerland)*, *5*(3). Https://Doi.Org/10.3390/Antibiotics5030028

Liu, Yiming, Carlsson Möller, M., Petersen, L., Söderberg, C. A. G., & Hederstedt, L. (2010). Penicillin-Binding Protein Spovd Disulphide Is A Target For Stoa In Bacillus Subtilis Forespores. *Molecular Microbiology*, *75*(1), 46–60. Https://Doi.Org/10.1111/J.1365-2958.2009.06964.X

Loewe, L., & Charlesworth, B. (2006). Inferring The Distribution Of Mutational Effects On Fitness In Drosophila. *Biology Letters*, *2*(3), 426–430. Https://Doi.Org/10.1098/Rsbl.2006.0481

Luk, L. Y. P., Javier Ruiz-Pernía, J., Dawson, W. M., Roca, M., Loveridge, E. J., Glowacki, D. R., … Allemann, R. K. (2013). Unraveling The Role Of Protein Dynamics In Dihydrofolate Reductase Catalysis. *Proceedings Of The National Academy Of Sciences*, *110*(41), 16344 LP – 16349. Https://Doi.Org/10.1073/Pnas.1312437110

Macbeath, G., Kast, P., & Hilvert, D. (1998). Redesigning Enzyme Topology By Directed Evolution. *Science*, *279*(5358), 1958 LP – 1961. Https://Doi.Org/10.1126/Science.279.5358.1958

Macheboeuf, P., Contreras-Martel, C., Job, V., Dideberg, O., & Dessen, A. (2006). Penicillin Binding Proteins: Key Players In Bacterial Cell Cycle And Drug Resistance Processes. *FEMS Microbiology Reviews*, *30*(5), 673–691. Https://Doi.Org/10.1111/J.1574-6976.2006.00024.X

Maclean, R. C., & Millan, A. S. (2019). The Evolution Of Antibiotic Resistance. *Science*, *365*(6458), 1082–1083. Https://Doi.Org/10.1126/Science.Aax3879

Magnet, S., Bellais, S., Dubost, L., Fourgeaud, M., Mainardi, J. L., Petit-Frère, S., … Gutmann, L. (2007). Identification Of The L,D-Transpeptidases Responsible For Attachment Of The Braun Lipoprotein To Escherichia Coli Peptidoglycan. *Journal Of Bacteriology*, *189*(10), 3927–3931. Https://Doi.Org/10.1128/JB.00084-07

Majiduddin, F. K., & Palzkill, T. (2003). Amino Acid Sequence Requirements At Residues 69 And 238 For The SME-1 B-Lactamase To Confer Resistance

To B-Lactam Antibiotics. *Antimicrobial Agents And Chemotherapy*, *47*(3), 1062–1067. Https://Doi.Org/10.1128/AAC.47.3.1062-1067.2003

Martínez, J. L., Baquero, F., & Andersson, D. I. (2007). Predicting Antibiotic Resistance. *Nature Reviews. Microbiology*, *5*(12), 958–965. Https://Doi.Org/10.1038/Nrmicro1796

Masel, J., & Siegal, M. L. (2009). Robustness: Mechanisms And Consequences. *Trends In Genetics*, *25*(9), 395–403. Https://Doi.Org/10.1016/J.Tig.2009.07.005

Masel, J., & Trotter, M. V. (2010). Robustness And Evolvability. *Trends In Genetics*, *26*(9), 406–414. Https://Doi.Org/Https://Doi.Org/10.1016/J.Tig.2010.06.002

Matagne, A., & Frère, J. M. (1995). Contribution Of Mutant Analysis To The Understanding Of Enzyme Catalysis: The Case Of Class A Beta-Lactamases. *Biochimica Et Biophysica Acta*, *1246*(2), 109–127. Https://Doi.Org/10.1016/0167-4838(94)00177-I

Matagne, A., Lamotte-Brasseur, J., & Frère, J. M. (1998). Catalytic Properties Of Class A Beta-Lactamases: Efficiency And Diversity. *The Biochemical Journal*, *330 (Pt 2*(Pt 2), 581–598. Https://Doi.Org/10.1042/Bj3300581

Mathelié-Guinlet, M., Asmar, A. T., Collet, J. F., & Dufrêne, Y. F. (2020). Lipoprotein Lpp Regulates The Mechanical Properties Of The *E. coli* Cell Envelope. *Nature Communications*, *11*(1). Https://Doi.Org/10.1038/S41467-020-15489-1

May, K. L., & Grabowicz, M. (2018). The Bacterial Outer Membrane Is An Evolving Antibiotic Barrier. *Proceedings Of The National Academy Of Sciences*, *115*(36), 8852 LP – 8854. Https://Doi.Org/10.1073/Pnas.1812779115

Maynard Smith, J. (1970). Natural Selection And The Concept Of A Protein Space. *Nature*, *225*(5232), 563–564. Https://Doi.Org/10.1038/225563a0

Mcdonald, J. F., & Ayala, F. J. (1978). Gene Regulation In Adaptive Evolution. *Canadian Journal Of Genetics And Cytology. Journal Canadien De Genetique Et De Cytologie*, *20*(2), 159–175. Https://Doi.Org/10.1139/G78-018

Medeiros, A. A. (1997). Evolution And Dissemination Of Beta-Lactamases Accelerated By Generations Of Beta-Lactam Antibiotics. *Clinical Infectious Diseases : An Official Publication Of The Infectious Diseases Society Of America*, *24 Suppl 1*, S19-45. Https://Doi.Org/10.1093/Clinids/24.Supplement_1.S19

Mehlhoff, J. D., Stearns, F. W., Rohm, D., Wang, B., Tsou, E.-Y., Dutta, N., … Ostermeier, M. (2020). Collateral Fitness Effects Of Mutations. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *117*(21), 11597–11607. Https://Doi.Org/10.1073/Pnas.1918680117

Merdanovic, M., Clausen, T., Kaiser, M., Huber, R., & Ehrmann, M. (2011). Protein Quality Control In The Bacterial Periplasm. *Annual Review Of Microbiology*, *65*(1), 149–168. Https://Doi.Org/10.1146/Annurev-Micro-090110-102925

Meroueh, S. O., Minasov, G., Lee, W., Shoichet, B. K., & Mobashery, S. (2003). Structural Aspects For Evolution Of B-Lactamases From Penicillin-Binding Proteins. *Journal Of The American Chemical Society*, *125*(32), 9612–9618. Https://Doi.Org/10.1021/Ja034861u

Messens, J., & Collet, J.-F. (2006). Pathways Of Disulfide Bond Formation In Escherichia Coli. *The International Journal Of Biochemistry & Cell Biology*, *38*(7), 1050–1062. Https://Doi.Org/Https://Doi.Org/10.1016/J.Biocel.2005.12.011

Messens, J., & Collet, J.-F. (2013). Thiol–Disulfide Exchange In Signaling: Disulfide Bonds As A Switch. *Antioxidants & Redox Signaling*, *18*(13), 1594–1596. Https://Doi.Org/10.1089/Ars.2012.5156

Millan, A. S., Peña-Miller, R., Toll-Riera, M., Halbert, Z. V, Mclean, A. R., Cooper, B. S., & Maclean, R. C. (2014). Positive Selection And Compensatory Adaptation Interact To Stabilize Non-Transmissible Plasmids. *Nature Communications*, *5*(1), 5208. Https://Doi.Org/10.1038/Ncomms6208

Miot, M., & Betton, J.-M. (2004). Protein Quality Control In The Bacterial Periplasm. *Microbial Cell Factories*, *3*(1), 4. Https://Doi.Org/10.1186/1475-2859-3-4

Modi, T., Risso, V. A., Martinez-Rodriguez, S., Gavira, J. A., Mebrat, M. D., Van Horn, W. D., … Banu Ozkan, S. (2021). Hinge-Shift Mechanism As A Protein Design Principle For The Evolution Of B-Lactamases From Substrate Promiscuity To Specificity. *Nature Communications*, *12*(1), 1852. Https://Doi.Org/10.1038/S41467-021-22089-0

Moon, T. M., D'Andréa, É. D., Lee, C. W., Soares, A., Jakoncic, J., Desbonnet, C., … Peti, W. (2018). The Structures Of Penicillin-Binding Protein 4 (PBP4) And PBP5 From Enterococci Provide Structural Insights Into B-

Lactam Resistance. *Journal Of Biological Chemistry*, *293*(48), 18574–18584. Https://Doi.Org/Https://Doi.Org/10.1074/Jbc.RA118.006052

Morè, N., Martorana, A. M., Biboy, J., Otten, C., Winkle, M., Serrano, C. K. G., … Polissi, A. (2019). Peptidoglycan Remodeling Enables Escherichia Coli To Survive Severe Outer Membrane Assembly Defect. *Mbio*, *10*(1). Https://Doi.Org/10.1128/Mbio.02729-18

Morlot, C., Straume, D., Peters, K., Hegnar, O. A., Simon, N., Villard, A.-M., … Zapun, A. (2018). Structure Of The Essential Peptidoglycan Amidotransferase Murt/Gatd Complex From Streptococcus Pneumoniae. *Nature Communications*, *9*(1), 3180. Https://Doi.Org/10.1038/S41467-018-05602-W

Mousseau, T. A., & Roff, D. A. (1987). Natural Selection And The Heritability Of Fitness Components. *Heredity*, *59*(2), 181–197. Https://Doi.Org/10.1038/Hdy.1987.113

Muheim, C., Götzke, H., Eriksson, A. U., Lindberg, S., Lauritsen, I., Nørholm, M. H. H., & Daley, D. O. (2017). Increasing The Permeability Of Escherichia Coli Using MAC13243. *Scientific Reports*, *7*(1), 17629. Https://Doi.Org/10.1038/S41598-017-17772-6

Münch, D., Roemer, T., Lee, S. H., Engeser, M., Sahl, H. G., & Schneider, T. (2012). Identification And In Vitro Analysis Of The Gatd/Murt Enzyme-Complex Catalyzing Lipid II Amidation In Staphylococcus Aureus. *PLOS Pathogens*, *8*(1), E1002509. Retrieved From Https://Doi.Org/10.1371/Journal.Ppat.1002509

Myhre, A. E., Stuestøl, J. F., K., D. M., Gunhild, Ø., Christoph, T., J., F. S., … E., W. J. (2004). Organ Injury And Cytokine Release Caused By Peptidoglycan Are Dependent On The Structural Integrity Of The Glycan Chain. *Infection And Immunity*, *72*(3), 1311–1317. Https://Doi.Org/10.1128/IAI.72.3.1311-1317.2004

Naas, T., Oueslati, S., Bonnin, R. A., Dabos, M. L., Zavala, A., Dortet, L., … Iorga, B. I. (2017). Beta-Lactamase Database (BLDB) - Structure And Function. *Journal Of Enzyme Inhibition And Medicinal Chemistry*, *32*(1), 917–919. Https://Doi.Org/10.1080/14756366.2017.1344235

Nadler, F., Bracharz, F., & Kabisch, J. (2019). Copyswitch—In Vivo Optimization Of Gene Copy Numbers For Heterologous Gene Expression In Bacillus Subtilis . *Frontiers In Bioengineering And Biotechnology*

, Vol. 6, P. 207. Retrieved From Https://Www.Frontiersin.Org/Article/10.3389/Fbioe.2018.00207

Nelson, D. E., & Young, K. D. (2001). Contributions Of PBP 5 And DD-Carboxypeptidase Penicillin Binding Proteins To Maintenance Of Cell Shape In Escherichia Coli. *Journal Of Bacteriology*, *183*(10), 3055–3064. Https://Doi.Org/10.1128/JB.183.10.3055-3064.2001

Nelson, E. C., & Elisha, B. G. (1999). Molecular Basis Of Ampc Hyperproduction In Clinical Isolates Of Escherichia Coli. *Antimicrobial Agents And Chemotherapy*, *43*(4), 957–959. Https://Doi.Org/10.1128/AAC.43.4.957

Neumann-Staubitz, P., & Neumann, H. (2016). The Use Of Unnatural Amino Acids To Study And Engineer Protein Function. *Current Opinion In Structural Biology*, *38*, 119–128. Https://Doi.Org/Https://Doi.Org/10.1016/J.Sbi.2016.06.006

Nikaido, H. (2003a). Molecular Basis Of Bacterial Outer Membrane Permeability Revisited. *Microbiology And Molecular Biology Reviews : MMBR*, *67*(4), 593–656. Https://Doi.Org/10.1128/MMBR.67.4.593-656.2003

Nikaido, H. (2003b). Molecular Basis Of Bacterial Outer Membrane Permeability Revisited. *Microbiology And Molecular Biology Reviews*, *67*(4), 593–656. Https://Doi.Org/10.1128/Mmbr.67.4.593-656.2003

Nikaido, H., & Normark, S. (1987). Sensitivity Of Escherichia Coli To Various Beta-Lactams Is Determined By The Interplay Of Outer Membrane Permeability And Degradation By Periplasmic Beta-Lactamases: A Quantitative Predictive Treatment. *Molecular Microbiology*, *1*(1), 29–36. Https://Doi.Org/10.1111/J.1365-2958.1987.Tb00523.X

Nöldeke, E. R., & Stehle, T. (2019). Unraveling The Mechanism Of Peptidoglycan Amidation By The Bifunctional Enzyme Complex Gatd/Murt: A Comparative Structural Approach. *International Journal Of Medical Microbiology : IJMM*, *309*(6), 151334. Https://Doi.Org/10.1016/J.Ijmm.2019.151334

O'Brien, P. J., & Herschlag, D. (1999). Catalytic Promiscuity And The Evolution Of New Enzymatic Activities. *Chemistry & Biology*, *6*(4), R91–R105. Https://Doi.Org/Https://Doi.Org/10.1016/S1074-5521(99)80033-7

O'Loughlin, T. L., Greene, D. N., & Matsumura, I. (2006). Diversification And

Specialization Of HIV Protease Function During In Vitro Evolution. *Molecular Biology And Evolution*, *23*(4), 764–772. Https://Doi.Org/10.1093/Molbev/Msj098

Otten, R., Liu, L., Kenner, L. R., Clarkson, M. W., Mavor, D., Tawfik, D. S., … Fraser, J. S. (2018). Rescue Of Conformational Dynamics In Enzyme Catalysis By Directed Evolution. *Nature Communications*, *9*(1), 1314. Https://Doi.Org/10.1038/S41467-018-03562-9

Pace, C. N., Grimsley, G. R., Thomson, J. A., & Barnett, B. J. (1988). Conformational Stability And Activity Of Ribonuclease T1 With Zero, One, And Two Intact Disulfide Bonds. *Journal Of Biological Chemistry*, *263*(24), 11820–11825. Https://Doi.Org/Https://Doi.Org/10.1016/S0021-9258(18)37859-1

Packiam, K. A. R., Ramanan, R. N., Ooi, C. W., Krishnaswamy, L., & Tey, B. T. (2020). Stepwise Optimization Of Recombinant Protein Production In Escherichia Coli Utilizing Computational And Experimental Approaches. *Applied Microbiology And Biotechnology*, *104*(8), 3253–3266. Https://Doi.Org/10.1007/S00253-020-10454-W

Pál, C., Papp, B., & Hurst, L. D. (2001). Highly Expressed Genes In Yeast Evolve Slowly. *Genetics*, *158*(2), 927–931. Https://Doi.Org/10.1093/Genetics/158.2.927

Palzkill, T. (2013). Metallo-B-Lactamase Structure And Function. *Annals Of The New York Academy Of Sciences*, *1277*, 91–104. Https://Doi.Org/10.1111/J.1749-6632.2012.06796.X

Pandya, C., Farelli, J. D., Dunaway-Mariano, D., & Allen, K. N. (2014). Enzyme Promiscuity: Engine Of Evolutionary Innovation*. *Journal Of Biological Chemistry*, *289*(44), 30229–30236. Https://Doi.Org/Https://Doi.Org/10.1074/Jbc.R114.572990

Park, H., & Kim, S. (2021). Gene-Specific Mutagenesis Enables Rapid Continuous Evolution Of Enzymes In Vivo. *Nucleic Acids Research*, *49*(6), E32. Https://Doi.Org/10.1093/Nar/Gkaa1231

Passam, F. J., & Chiu, J. (2019). Allosteric Disulphide Bonds As Reversible Mechano-Sensitive Switches That Control Protein Functions In The Vasculature. *Biophysical Reviews*, *11*(3), 419–430. Https://Doi.Org/10.1007/S12551-019-00543-0

Payne, J. L., & Wagner, A. (2019). The Causes Of Evolvability And Their

Evolution. *Nature Reviews Genetics*, *20*(1), 24–38. Https://Doi.Org/10.1038/S41576-018-0069-Z

Pazos, M., & Peters, K. (2019a). Peptidoglycan. *Sub-Cellular Biochemistry*, *92*, 127–168. Https://Doi.Org/10.1007/978-3-030-18768-2_5

Pazos, M., & Peters, K. (2019b). *Peptidoglycan BT - Bacterial Cell Walls And Membranes* (A. Kuhn, Ed.). Https://Doi.Org/10.1007/978-3-030-18768-2_5

Peter-Getzlaff, S., Polsfuss, S., Poledica, M., Hombach, M., Giger, J., Böttger, E. C., … Bloemberg, G. V. (2011). Detection Of Ampc Beta-Lactamase In Escherichia Coli: Comparison Of Three Phenotypic Confirmation Assays And Genetic Analysis. *Journal Of Clinical Microbiology*, *49*(8), 2924–2932. Https://Doi.Org/10.1128/JCM.00091-11

Petrović, D., Risso, V. A., Kamerlin, S. C. L., & Sanchez-Ruiz, J. M. (2018). Conformational Dynamics And Enzyme Evolution. *Journal Of The Royal Society Interface*, *15*(144), 20180330. Https://Doi.Org/10.1098/Rsif.2018.0330

Philippon, A., Slama, P., Dény, P., & Labia, R. (2016, January 1). A Structure-Based Classification Of Class A B-Lactamases, A Broadly Diverse Family Of Enzymes. *Clinical Microbiology Reviews*, Vol. 29, Pp. 29–57. Https://Doi.Org/10.1128/CMR.00019-15

Plotkin, J. B., & Kudla, G. (2011). Synonymous But Not The Same: The Causes And Consequences Of Codon Bias. *Nature Reviews Genetics*, *12*(1), 32–42. Https://Doi.Org/10.1038/Nrg2899

Plückthun, A. (2012). *Ribosome Display: A Perspective* (J. A. Douthwaite & R. H. Jackson, Eds.). Https://Doi.Org/10.1007/978-1-61779-379-0_1

Poon, A., & Chao, L. (2005). The Rate Of Compensatory Mutation In The DNA Bacteriophage Phix174. *Genetics*, *170*(3), 989–999. Https://Doi.Org/10.1534/Genetics.104.039438

Popham, D. L., & Young, K. D. (2003). Role Of Penicillin-Binding Proteins In Bacterial Cell Morphogenesis. *Current Opinion In Microbiology*, *6*(6), 594–599. Https://Doi.Org/Https://Doi.Org/10.1016/J.Mib.2003.10.002

Püllmann, P., Ulpinnis, C., Marillonnet, S., Gruetzner, R., Neumann, S., & Weissenborn, M. J. (2019). Golden Mutagenesis: An Efficient Multi-Site-Saturation Mutagenesis Approach By Golden Gate Cloning With

Automated Primer Design. *Scientific Reports*, *9*(1), 10932. Https://Doi.Org/10.1038/S41598-019-47376-1

Qi, D., & Scholthof, K.-B. G. (2008). A One-Step PCR-Based Method For Rapid And Efficient Site-Directed Fragment Deletion, Insertion, And Substitution Mutagenesis. *Journal Of Virological Methods*, *149*(1), 85–90. Https://Doi.Org/10.1016/J.Jviromet.2008.01.002

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome Engineering Using The CRISPR-Cas9 System. *Nature Protocols*, *8*, 2281.

Raquet, X., Lamotte-Brasseur, J., Bouillenne, F., & Frere, J.-M. (1997). A Disulfide Bridge Near The Active Site Of Carbapenem-Hydrolyzing Class A-Lactamases Might Explain Their Unusual Substrate Profile. In *Proteins* (Vol. 27). Wiley-Liss, Inc.

Ravikumar, A., Arrieta, A., & Liu, C. C. (2014). An Orthogonal DNA Replication System In Yeast. *Nature Chemical Biology*, *10*(3), 175–177. Https://Doi.Org/10.1038/Nchembio.1439

Ravikumar, A., Arzumanyan, G. A., Obadi, M. K. A., Javanpour, A. A., & Liu, C. C. (2018). Scalable, Continuous Evolution Of Genes At Mutation Rates Above Genomic Error Thresholds. *Cell*, *175*(7), 1946-1957.E13. Https://Doi.Org/10.1016/J.Cell.2018.10.021

Reardon-Robinson, M. E., & Ton-That, H. (2015). Disulfide-Bond-Forming Pathways In Gram-Positive Bacteria. *Journal Of Bacteriology*, *198*(5), 746–754. Https://Doi.Org/10.1128/JB.00769-15

Reetz, M. T., Kahakeaw, D., & Lohmer, R. (2008). Addressing The Numbers Problem In Directed Evolution. *Chembiochem : A European Journal Of Chemical Biology*, *9*(11), 1797–1804. Https://Doi.Org/10.1002/Cbic.200800298

Regeimbal, J., & Bardwell, J. C. A. (2002). *7 - Disulfide Bond Formation In Prokaryotes And Eukaryotes* (R. E. Dalbey & G. B. T.-P. T. Von Heijne Transport, And Translocation, Eds.). Https://Doi.Org/Https://Doi.Org/10.1016/B978-012200731-6.50009-4

Rhazi, N., Galleni, M., Page, M. I., & Frère, J. M. (1999). Peptidase Activity Of Beta-Lactamases. *The Biochemical Journal*, *341 (Pt 2*(Pt 2), 409–413. Retrieved From Https://Pubmed.Ncbi.Nlm.Nih.Gov/10393100

Rix, G., & Liu, C. C. (2021). Systems For In Vivo Hypermutation: A Quest For Scale And Depth In Directed Evolution. *Current Opinion In Chemical Biology*, *64*, 20–26. Https://Doi.Org/10.1016/J.Cbpa.2021.02.008

Rix, G., Watkins-Dulaney, E. J., Almhjell, P. J., Boville, C. E., Arnold, F. H., & Liu, C. C. (2020). Scalable Continuous Evolution For The Generation Of Diverse Enzyme Variants Encompassing Promiscuous Activities. *Nature Communications*, *11*(1), 5644. Https://Doi.Org/10.1038/S41467-020-19539-6

Rockah-Shmuel, L., Tóth-Petróczy, Á., & Tawfik, D. S. (2015). Systematic Mapping Of Protein Mutational Space By Prolonged Drift Reveals The Deleterious Effects Of Seemingly Neutral Mutations. *Plos Computational Biology*, *11*(8), 1–28. Https://Doi.Org/10.1371/Journal.Pcbi.1004421

Rodríguez-Beltrán, J., Sørum, V., Toll-Riera, M., De La Vega, C., Peña-Miller, R., & San Millán, Á. (2020). Genetic Dominance Governs The Evolution And Spread Of Mobile Genetic Elements In Bacteria. *Proceedings Of The National Academy Of Sciences*, *117*(27), 15755 LP – 15762. Https://Doi.Org/10.1073/Pnas.2001240117

Rogers, H. ., Perkins, H. ., & Ward, J. B. (1980). *Microbial Cell Walls And Membranes*. London: Chapman And Hall.

Rojas, E. R., & Huang, K. C. (2018). Regulation Of Microbial Growth By Turgor Pressure. *Current Opinion In Microbiology*, *42*, 62–70. Https://Doi.Org/10.1016/J.Mib.2017.10.015

Romero, P. A., & Arnold, F. H. (2009a). Exploring Protein Fitness Landscapes By Directed Evolution. *Nature Reviews. Molecular Cell Biology*, *10*(12), 866–876. Https://Doi.Org/10.1038/Nrm2805

Romero, P. A., & Arnold, F. H. (2009b). Exploring Protein Fitness Landscapes By Directed Evolution. *Nature Reviews Molecular Cell Biology*, *10*(12), 866–876. Https://Doi.Org/10.1038/Nrm2805

Ruiz, C., & Levy, S. B. (2010). Many Chromosomal Genes Modulate Mara-Mediated Multidrug Resistance In Escherichia Coli. *Antimicrobial Agents And Chemotherapy*, *54*(5), 2125–2134. Https://Doi.Org/10.1128/AAC.01420-09

Ryan, O. W., Skerker, J. M., Maurer, M. J., Li, X., Tsai, J. C., Poddar, S., … Cate, J. H. D. (2014). Selection Of Chromosomal DNA Libraries Using A Multiplex CRISPR System. *Elife*, *3*, E03703.

Https://Doi.Org/10.7554/Elife.03703

Salverda, M. L. M., De Visser, J. A. G. M., & Barlow, M. (2010). Natural Evolution Of TEM-1 B-Lactamase: Experimental Reconstruction And Clinical Relevance. *FEMS Microbiology Reviews*, *34*(6), 1015–1036. Https://Doi.Org/10.1111/J.1574-6976.2010.00222.X

Sandoz, K. M., Moore, R. A., Beare, P. A., Patel, A. V, Smith, R. E., Bern, M., … Heinzen, R. A. (2021). B-Barrel Proteins Tether The Outer Membrane In Many Gram-Negative Bacteria. *Nature Microbiology*, *6*(1), 19–26. Https://Doi.Org/10.1038/S41564-020-00798-4

Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A., & Charlier, P. (2008a). The Penicillin-Binding Proteins: Structure And Role In Peptidoglycan Biosynthesis. *FEMS Microbiology Reviews*, *32*(2), 234–258. Https://Doi.Org/10.1111/J.1574-6976.2008.00105.X

Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A., & Charlier, P. (2008b). The Penicillin-Binding Proteins: Structure And Role In Peptidoglycan Biosynthesis. *FEMS Microbiology Reviews*, *32*(2), 234–258. Https://Doi.Org/10.1111/J.1574-6976.2008.00105.X

Schenkmayerova, A., Pinto, G. P., Toul, M., Marek, M., Hernychova, L., Planas-Iglesias, J., … Damborsky, J. (2021). Engineering The Protein Dynamics Of An Ancestral Luciferase. *Nature Communications*, *12*(1), 3616. Https://Doi.Org/10.1038/S41467-021-23450-Z

Schierle, C. F., Berkmen, M., Huber, D., Kumamoto, C., Boyd, D., & Beckwith, J. (2003). The Dsba Signal Sequence Directs Efficient, Cotranslational Export Of Passenger Proteins To The Escherichia Coli Periplasm Via The Signal Recognition Particle Pathway. *Journal Of Bacteriology*, *185*(19), 5706–5713. Https://Doi.Org/10.1128/JB.185.19.5706-5713.2003

Schlapschy, M., & Skerra, A. (2011). Periplasmic Chaperones Used To Enhance Functional Secretion Of Proteins In *E. coli*. *Methods In Molecular Biology (Clifton, N.J.)*, *705*, 211–224. Https://Doi.Org/10.1007/978-1-61737-967-3_12

Schlegel, S., Rujas, E., Ytterberg, A. J., Zubarev, R. A., Luirink, J., & De Gier, J.-W. (2013). Optimizing Heterologous Protein Production In The Periplasm Of *E. coli* By Regulating Gene Expression Levels. *Microbial Cell Factories*, *12*, 24. Https://Doi.Org/10.1186/1475-2859-12-24

Schleifer, K. H., & Kandler, O. (1972). Peptidoglycan Types Of Bacterial Cell

Walls And Their Taxonomic Implications. *Bacteriological Reviews*, *36*(4), 407–477. Https://Doi.Org/10.1128/Br.36.4.407-477.1972

Schleyer, M., Schmid, R., & Bakker, E. P. (1993). Transient, Specific And Extremely Rapid Release Of Osmolytes From Growing Cells Of Escherichia Coli K-12 Exposed To Hypoosmotic Shock. *Archives Of Microbiology*, *160*(6), 424–431. Https://Doi.Org/10.1007/BF00245302

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image To Imagej: 25 Years Of Image Analysis. *Nature Methods*, *9*(7), 671–675. Https://Doi.Org/10.1038/Nmeth.2089

Schultz, S. C., Dalbadie-M~Farland, G., Neit~El, J. J., & Richards', J. H. (1987). Stability Of Wild-Type And Mutant RTEM-1 ,8-Lactamases: Effect Of The Disulfide Bond. In *PROTEINS: Structure, Function, And Genetics*.

Seoane, A. S., & Levy, S. B. (1995). Characterization Of Marr, The Repressor Of The Multiple Antibiotic Resistance (Mar) Operon In Escherichia Coli. *Journal Of Bacteriology*, *177*(12), 3414–3419. Https://Doi.Org/10.1128/Jb.177.12.3414-3419.1995

Serohijos, A. W. R., Rimas, Z., & Shakhnovich, E. I. (2012a). Protein Biophysics Explains Why Highly Abundant Proteins Evolve Slowly. *Cell Reports*, *2*(2), 249–256.
Https://Doi.Org/Https://Doi.Org/10.1016/J.Celrep.2012.06.022

Serohijos, A. W. R., Rimas, Z., & Shakhnovich, E. I. (2012b). Protein Biophysics Explains Why Highly Abundant Proteins Evolve Slowly. *Cell Reports*, *2*(2), 249–256. Https://Doi.Org/10.1016/J.Celrep.2012.06.022

Sham, L.-T., Butler, E. K., Lebar, M. D., Kahne, D., Bernhardt, T. G., & Ruiz, N. (2014). Bacterial Cell Wall. Murj Is The Flippase Of Lipid-Linked Precursors For Peptidoglycan Biogenesis. *Science (New York, N.Y.)*, *345*(6193), 220–222. Https://Doi.Org/10.1126/Science.1254522

Shimizu-Ibuka, A., Matsuzawa, H., & Sakai, H. (2004). An Engineered Disulfide Bond Between Residues 69 And 238 In Extended-Spectrum B-Lactamase Toho-1 Reduces Its Activity Toward Third-Generation Cephalosporins. *Biochemistry*, *43*(50), 15737–15745. Https://Doi.Org/10.1021/Bi048488u

Shimizu-Ibuka, A., Matsuzawa, H., & Sakai, H. (2006). Effect Of Disulfide-Bond Introduction On The Activity And Stability Of The Extended-Spectrum Class A B-Lactamase Toho-1. *Biochimica Et Biophysica Acta - Proteins And Proteomics*, *1764*(8), 1349–1355.

Https://Doi.Org/10.1016/J.Bbapap.2006.06.004

Shinkai, A., & Loeb, L. A. (2001). In Vivo Mutagenesis By Escherichia Coli DNA Polymerase I: Ile 709 In Motif A Functions In Base Selection. *Journal Of Biological Chemistry*, *276*(50), 46759–46764. Https://Doi.Org/10.1074/Jbc.M104780200

Shortle, D., & Sondek, J. (1995). The Emerging Role Of Insertions And Deletions In Protein Engineering. *Current Opinion In Biotechnology*, *6*(4), 387–393. Https://Doi.Org/10.1016/0958-1669(95)80067-0

Silvers, R., Sziegat, F., Tachibana, H., Segawa, S., Whittaker, S., Günther, U. L., … Schwalbe, H. (2012). Modulation Of Structure And Dynamics By Disulfide Bond Formation In Unfolded States. *Journal Of The American Chemical Society*, *134*(15), 6846–6854. Https://Doi.Org/10.1021/Ja3009506

Simakov, N., Leonard, D. A., Smith, J. C., Wymore, T., & Szarecka, A. (2017). A Distal Disulfide Bridge In OXA-1 B-Lactamase Stabilizes The Catalytic Center And Alters The Dynamics Of The Specificity Determining Ω Loop. *Journal Of Physical Chemistry B*, *121*(15), 3285–3296. Https://Doi.Org/10.1021/Acs.Jpcb.6b07884

Sjodt, M., Rohs, P. D. A., Gilman, M. S. A., Erlandson, S. C., Zheng, S., Green, A. G., … Kruse, A. C. (2020). Structural Coordination Of Polymerization And Crosslinking By A SEDS–Bpbp Peptidoglycan Synthase Complex. *Nature Microbiology*, *5*(6), 813–820. Https://Doi.Org/10.1038/S41564-020-0687-Z

Skamaki, K., Emond, S., Chodorge, M., Andrews, J., Rees, D. G., Cannon, D., … Hollfelder, F. (2020). In Vitro Evolution Of Antibody Affinity Via Insertional Scanning Mutagenesis Of An Entire Antibody Variable Region. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *117*(44), 27307–27318. Https://Doi.Org/10.1073/Pnas.2002954117

Sleator, R. D., & Hill, C. (2002). Bacterial Osmoadaptation: The Role Of Osmolytes In Bacterial Stress And Virulence. *FEMS Microbiology Reviews*, *26*(1), 49–71. Https://Doi.Org/10.1111/J.1574-6976.2002.Tb00598.X

Slonczewski, J. L., Rosen, B. P., Alger, J. R., & Macnab, R. M. (1981). Ph Homeostasis In Escherichia Coli: Measurement By 31P Nuclear Magnetic Resonance Of Methylphosphonate And Phosphate. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *78*(10), 6271–6275. Https://Doi.Org/10.1073/Pnas.78.10.6271

Smith, C. A., Nossoni, Z., Toth, M., Stewart, N. K., Frase, H., & Vakulenko, S. B. (2016). Role Of The Conserved Disulfide Bridge In Class A Carbapenemases. *Journal Of Biological Chemistry*, *291*(42), 22196–22206. Https://Doi.Org/10.1074/Jbc.M116.749648

Smith, G. P., & Petrenko, V. A. (1997). Phage Display. *Chemical Reviews*, *97*(2), 391–410. Https://Doi.Org/10.1021/Cr960065d

Smith, J. D., Kumarasiri, M., Zhang, W., Hesek, D., Lee, M., Toth, M., … Chen, Y. (2013). Structural Analysis Of The Role Of Pseudomonas Aeruginosa Penicillin-Binding Protein 5 In B-Lactam Resistance. *Antimicrobial Agents And Chemotherapy*, *57*(7), 3137–3146. Https://Doi.Org/10.1128/AAC.00505-13

Soumillion, P., & Fastrez, J. (2002, April 22). Investigation Of Phage Display For The Directed Evolution Of Enzymes. *Directed Molecular Evolution Of Proteins*, Pp. 79–110. Https://Doi.Org/Https://Doi.Org/10.1002/3527600647.Ch5

Sousa, A., Magalhães, S., & Gordo, I. (2012). Cost Of Antibiotic Resistance And The Geometry Of Adaptation. *Molecular Biology And Evolution*, *29*(5), 1417–1428. Https://Doi.Org/10.1093/Molbev/Msr302

Speight, R. E., Hart, D. J., Sutherland, J. D., & Blackburn, J. M. (2001). A New Plasmid Display Technology For The In Vitro Selection Of Functional Phenotype-Genotype Linked Proteins. *Chemistry & Biology*, *8*(10), 951–965. Https://Doi.Org/10.1016/S1074-5521(01)00066-7

Sperandeo, P., Martorana, A. M., & Polissi, A. (2019). *Lipopolysaccharide Biosynthesis And Transport To The Outer Membrane Of Gram-Negative Bacteria BT - Bacterial Cell Walls And Membranes* (A. Kuhn, Ed.). Https://Doi.Org/10.1007/978-3-030-18768-2_2

Stancik, L. M., Stancik, D. M., Schmidt, B., Barnhart, D. M., Yoncheva, Y. N., & Slonczewski, J. L. (2002). Ph-Dependent Expression Of Periplasmic Proteins And Amino Acid Catabolism In Escherichia Coli. *Journal Of Bacteriology*, *184*(15), 4246–4258. Https://Doi.Org/10.1128/JB.184.15.4246-4258.2002

Starr, T. N., Picton, L. K., & Thornton, J. W. (2017). Alternative Evolutionary Histories In The Sequence Space Of An Ancient Protein. *Nature*, *549*(7672), 409–413. Https://Doi.Org/10.1038/Nature23902

Stein, V., Sielaff, I., Johnsson, K., & Hollfelder, F. (2007). A Covalent Chemical

Genotype–Phenotype Linkage For In Vitro Protein Evolution. *Chembiochem*, *8*(18), 2191–2194. Https://Doi.Org/Https://Doi.Org/10.1002/Cbic.200700459

Stemmer, W. P. (1994). Rapid Evolution Of A Protein In Vitro By DNA Shuffling. *Nature*, *370*(6488), 389–391. Https://Doi.Org/10.1038/370389a0

Stern, A., Bianco, S., Yeh, M. Te, Wright, C., Butcher, K., Tang, C., … Andino, R. (2014). Costs And Benefits Of Mutational Robustness In RNA Viruses. *Cell Reports*, *8*(4), 1026–1036. Https://Doi.Org/10.1016/J.Celrep.2014.07.011

Stokes, J. M., French, S., Ovchinnikova, O. G., Bouwman, C., Whitfield, C., & Brown, E. D. (2016). Cold Stress Makes Escherichia Coli Susceptible To Glycopeptide Antibiotics By Altering Outer Membrane Integrity. *Cell Chemical Biology*, *23*(2), 267–277. Https://Doi.Org/10.1016/J.Chembiol.2015.12.011

Straume, D., Piechowiak, K. W., Olsen, S., Stamsås, G. A., Berg, K. H., Kjos, M., … Håvarstein, L. S. (2020). Class A Pbps Have A Distinct And Unique Role In The Construction Of The Pneumococcal Cell Wall. *Proceedings Of The National Academy Of Sciences*, *117*(11), 6129 LP – 6138. Https://Doi.Org/10.1073/Pnas.1917820117

Sun, X., Skorstengaard, K., & Mosher, D. F. (1992). Disulfides Modulate RGD-Inhibitable Cell Adhesive Activity Of Thrombospondin. *The Journal Of Cell Biology*, *118*(3), 693–701. Https://Doi.Org/10.1083/Jcb.118.3.693

Szarecka, A., Lesnock, K. R., Ramirez-Mondragon, C. A., Nicholas Jr, H. B., & Wymore, T. (2011). The Class D Beta-Lactamase Family: Residues Governing The Maintenance And Diversity Of Function. *Protein Engineering, Design & Selection : PEDS*, *24*(10), 801–809. Https://Doi.Org/10.1093/Protein/Gzr041

Taguchi, A., Welsh, M. A., Marmont, L. S., Lee, W., Sjodt, M., Kruse, A. C., … Walker, S. (2019). Ftsw Is A Peptidoglycan Polymerase That Is Functional Only In Complex With Its Cognate Penicillin-Binding Protein. *Nature Microbiology*, *4*(4), 587–594. Https://Doi.Org/10.1038/S41564-018-0345-X

Takahashi, T. T., & Roberts, R. W. (2009). In Vitro Selection Of Protein And Peptide Libraries Using Mrna Display. *Methods In Molecular Biology (Clifton,*

N.J.), *535*, 293–314. Https://Doi.Org/10.1007/978-1-59745-557-2_17

Tapia-Rojo, R., Alonso-Caballero, A., Badilla, C. L., & Fernandez, J. M. (2021). Identical Sequences, Different Behaviors: Protein Diversity Captured At The Single-Molecule Level. *Biorxiv*, 2021.02.24.432730. Https://Doi.Org/10.1101/2021.02.24.432730

Tawfik, D. S., & Griffiths, A. D. (1998). Man-Made Cell-Like Compartments For Molecular Evolution . *Nature Biotechnology*, *16*(7), 652–656. Https://Doi.Org/10.1038/Nbt0798-652

Thompson, M. G., Sedaghatian, N., Barajas, J. F., Wehrs, M., Bailey, C. B., Kaplan, N., … Keasling, J. D. (2018). Isolation And Characterization Of Novel Mutations In The Psc101 Origin That Increase Copy Number. *Scientific Reports*, *8*(1), 1590. Https://Doi.Org/10.1038/S41598-018-20016-W

Tinafar, A., Jaenes, K., & Pardee, K. (2019). Synthetic Biology Goes Cell-Free. *BMC Biology*, *17*(1), 64. Https://Doi.Org/10.1186/S12915-019-0685-X

Tizei, P. A. G., Harris, E., Withanage, S., Renders, M., & Pinheiro, V. B. (2021). A Novel Framework For Engineering Protein Loops Exploring Length And Compositional Variation. *Scientific Reports*, *11*(1), 9134. Https://Doi.Org/10.1038/S41598-021-88708-4

Tokuriki, N., & Tawfik, D. S. (2009a). Protein Dynamism And Evolvability. *Science (New York, N.Y.)*, *324*(5924), 203–207. Https://Doi.Org/10.1126/Science.1169375

Tokuriki, N., & Tawfik, D. S. (2009b, October). Stability Effects Of Mutations And Protein Evolvability. *Current Opinion In Structural Biology*, Vol. 19, Pp. 596–604. Https://Doi.Org/10.1016/J.Sbi.2009.08.003

Tooke, C. L., Hinchliffe, P., Bonomo, R. A., Schofield, C. J., Mulholland, A. J., & Spencer, J. (2021). Natural Variants Modify Klebsiella Pneumoniae Carbapenemase (KPC) Acyl–Enzyme Conformational Dynamics To Extend Antibiotic Resistance. *Journal Of Biological Chemistry*, *296*, 100126. Https://Doi.Org/Https://Doi.Org/10.1074/Jbc.RA120.016461

Tooke, C. L., Hinchliffe, P., Bragginton, E. C., Colenso, C. K., Hirvonen, V. H. A., Takebayashi, Y., & Spencer, J. (2019). B-Lactamases And B-Lactamase Inhibitors In The 21st Century. *Journal Of Molecular Biology*, *431*(18), 3472–3500. Https://Doi.Org/10.1016/J.Jmb.2019.04.002

Trabelsi, H. (2014). *Directed Evolution Of A DD-Peptidase Into A B-Lactamase* (Université Catholique De Louvain). Retrieved From Http://Hdl.Handle.Net/2078.1/151945

Tracz, D. M., Boyd, D. A., Bryden, L., Hizon, R., Giercke, S., Van Caeseele, P., & Mulvey, M. R. (2005). Increase In Ampc Promoter Strength Due To Mutations And Deletion Of The Attenuator In A Clinical Isolate Of Cefoxitin-Resistant Escherichia Coli As Determined By RT-PCR. *The Journal Of Antimicrobial Chemotherapy*, *55*(5), 768–772. Https://Doi.Org/10.1093/Jac/Dki074

Tracz, D. M., Boyd, D. A., Hizon, R., Bryce, E., Mcgeer, A., Ofner-Agostini, M., … Mulvey, M. R. (2007). Ampc Gene Expression In Promoter Mutants Of Cefoxitin-Resistant Escherichia Coli Clinical Isolates. *FEMS Microbiology Letters*, *270*(2), 265–271. Https://Doi.Org/10.1111/J.1574-6968.2007.00672.X

Trösemeier, J.-H., Rudorf, S., Loessner, H., Hofner, B., Reuter, A., Schulenborg, T., … Kamp, C. (2019). Optimizing The Dynamics Of Protein Expression. *Scientific Reports*, *9*(1), 7511. Https://Doi.Org/10.1038/S41598-019-43857-5

Typas, A., Banzhaf, M., Gross, C. A., & Vollmer, W. (2011). From The Regulation Of Peptidoglycan Synthesis To Bacterial Growth And Morphology. *Nature Reviews. Microbiology*, *10*(2), 123–136. Https://Doi.Org/10.1038/Nrmicro2677

Typas, A., Banzhaf, M., Van Den Berg Van Saparoea, B., Verheul, J., Biboy, J., Nichols, R. J., … Vollmer, W. (2010). Regulation Of Peptidoglycan Synthesis By Outer-Membrane Proteins. *Cell*, *143*(7), 1097–1109. Https://Doi.Org/10.1016/J.Cell.2010.11.038

Urbach, C., Evrard, C., Pudzaitis, V., Fastrez, J., Soumillion, P., & Declercq, J.-P. (2009). Structure Of PBP-A From Thermosynechococcus Elongatus, A Penicillin-Binding Protein Closely Related To Class A B-Lactamases. In *Journal Of Molecular Biology* (Vol. 386).

Urbach, C., Fastrez, J., & Soumillion, P. (2008). A New Family Of Cyanobacterial Penicillin-Binding Proteins. A Missing Link In The Evolution Of Class A Beta-Lactamases. *The Journal Of Biological Chemistry*, *283*(47), 32516–32526. Https://Doi.Org/10.1074/Jbc.M805375200

Urban, J. H., & Merten, C. A. (2011). Retroviral Display In Gene Therapy,

Protein Engineering, And Vaccine Development. *ACS Chemical Biology*, *6*(1), 61–74. Https://Doi.Org/10.1021/Cb100285n

Urvoas, A., Minard, P., & Soumillion, P. (2020). *Phage Display Methodologies BT - Peptide And Protein Engineering: From Concepts To Biotechnological Applications* (O. Iranzo & A. C. Roque, Eds.). Https://Doi.Org/10.1007/978-1-0716-0720-6_8

Vaara, M. (1993). Outer Membrane Permeability Barrier To Azithromycin, Clarithromycin, And Roxithromycin In Gram-Negative Enteric Bacteria. *Antimicrobial Agents And Chemotherapy*, *37*(2), 354–356. Https://Doi.Org/10.1128/AAC.37.2.354

Van Heijenoort, J. (2011). Peptidoglycan Hydrolases Of Escherichia Coli. *Microbiology And Molecular Biology Reviews : MMBR*, *75*(4), 636–663. Https://Doi.Org/10.1128/MMBR.00022-11

Van Laer, K., Hamilton, C. J., & Messens, J. (2013). Low-Molecular-Weight Thiols In Thiol-Disulfide Exchange. *Antioxidants & Redox Signaling*, *18*(13), 1642–1653. Https://Doi.Org/10.1089/Ars.2012.4964

Van Rossum, T., Kengen, S. W. M., & Van Der Oost, J. (2013). Reporter-Based Screening And Selection Of Enzymes. *The FEBS Journal*, *280*(13), 2979–2996. Https://Doi.Org/10.1111/Febs.12281

Veiga-Da-Cunha, M., Van Schaftingen, E., & Bommer, G. T. (2020). Inborn Errors Of Metabolite Repair. *Journal Of Inherited Metabolic Disease*, *43*(1), 14–24. Https://Doi.Org/10.1002/Jimd.12187

Verheul, J., Lodge, A., Yau, H. C. L., Liu, X., Solovyova, A. S., Typas, A., … Den Blaauwen, T. (2020). Midcell Localization Of PBP4 Of Escherichia Coli Is Essential For The Timing Of Divisome Assembly. *Biorxiv*, 2020.07.30.230052. Https://Doi.Org/10.1101/2020.07.30.230052

Vermassen, A., Leroy, S., Talon, R., Provot, C., Popowska, M., & Desvaux, M. (2019). Cell Wall Hydrolases In Bacteria: Insight On The Diversity Of Cell Wall Amidases, Glycosidases And Peptidases Toward Peptidoglycan . *Frontiers In Microbiology* , Vol. 10, P. 331. Retrieved From Https://Www.Frontiersin.Org/Article/10.3389/Fmicb.2019.00331

Verwer, R. W., Nanninga, N., Keck, W., & Schwarz, U. (1978). Arrangement Of Glycan Chains In The Sacculus Of Escherichia Coli. *Journal Of Bacteriology*, *136*(2), 723–729. Https://Doi.Org/10.1128/JB.136.2.723-729.1978

Voedts, H., Dorchêne, D., Lodge, A., Vollmer, W., Arthur, M., & Hugonnet, J.-E. (2021). Role Of Endopeptidases In Peptidoglycan Synthesis Mediated By Alternative Cross-Linking Enzymes In Escherichia Coli; *Biorxiv*, 2021.02.12.430937. Https://Doi.Org/10.1101/2021.02.12.430937

Vollmer, W., & Bertsche, U. (2008). Murein (Peptidoglycan) Structure, Architecture And Biosynthesis In Escherichia Coli. *Biochimica Et Biophysica Acta - Biomembranes*, *1778*(9), 1714–1734. Https://Doi.Org/10.1016/J.Bbamem.2007.06.007

Vollmer, W., Blanot, D., & De Pedro, M. A. (2008). Peptidoglycan Structure And Architecture. *FEMS Microbiology Reviews*, *32*(2), 149–167. Https://Doi.Org/10.1111/J.1574-6976.2007.00094.X

Vollmer, W., & Höltje, J.-V. (2004). The Architecture Of The Murein (Peptidoglycan) In Gram-Negative Bacteria: Vertical Scaffold Or Horizontal Layer(S)? *Journal Of Bacteriology*, *186*(18), 5978–5987. Https://Doi.Org/10.1128/JB.186.18.5978-5987.2004

Vollmer, W., Joris, B., Charlier, P., & Foster, S. (2008). Bacterial Peptidoglycan (Murein) Hydrolases. *FEMS Microbiology Reviews*, *32*(2), 259–286. Https://Doi.Org/10.1111/J.1574-6976.2007.00099.X

Wagner, A. (2005a). Robustness, Evolvability, And Neutrality. *FEBS Letters*, *579*(8), 1772–1778. Https://Doi.Org/10.1016/J.Febslet.2005.01.063

Wagner, A. (2005b). *Robustness And Evolvability In Living Systems:* Https://Doi.Org/10.2307/J.Ctt4cgcnc

Wagner, A. (2008). Robustness And Evolvability: A Paradox Resolved. *Proceedings Of The Royal Society B: Biological Sciences*, *275*(1630), 91–100. Https://Doi.Org/10.1098/Rspb.2007.1137

Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., & Church, G. M. (2009). Programming Cells By Multiplex Genome Engineering And Accelerated Evolution. *Nature*, *460*(7257), 894–898. Https://Doi.Org/10.1038/Nature08187

Wang, S., Li, W., Liu, S., & Xu, J. (2016). Raptorx-Property: A Web Server For Protein Structure Property Prediction. *Nucleic Acids Research*, *44*(W1), W430-5. Https://Doi.Org/10.1093/Nar/Gkw306

Wetzel, R. (1987). Harnessing Disulfide Bonds Using Protein Engineering. *Trends In Biochemical Sciences*, *12*, 478–482.

Https://Doi.Org/Https://Doi.Org/10.1016/0968-0004(87)90234-9

Wickner, W., Driessen, A. J. M., & Haril, F.-U. (1991). THE ENZYMOLOGY OF PROTEIN TRANSLOCATION ACROSS THE Escherichia Coli PLASMA MEMBRANE. *Annual Review Of Biochemistry*, *60*(1), 101–124. Https://Doi.Org/10.1146/Annurev.Bi.60.070191.000533

Wilks, J. C., & Slonczewski, J. L. (2007). Ph Of The Cytoplasm And Periplasm Of Escherichia Coli: Rapid Measurement By Green Fluorescent Protein Fluorimetry. *Journal Of Bacteriology*, *189*(15), 5601–5607. Https://Doi.Org/10.1128/JB.00615-07

Wilson, A. J., Pemberton, J. M., Pilkington, J. G., Coltman, D. W., Mifsud, D. V, Clutton-Brock, T. H., & Kruuk, L. E. B. (2006). Environmental Coupling Of Selection And Heritability Limits Evolution. *PLOS Biology*, *4*(7), E216. Retrieved From Https://Doi.Org/10.1371/Journal.Pbio.0040216

Wong, J. W. H., Ho, S. Y. W., & Hogg, P. J. (2011). Disulfide Bond Acquisition Through Eukaryotic Protein Evolution. *Molecular Biology And Evolution*, *28*(1), 327–334. Https://Doi.Org/10.1093/Molbev/Msq194

Worthington, R. J., & Melander, C. (2013). Overcoming Resistance To B-Lactam Antibiotics. *The Journal Of Organic Chemistry*, *78*(9), 4207–4213. Https://Doi.Org/10.1021/Jo400236f

Wyke, A. W., Ward, J. B., Hayes, M. V, & Curtis, N. A. (1981). A Role In Vivo For Penicillin-Binding Protein-4 Of Staphylococcus Aureus. *European Journal Of Biochemistry*, *119*(2), 389–393. Https://Doi.Org/10.1111/J.1432-1033.1981.Tb05620.X

Xiao, H., Bao, Z., & Zhao, H. (2015). High Throughput Screening And Selection Methods For Directed Enzyme Evolution. *Industrial & Engineering Chemistry Research*, *54*(16), 4011–4020. Https://Doi.Org/10.1021/Ie503060a

Yang, H., Hu, J., Lu, X., Wang, F., Shen, W., Hu, W., … Liu, L. (2019). Improving Extracellular Protein Production In Escherichia Coli By Overexpressing D,D-Carboxypeptidase To Perturb Peptidoglycan Network Synthesis And Structure. *Applied Microbiology And Biotechnology*, *103*(2), 793–806. Https://Doi.Org/10.1007/S00253-018-9510-7

Yasukawa, T., Kanei-Ishii, C., Maekawa, T., Fujimoto, J., Yamamoto, T., & Ishii, S. (1995). Increase Of Solubility Of Foreign Proteins In Escherichia Coli By Coproduction Of The Bacterial Thioredoxin. *The Journal Of Biological Chemistry*, *270*(43), 25328–25331.

Https://Doi.Org/10.1074/Jbc.270.43.25328

Yokobayashi, Y., Weiss, R., & Arnold, F. H. (2002). Directed Evolution Of A Genetic Circuit. *Proceedings Of The National Academy Of Sciences*, *99*(26), 16587 LP – 16591. Https://Doi.Org/10.1073/Pnas.252535999

York, A., Lloyd, A. J., Del Genio, C. I., Shearer, J., Hinxman, K. J., Fritz, K., … Roper, D. I. (2020). Structure-Based Modelling And Dynamics Of MurM, A; Streptococcus Pneumoniae; Penicillin Resistance Determinant That Functions At The Cytoplasmic Membrane Interface. *Biorxiv*, 2020.06.18.158840. Https://Doi.Org/10.1101/2020.06.18.158840

Zapun, A., Philippe, J., Abrahams, K. A., Signor, L., Roper, D. I., Breukink, E., & Vernet, T. (2013). In Vitro Reconstitution Of Peptidoglycan Assembly From The Gram-Positive Pathogen Streptococcus Pneumoniae. *ACS Chemical Biology*, *8*(12), 2688–2696. Https://Doi.Org/10.1021/Cb400575t

Zeymer, C., & Hilvert, D. (2018). Directed Evolution Of Protein Catalysts. *Annual Review Of Biochemistry*, *87*(1), 131–157. Https://Doi.Org/10.1146/Annurev-Biochem-062917-012034

Zhang, Haoming, Kenaan, C., Hamdane, D., Hoa, G. H. B., & Hollenberg, P. F. (2009). Effect Of Conformational Dynamics On Substrate Recognition And Specificity As Probed By The Introduction Of A De Novo Disulfide Bond Into Cytochrome P450 2B1. *The Journal Of Biological Chemistry*, *284*(38), 25678–25686. Https://Doi.Org/10.1074/Jbc.M109.032748

Zhang, Hong, Cheng, Q. X., Liu, A. M., Zhao, G. P., & Wang, J. (2017). A Novel And Efficient Method For Bacteria Genome Editing Employing Both CRISPR/Cas9 And An Antibiotic Resistance Cassette. *Frontiers In Microbiology*, *8*(MAY), 1–11. Https://Doi.Org/10.3389/Fmicb.2017.00812

Zhao, G., Meier, T. I., Kahl, S. D., Gee, K. R., & Blaszczak, L. C. (1999). BOCILLIN FL, A Sensitive And Commercially Available Reagent For Detection Of Penicillin-Binding Proteins. *Antimicrobial Agents And Chemotherapy*, *43*(5), 1124–1128. Https://Doi.Org/10.1128/AAC.43.5.1124

Zheng, J., Guo, N., & Wagner, A. (2020). Selection Enhances Protein Evolvability By Increasing Mutational Robustness And Foldability. *Science*, *370*(6521), Eabb5962. Https://Doi.Org/10.1126/Science.Abb5962

Zhong, Z., Wong, B. G., Ravikumar, A., Arzumanyan, G. A., Khalil, A. S., & Liu, C. C. (2020). Automated Continuous Evolution Of Proteins In Vivo. *ACS*

Synthetic Biology, *9*(6), 1270–1276. Https://Doi.Org/10.1021/Acssynbio.0c00135

9. Appendices

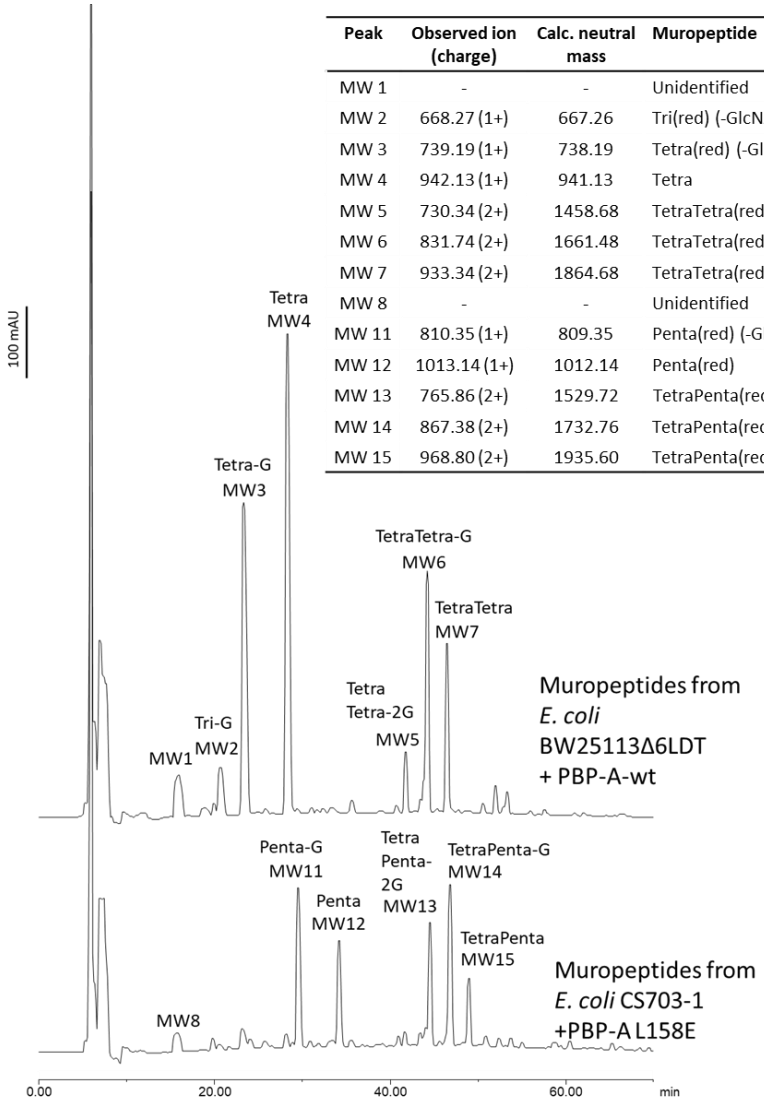
Appendix 9-1: Abundance of molecules detected in MS spectra of PBP-A-wt and PBP-A-L158E assays with purified PG from from *E. coli* Δ6LDT.

	Δ6LDT	**Δ6LDT**	**Δ6LDT**
	control	**PBPA**	**PBPA L158E**
	sample1	**sample2**	**sample3**
Muropeptide	% Area	% Area	% Area
?	0.0	8.7	9.2
Tri-G	0.0	6.5	4.5
Tetra-G	0.0	31.2	32.0
Tetra	56.6	32.7	25.0
TetraTetra-2G	0.0	3.0	5.6
TetraTetra-G	0.0	12.2	17.2
TetraTetra	39.7	5.8	6.5
all known	96.3	100.0	100.0
monomers (total)	58.8	79.1	70.8
monomer tri	0.0	6.5	4.5
monomer tetra	58.8	63.9	57.0
dimers (total)	41.2	20.9	29.2
dimers(DD)	41.2	20.9	29.2
tripeptides (total)	0.0	6.5	4.5
tetrapeptides (total)	100.0	84.8	86.3
degree of cross-linkage	20.6	10.5	14.6
% peptides in cross-links	41.2	20.9	29.2

Appendix 9-2: Abundance of molecules detected in MS spectra of PBP-A-wt and PBP-A-L158E assays with purified PG from from *E. coli* CS703-1.

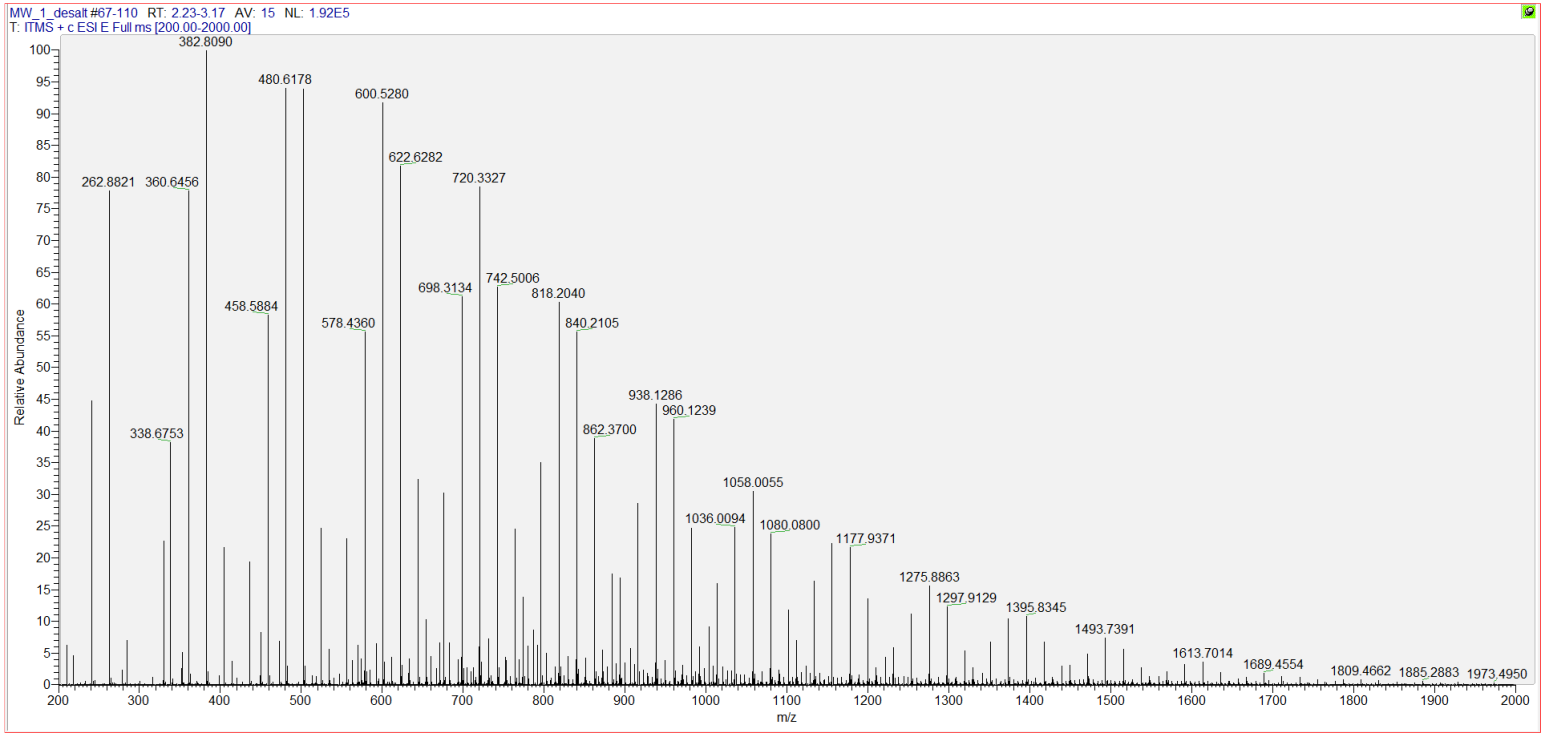
		CS703-1	CS703-1	CS703-1
		control	PBPA	PBPA L158E
		sample5	sample6	sample7
Peak No.	**Muropeptide**	% Area	% Area	% Area
1	Tri	0.5	0.8	0.58
2	TetraGly 4	0.4	0.5	0.42
3	Tetra	5.5	15.0	8.3
4	PentaGly5	1.82	1.76	1.79
5	Di	0.0	0.0	0
6	Penta	29.3	29.2	28.61
7	Tri-Lys-Arg	0.8	1.3	1.28
8	TetraPentaGly5	3.4	2.3	2.3
9	TetraTetra	4.8	3.7	3.63
10	TetraPenta	37.5	33.4	34.25
11	Unknown	0.3	0.4	2.32
12	Unknown	0.3	0.0	1.54
13	TetraTri-LysArg	0.3	0.3	0
14	TetraTetraTetra	0.3	0.9	0.77
15	TetraTetraPenta	3.0	2.7	2.39
16	Unknown	0.4	1.3	1.31
17	TetraPentaAnh I	2.0	1.5	1.97
18	TetraPentaAnh II	1.8	0.9	1.07
19	TetraTetraPentaAnh	0.9	0.7	0.71
1 - 19	all known	93.1	96.4	93.2
	monomers (total)	41.1	50.3	44.0
	monomer di	0.0	0.0	0
	monomer tri	0.5	0.8	0.6
	monomer tetra	5.9	15.6	8.9
	monomer tetraGly4	0.4	0.5	0.5
	monomer penta	31.5	30.3	30.7
	monomer pentaGly5	2.0	1.8	1.9
	dimers (total)	53.3	43.6	46.4
	dimers(DD)	53.3	43.6	46.4
	dimers(LD)	0.0	0.0	0
	dimers anhydro	4.0	2.4	3.3
	Trimers (Total)	4.5	4.4	4.2
	trimer (anhydro)	0.9	0.7	0.8
	dipeptides (total)	0.0	0.0	0
	tripeptides (total)	1.5	2.3	2.0
	tetrapeptides (total)	38.6	42.9	37.5
	pentapeptides (total)	58.8	53.0	54.9
	3-3 linkage	0.0	0.0	0
	chain ends (anhydros)	2.0	1.2	1.5
	degree of cross-linkage	29.7	24.7	25.9
	% peptides in cross-links	58.9	49.7	56.0

Appendix 9-3: HPLC peak detection in assay of PBP-As with purified PG. The table on the right represent MS analysis for molecular weight (MW) of labelled peaks obtained in HPLC. MS spectra for all the peaks are available in Appendices 2-6 to 2-19.

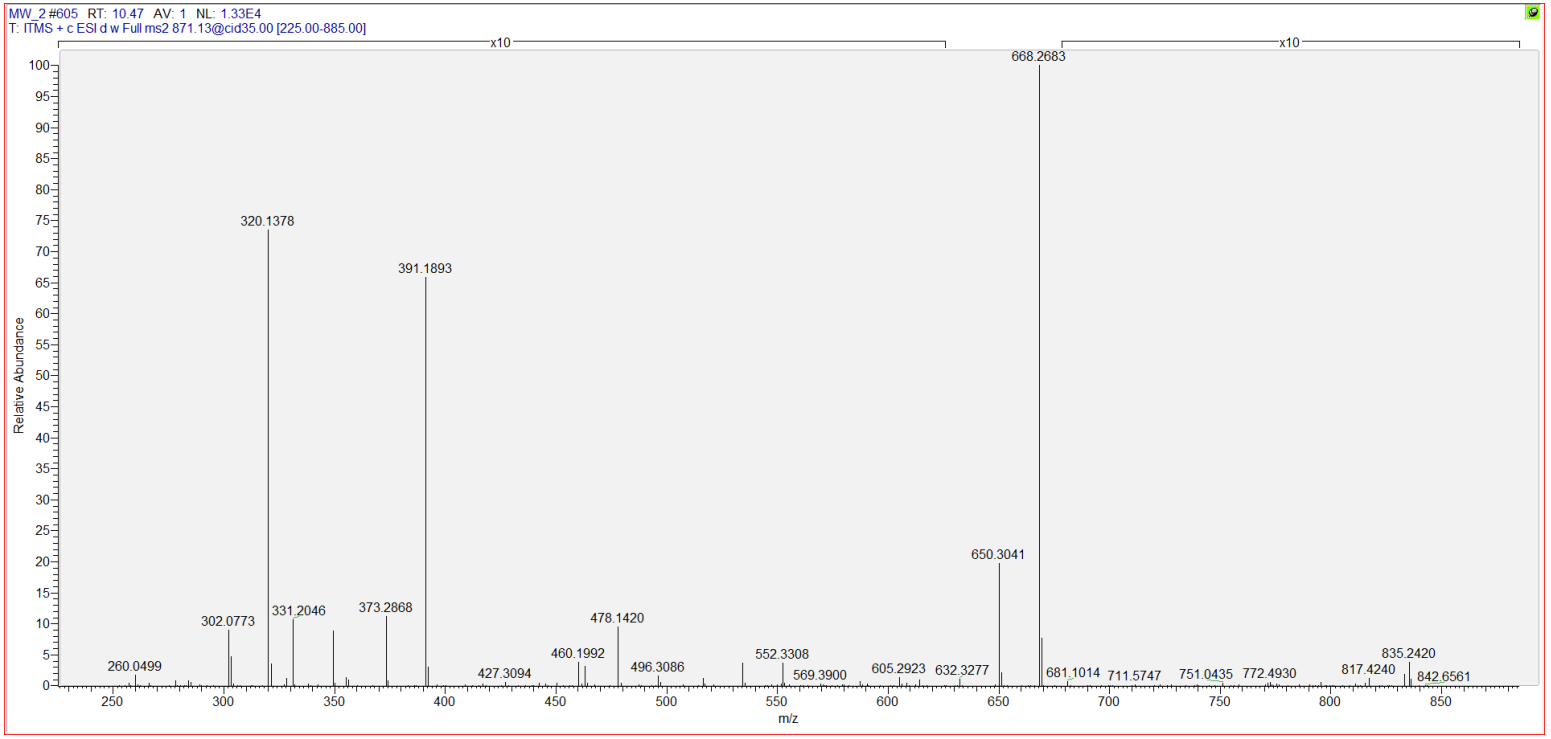


Peak	Observed ion (charge)	Calc. neutral mass	Muropeptide
MW 1	-	-	Unidentified
MW 2	668.27 (1+)	667.26	Tri(red) (-GlcNAc)
MW 3	739.19 (1+)	738.19	Tetra(red) (-GlcNAc)
MW 4	942.13 (1+)	941.13	Tetra
MW 5	730.34 (2+)	1458.68	TetraTetra(red) (-2 GlcNAc)
MW 6	831.74 (2+)	1661.48	TetraTetra(red) (-GlcNAc)
MW 7	933.34 (2+)	1864.68	TetraTetra(red)
MW 8	-	-	Unidentified
MW 11	810.35 (1+)	809.35	Penta(red) (-GlcNAc)
MW 12	1013.14 (1+)	1012.14	Penta(red)
MW 13	765.86 (2+)	1529.72	TetraPenta(red) (-2GlcNAc)
MW 14	867.38 (2+)	1732.76	TetraPenta(red) (-GlcNAc)
MW 15	968.80 (2+)	1935.60	TetraPenta(red)

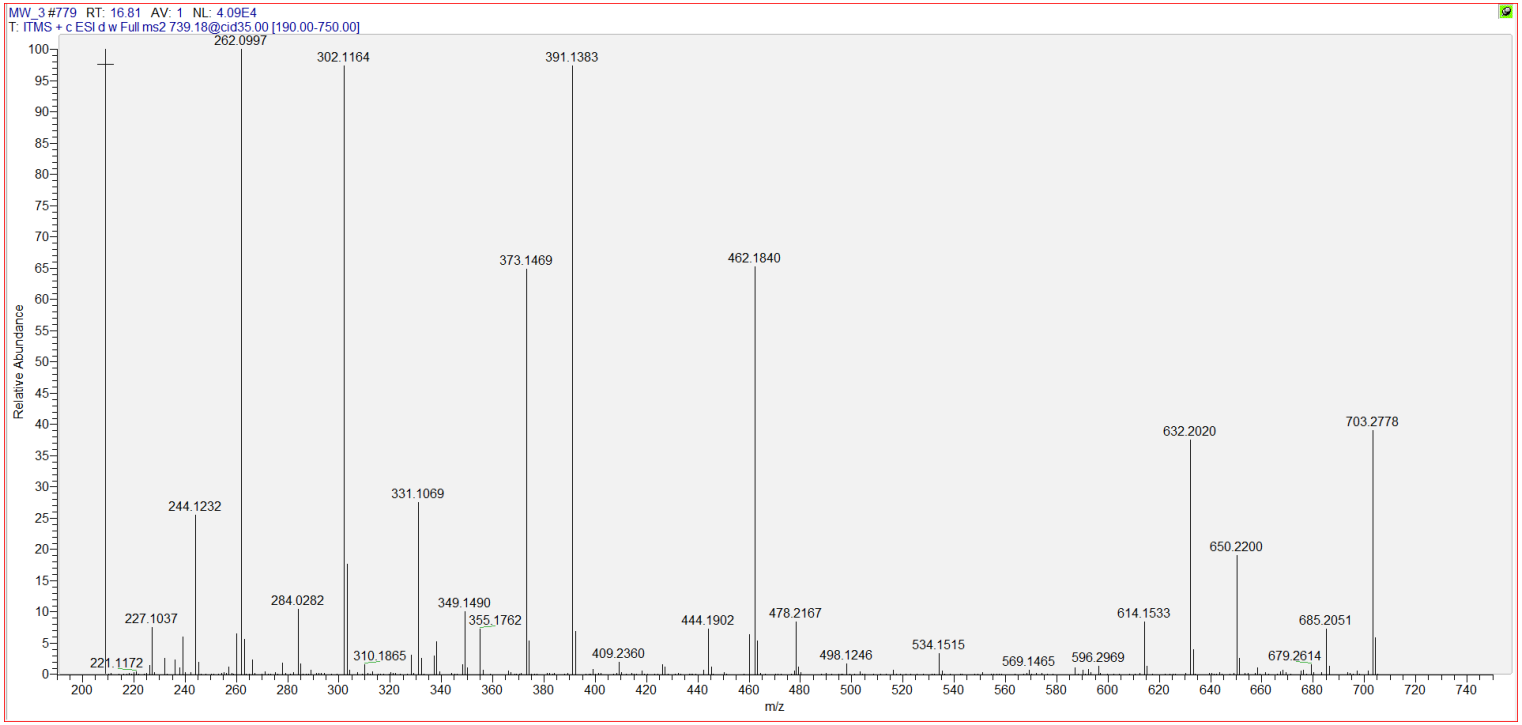
Appendix 9-4: MS spectra of peak MW1 in Appendix 9-3.



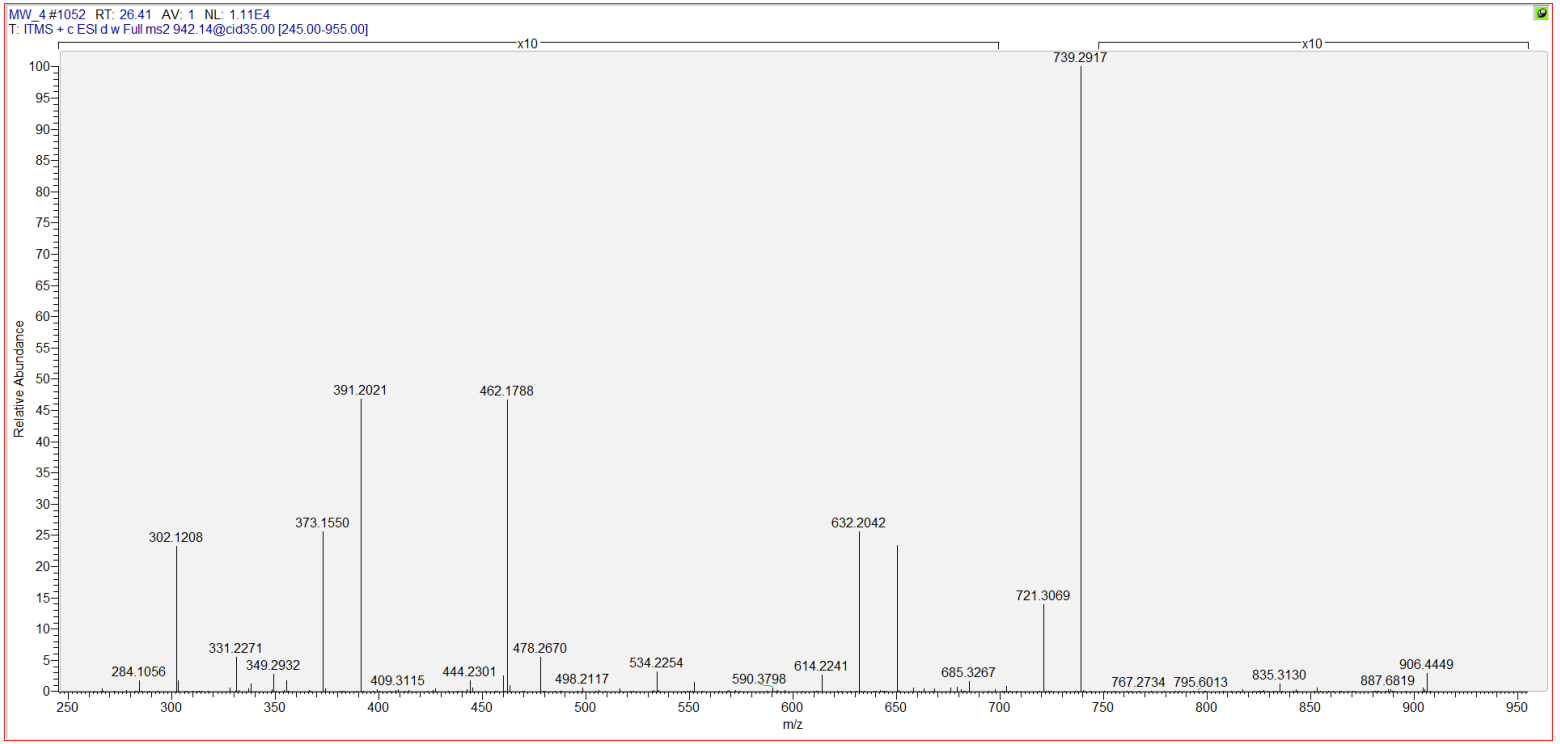
Appendix 9-5: MS spectra of peak MW2 in Appendix 9-3.



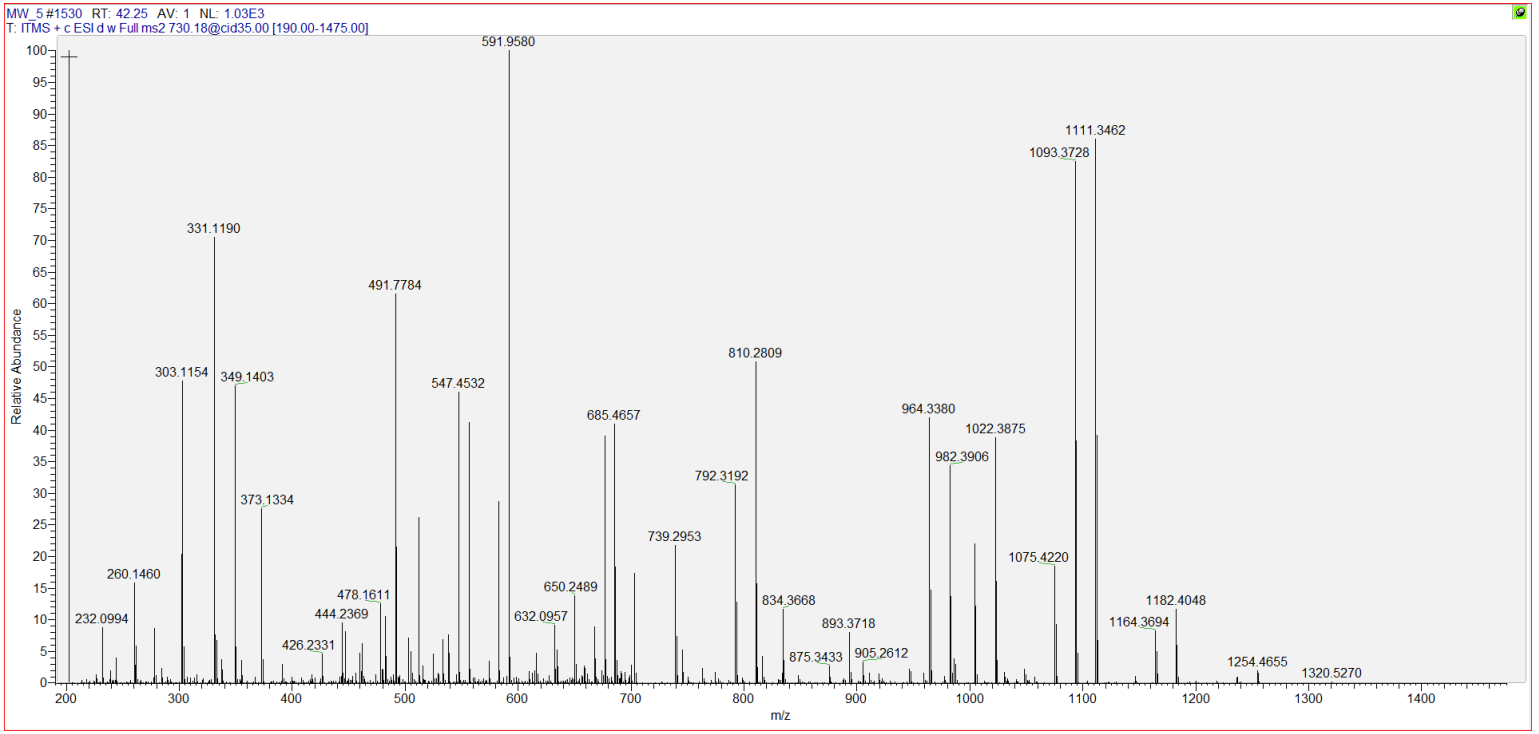
Appendix 9-6: MS spectra of peak MW3 in Appendix 9-3.



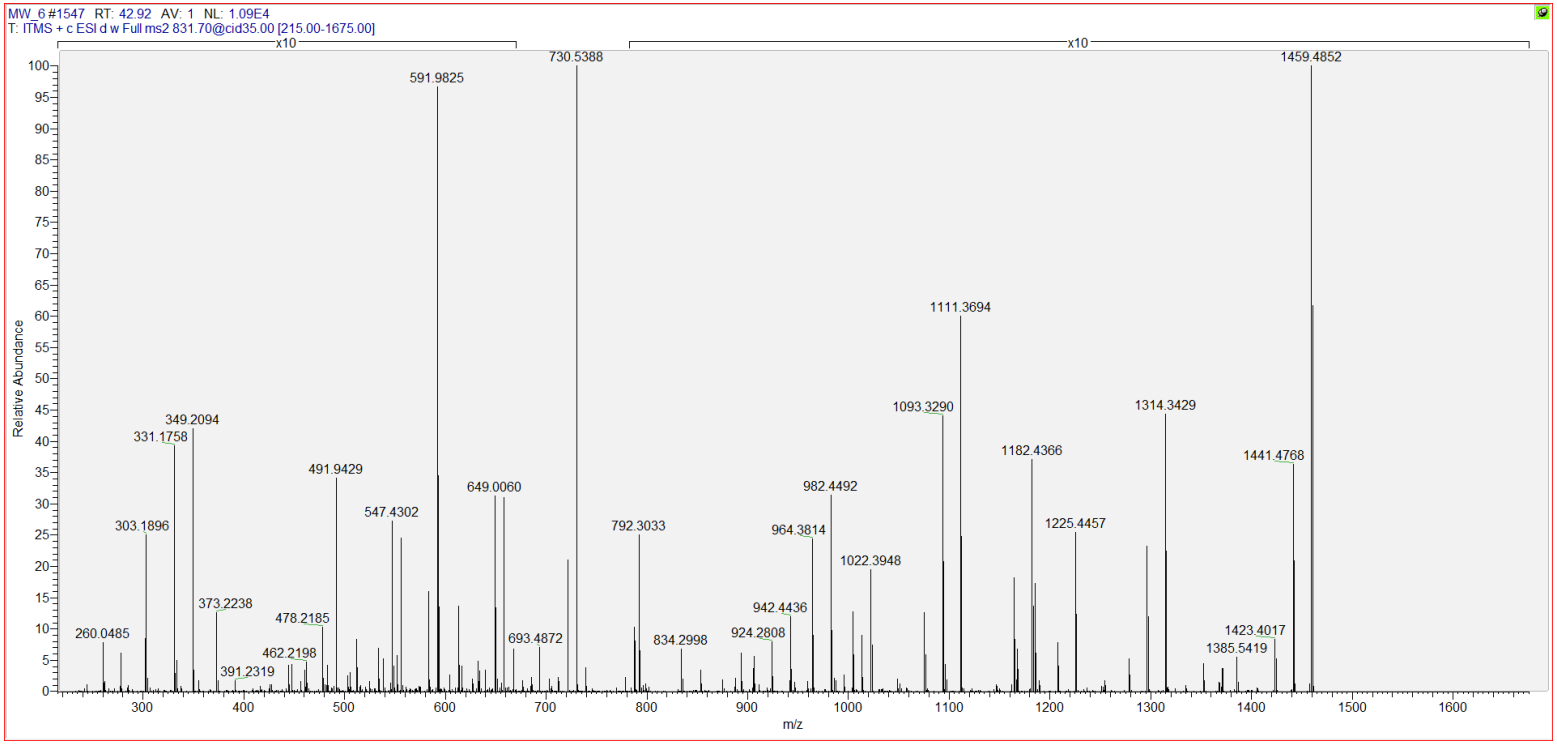
Appendix 9-7: MS spectra of peak MW4 in Appendix 9-3.



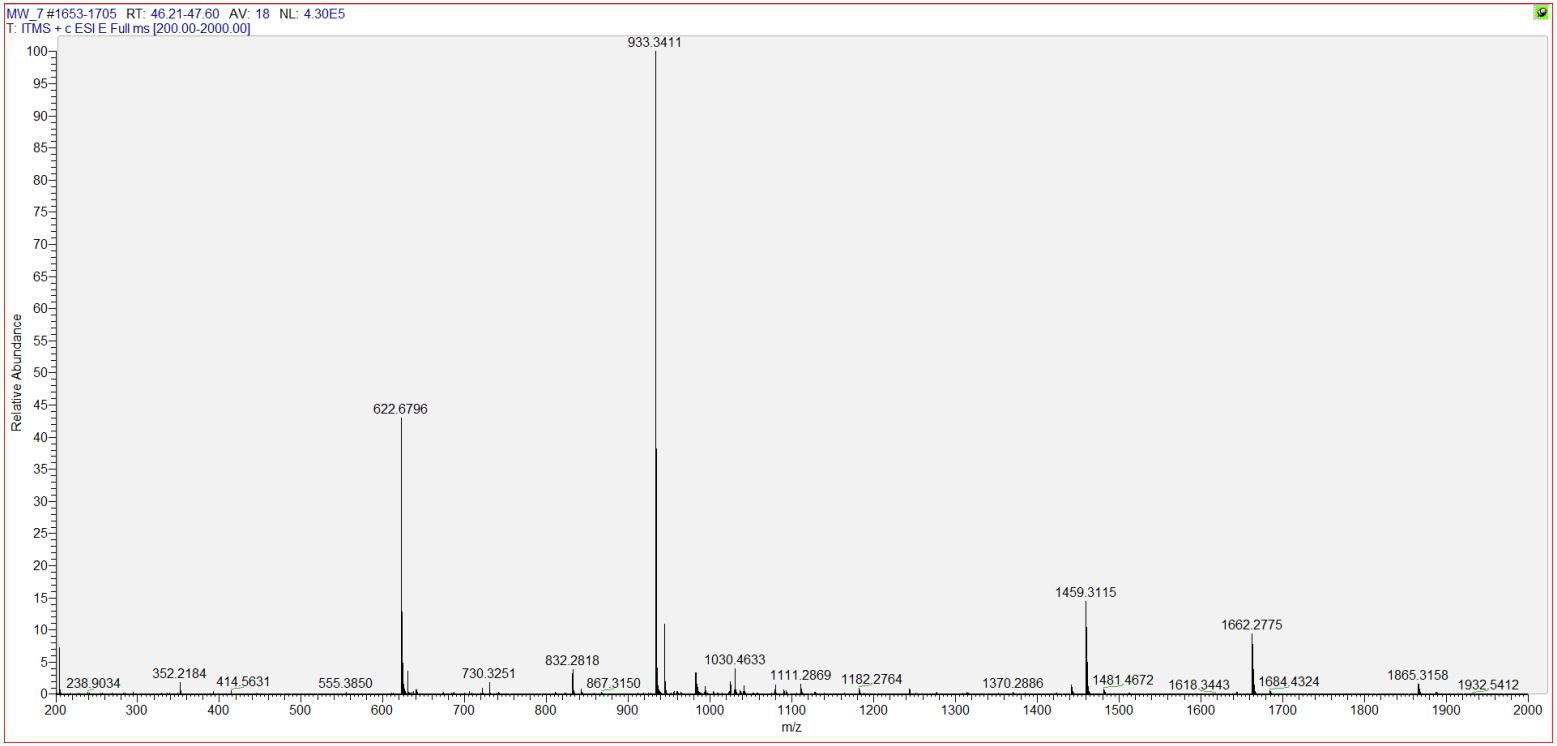
Appendix 9-8: MS spectra of peak MW5 in Appendix 9-3.



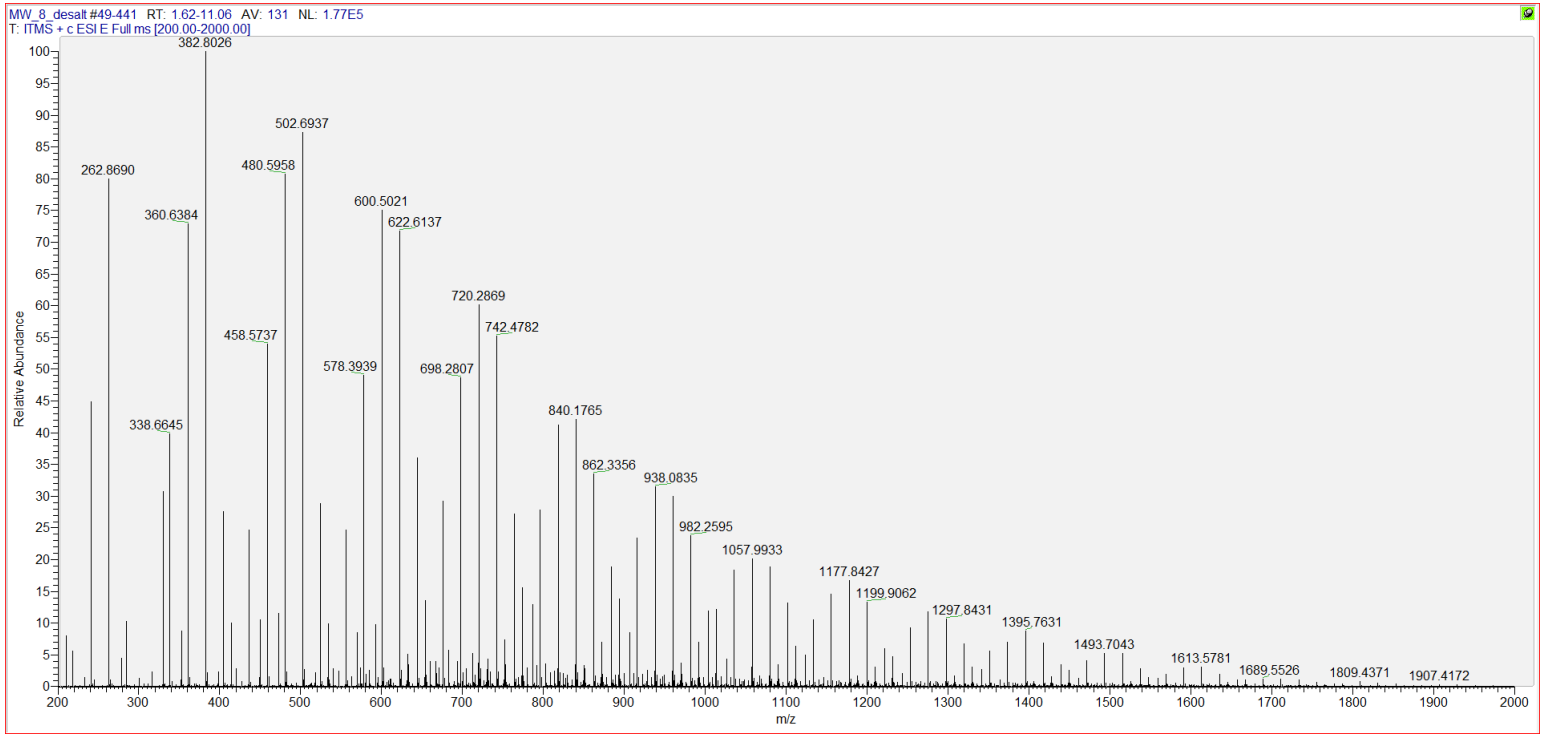
Appendix 9-9: MS spectra of peak MW6 in Appendix 9-3.



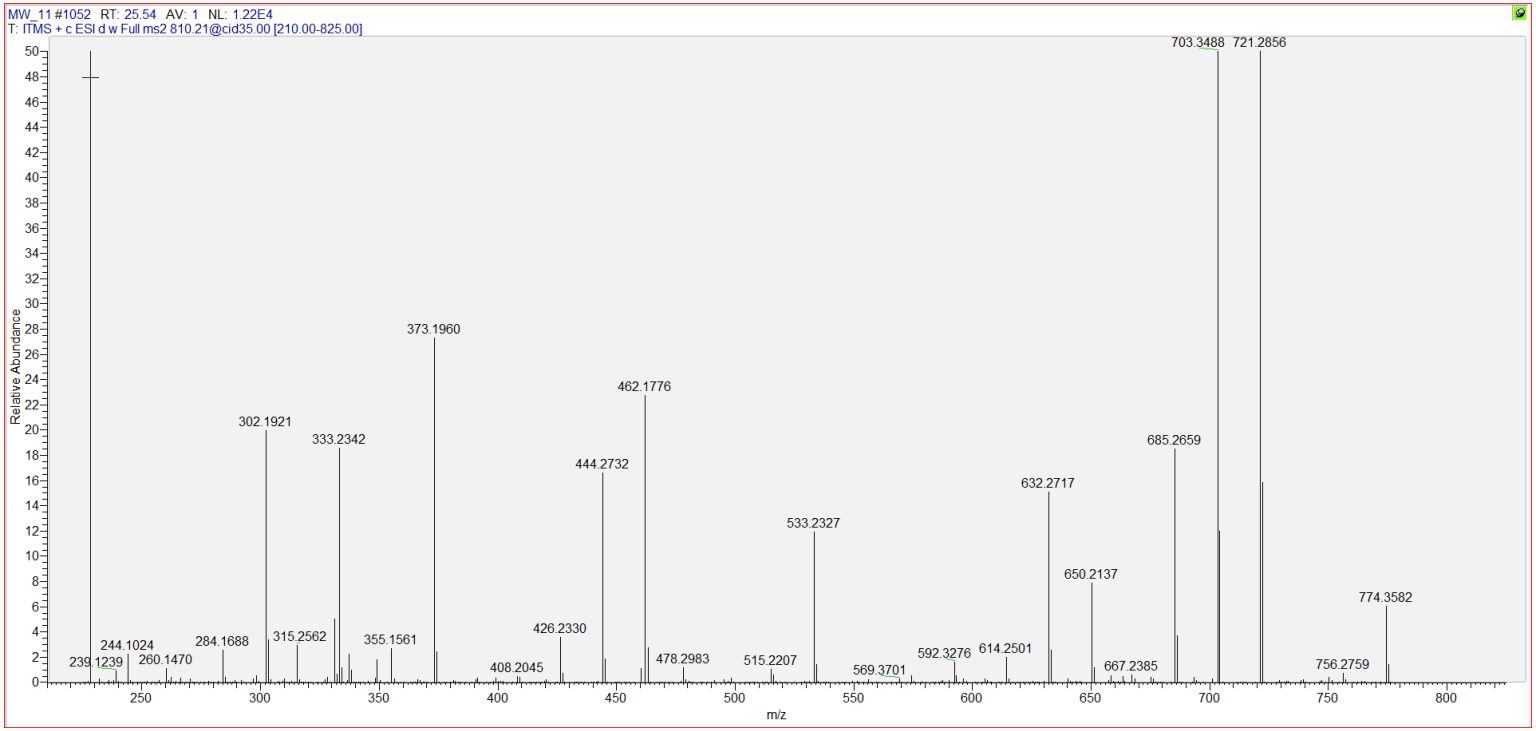
Appendix 9-10: MS spectra of peak MW7 in Appendix 9-3.



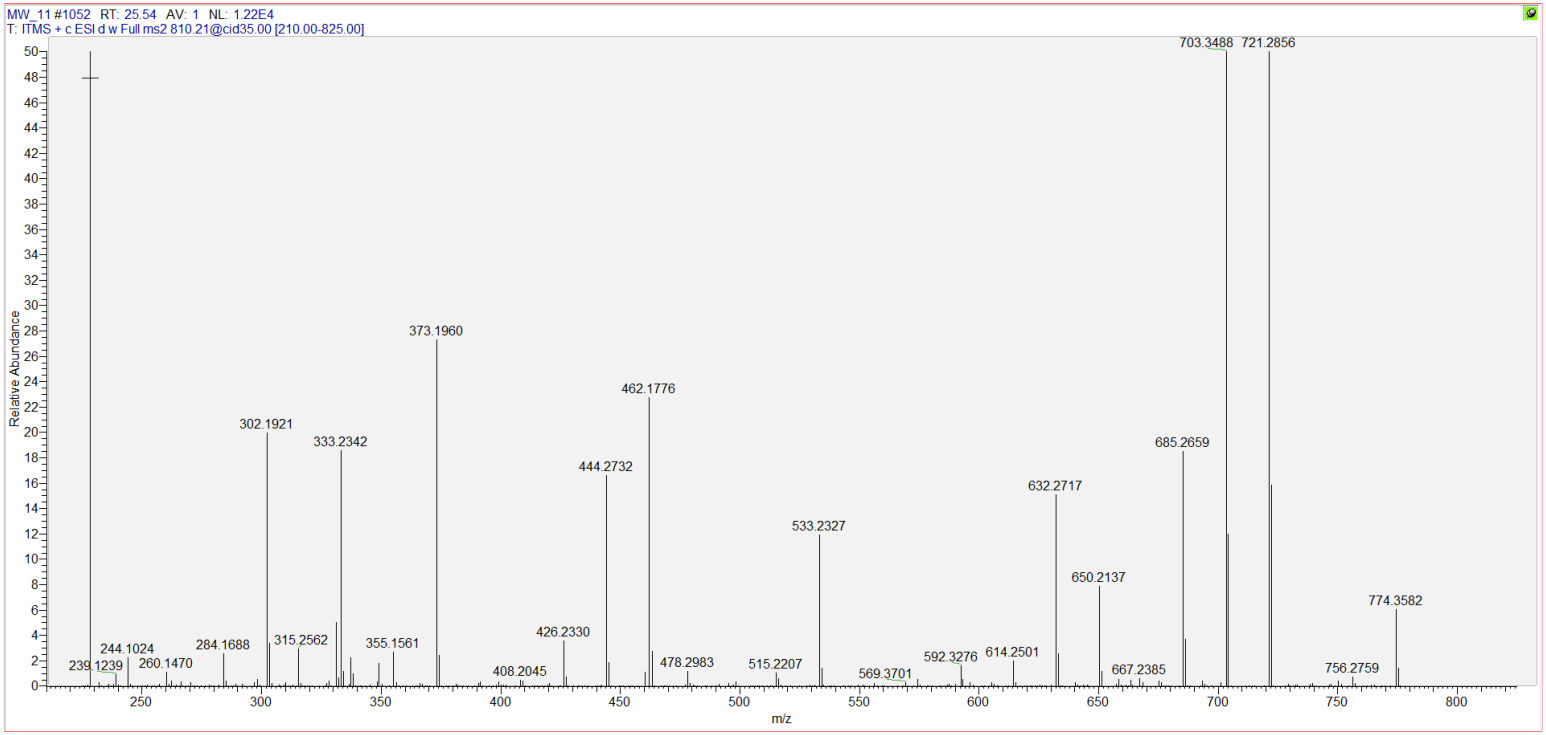
Appendix 9-11: MS spectra of peak MW8 in Appendix 9-3.



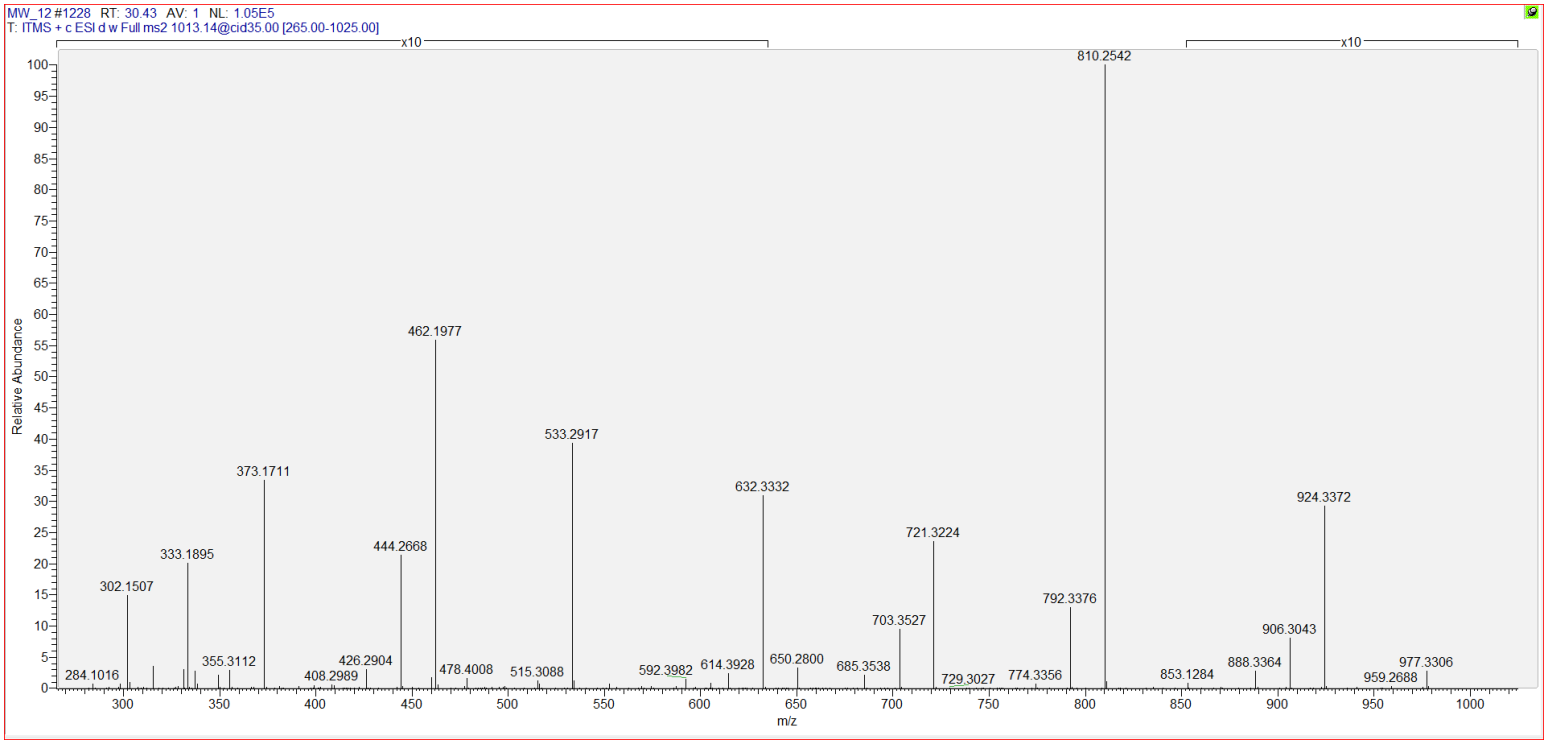
Appendix 9-12: MS spectra of peak MW9 in Appendix 9-3.



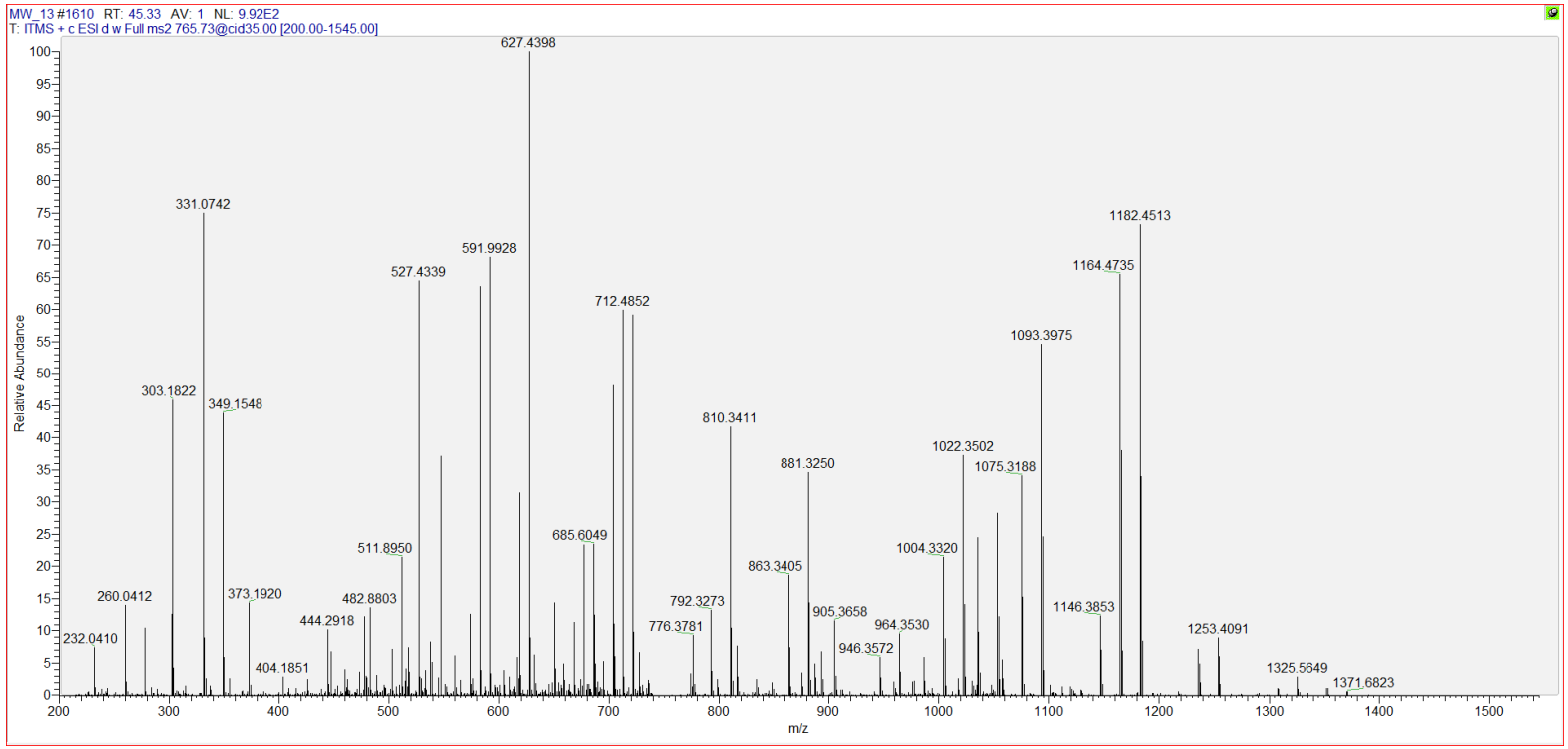
Appendix 9-13: MS spectra of peak MW11 in Appendix 9-3.



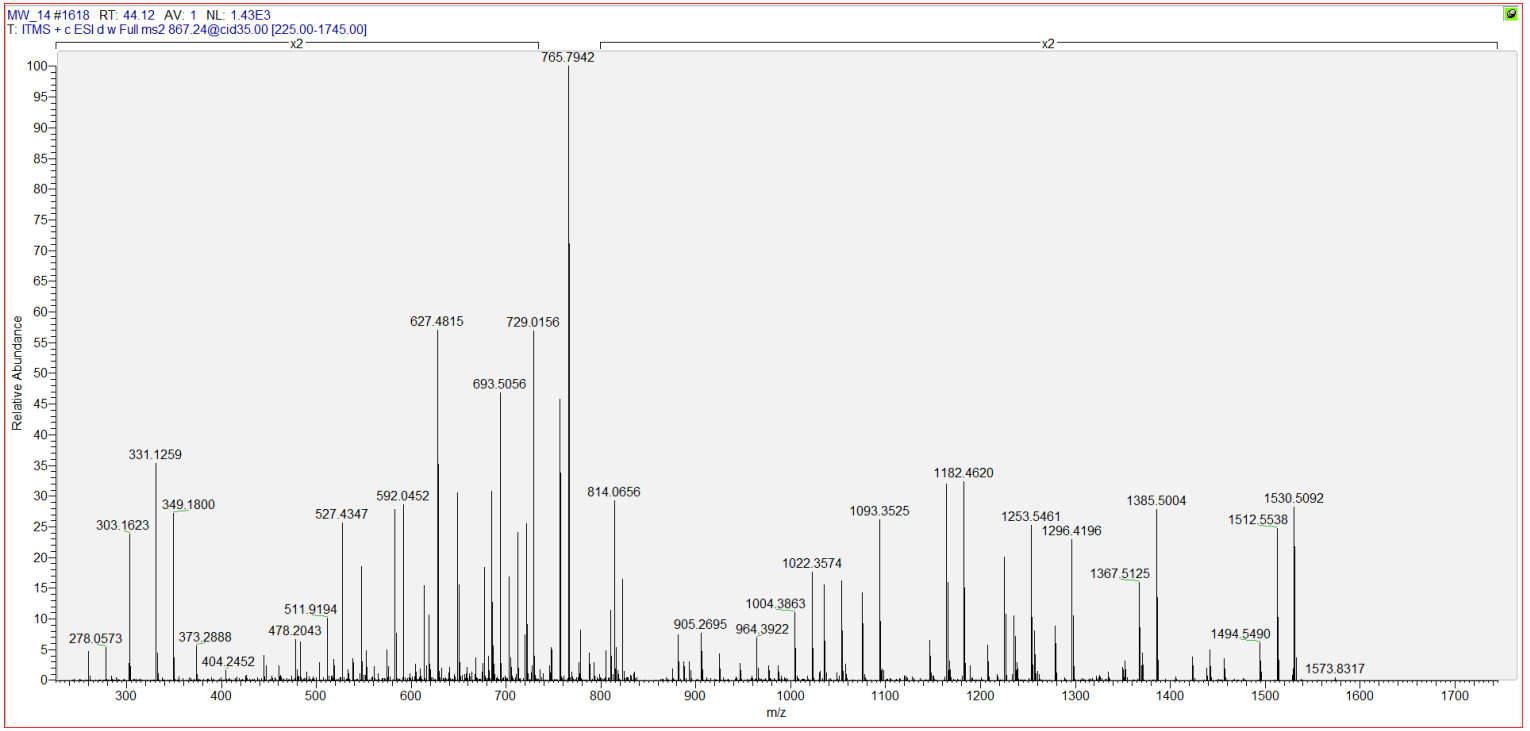
Appendix 9-14: MS spectra of peak MW12 in Appendix 9-3.



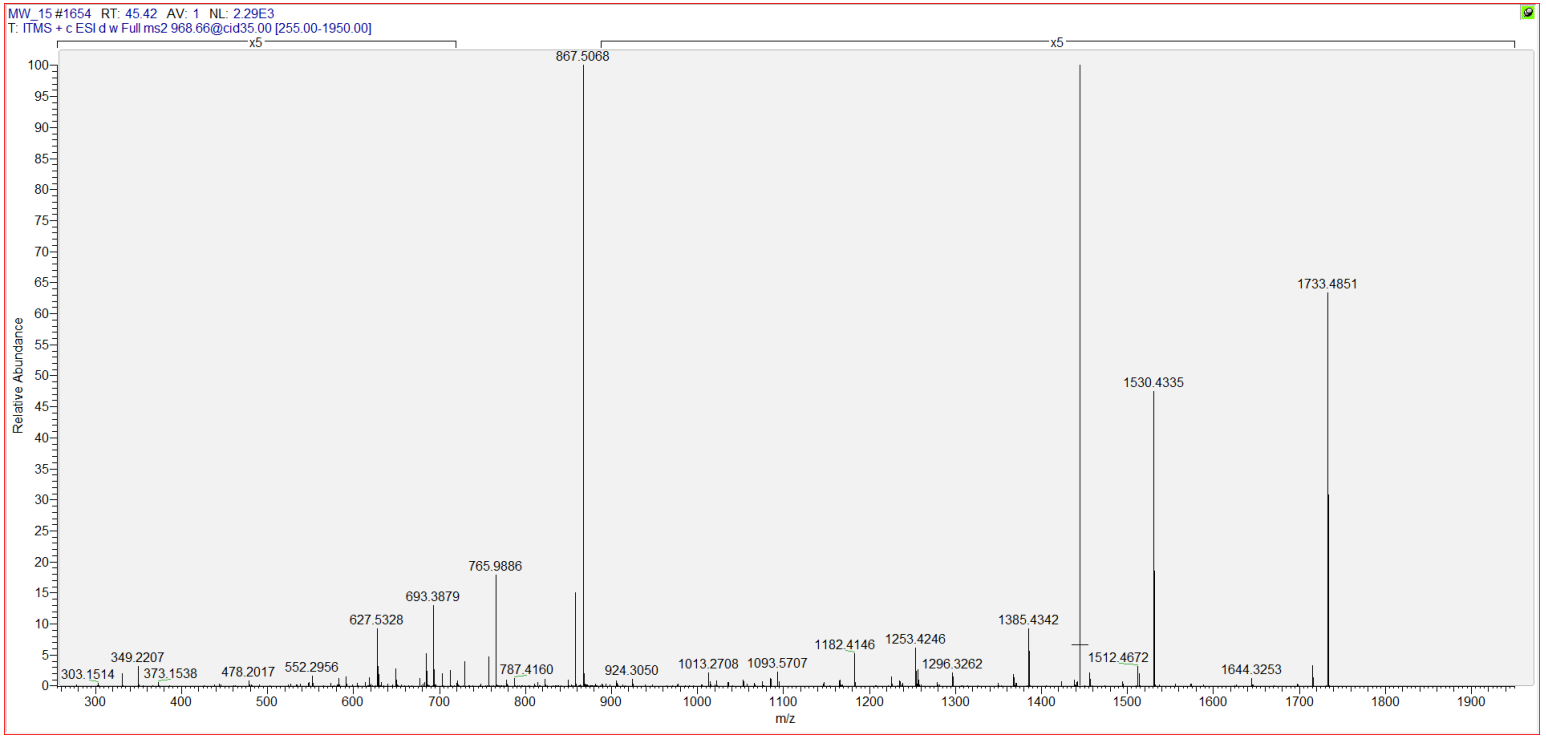
Appendix 9-15: MS spectra of peak MW13 in Appendix 9-3.



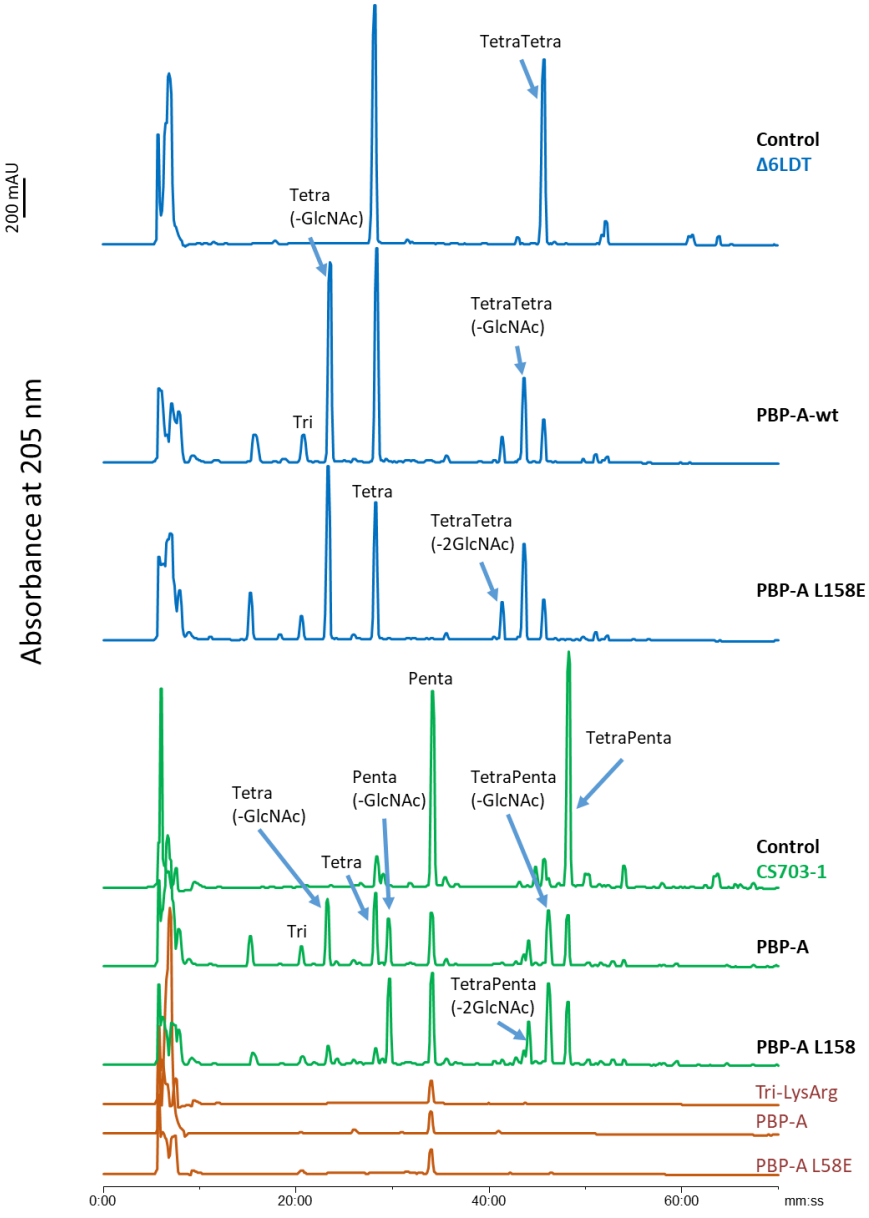
Appendix 9-16: MS spectra of peak MW14 in Appendix 9-3.



Appendix 9-17: MS spectra of peak MW15 in Appendix 9-3.



Appendix 9-18: MS spectra of activity assay of PBP-As on purified PG of different *E. coli* strains. Despite its detection in both assays of PBP-As against purified PG of *E. coli* Δ6LDT and CS703-1, no Glc*NAc* removal was observed in assay with TriLysArg muropeptides.



Appendix 9-19: Abundance of muropeptides detected in assay of PBP-A on PG extracts of *E. coli* Top10 at pH 7.5 and 5.

		pH 7.5			pH 5.0		
		E. coli Top10	E. coli Top10	E. coli Top10	*E. coli* Top10	*E. coli* Top10	*E. coli* Top10
		control	PBPA	PBPA L158E	control	PBPA	PBPA L158E
		sample1	sample2	sample3	sample1	sample2	sample3
Peak No.	**Muropeptide**	% Area	% Area	% Area	% Area	% Area	% Area
1	Tri	5.3	6.1	5.8	5.3	5.9	5.8
2	TetraGly 4	2.2	2.2	2.2	2.1	2.1	2.1
3	Tetra	38.3	41.4	38.1	39.3	41.6	39.1
4	Di	1.9	1.6	1.4	1.4	1.4	1.5
5	Tri-Lys-Arg	5.0	5.2	5.3	5.0	5.4	5.0
6	TetraTri(DAP)	2.6	2.7	2.8	3.1	2.6	2.5
7	TetraTetraGly4	1.6	1.6	1.5	1.6	1.2	1.2
8	TetraTri	2.4	2.7	3.1	2.7	2.2	2.5
9	TetraTetra	26.2	24.1	27.1	29.3	22.2	23.7
10	TetraPenta	0.6	0.5	0.6	0.5	0.4	0.5
11	TetraTetraTri	0.0	0.0	0.0	0.0	0.0	0.0
12	TetraTri-Lys-Arg	3.0	2.9	2.8	3.4	2.9	2.7
13	TetraTetraTetra	2.9	3.1	2.8	2.9	3.4	3.7
14	TetraTriAnh I	0.0	0.0	0.0	0.3	0.3	0.3
15	TeraTriAnh II	0.0	0.0	0.3	0.0	0.0	0.0
16	TetraTetraAnh I	0.6	0.0	0.5	0.6	0.4	0.4
17	TeraTetraAnh II	1.1	1.0	1.0	1.1	1.0	1.1
18	TetraTetraTetra Anh	1.1	1.0	0.8	1.2	0.8	0.7
19	Unknown	2.8	1.8	1.0	0.0	2.5	2.6
1 - 19	all known	97.5	98.0	96.9	99.7	96.0	95.4
	monomers (total)	53.9	57.6	54.4	53.2	58.7	56.1
	monomer di	1.9	1.6	1.5	1.4	1.4	1.5
	monomer tri	5.4	6.2	6.0	5.3	6.2	6.1
	monomer tetra	39.3	42.3	39.3	39.4	43.3	41.0
	monomer tetraGly4	2.2	2.2	2.2	2.1	2.2	2.2
	dimers (total)	39.1	36.3	40.8	42.7	34.4	36.6
	dimers(DD)	33.3	30.5	35.1	36.2	28.7	31.1
	dimers(LD)	2.7	2.8	2.8	3.1	2.7	2.6
	dimers anhydro	1.7	1.0	1.8	2.0	1.7	1.9
	Trimers (Total)	4.1	4.2	3.7	4.1	4.3	4.6
	trimer (anhydro)	1.1	1.0	0.8	1.2	0.8	0.7
	dipeptides (total)	1.9	1.6	1.5	1.4	1.4	1.5
	tripeptides (total)	8.0	9.0	9.2	8.4	8.8	8.9
	tetrapeptides (total)	77.6	77.6	78.5	80.2	77.1	77.3
	pentapeptides (total)	0.3	0.2	0.3	0.2	0.2	0.3
	3-3 linkage	1.4	1.4	1.4	1.6	1.4	1.3
	chain ends (anhydros)	1.2	0.9	1.2	1.4	1.1	1.2
	degree of cross-linkage	22.3	21.0	22.9	24.1	20.1	21.4
	% peptides in cross-links	46.1	42.4	45.6	46.8	41.3	43.9

Appendix 9-20: Abundance of muropeptides detected in assay of PBP-A on PG extracts of *E. coli* CS703-1 at pH 7.5 and 5.

		pH 7.5			pH 5.0		
		CS703-1	CS703-1	CS703-1	*CS703-1*	CS703-1	CS703-1
		control	PBPA	PBPA L158E	control	PBPA	PBPA L158E
		sample5	sample6	sample7	sample5	sample6	sample7
Peak No.	**Muropeptide**	% Area	% Area	% Area	% Area	% Area	% Area
1	Tri	0.5	0.8	0	0.5	0.8	0.58
2	TetraGly 4	0.5	0.4	0.75	0.4	0.5	0.42
3	Tetra	5.7	13.0	7.63	5.5	15.0	8.3
4	PentaGly5	1.85	1.67	1.7	1.82	1.76	1.79
5	Di	0.0	0.0	0	0.0	0.0	0
6	Penta	29.4	24.3	28.96	29.3	29.2	28.61
7	Tri-Lys-Arg	0.8	1.0	1.32	0.8	1.3	1.28
8	TetraPentaGly5	3.1	2.6	3.24	3.4	2.3	2.3
9	TetraTetra	4.8	4.9	5.28	4.8	3.7	3.63
10	TetraPenta	39.5	36.0	39.59	37.5	33.4	34.25
11	Unknown	0.5	1.5	0.51	0.3	0.4	2.32
12	Unknown	0.3	1.2	0	0.3	0.0	1.54
13	TetraTri-LysArg	0.0	0.0	0	0.3	0.3	0
14	TetraTetraTetra	0.3	0.5	0.35	0.3	0.9	0.77
15	TetraTetraPenta	2.8	2.6	2.91	3.0	2.7	2.39
16	Unknown	0.0	0.5	0.51	0.4	1.3	1.31
17	TetraPentaAnh I	2.0	1.6	1.76	2.0	1.5	1.97
18	TetraPentaAnh II	1.7	1.5	1.66	1.8	0.9	1.07
19	TetraTetraPentaAnh	1.1	1.1	0.67	0.9	0.7	0.71
1 - 19	all known	94.8	95.2	96.8	93.1	96.4	93.2
	monomers (total)	40.8	43.2	41.7	41.1	50.3	44.0
	monomer di	0.0	1.0	2.0	0.0	0.0	0
	monomer tri	0.5	0.9	0.0	0.5	0.8	0.6
	monomer tetra	6.0	13.6	7.9	5.9	15.6	8.9
	monomer tetraGly4	0.5	0.4	0.8	0.4	0.5	0.5
	monomer penta	31.0	25.5	29.9	31.5	30.3	30.7
	monomer pentaGly5	2.0	1.8	1.8	2.0	1.8	1.9
	dimers (total)	53.9	49.0	53.2	53.3	43.6	46.4
	dimers(DD)	53.9	49.0	53.2	53.3	43.6	46.4
	dimers(LD)	0.0	1.0	2.0	0.0	0.0	0
	dimers anhydro	3.9	3.3	3.5	4.0	2.4	3.3
	Trimers (Total)	4.5	4.4	4.1	4.5	4.4	4.2
	trimer (anhydro)	1.2	1.1	0.7	0.9	0.7	0.8
	dipeptides (total)	0.0	0.0	0.0	0.0	0.0	0
	tripeptides (total)	1.4	1.9	1.4	1.5	2.3	2.0
	tetrapeptides (total)	39.0	44.2	40.8	38.6	42.9	37.5
	pentapeptides (total)	58.8	50.5	56.8	58.8	53.0	54.9
	3-3 linkage	0.0	0.0	0.0	0.0	0.0	0
	chain ends (anhydros)	2.0	1.7	1.7	2.0	1.2	1.5
	degree of cross-linkage	30.0	27.4	29.3	29.7	24.7	25.9
	% peptides in cross-links	59.2	56.8	58.3	58.9	49.7	56.0

Appendix 9-21: Abundance (%) of all peptides detected in PG isolates of *E. coli* Top10 (control), *E. coli Top10/rhaA::PBP-A-wt, E. coli Top10/pBAD43_PBP-A-wt, E. coli Top10/rhaA::PBP-A-L158E* and *E. coli Top10/pBAD43_PBP-A-L158E.*

		E. coli Top10				
		no PBP-A	**rhaA::pbp-a-wt**	**pBAD43_pbp-a-wt**	**rhaA::pbp-a-L158E**	**pBAD43_pbp-a-L158E**
		No copy	**Chromosomal**	**Plasmid**	**Chromosomal**	**Plasmid**
Peak No.	Muropeptide	% Area	% Area	% Area	% Area	% Area
1	Tri	4.97	4.35	8.26	4.38	8.22
2	TetraGly 4	1.77	1.81	2.44	1.69	2.18
3	Tetra	35.73	37.84	33.24	38.38	32.23
4	Di	1.11	1.05	1.17	1.08	1.2
5	Tri-Lys-Arg	5.98	6.14	6.37	5.9	5.1
6	TetraTri(DAP)Gly 4	0	0	0.44	0	0.4
7	TriTri(DAP)	0.46	0.4	1.15	0.38	1.04
8	TetraTri(DAP)	3.25	3.43	4.66	3.29	4
9	TetraTetraGly 4	1.66	1.55	2.08	1.48	2.59
10	TetraTri	2.04	1.79	2.83	1.76	3.21
11	TetraTetra	29.26	29.43	20.22	29.73	22.99
12	TetraPenta	1.01	0.87	1.58	0.84	1.12
13	TetraTetraTri	0.84	0.72	0.74	0.83	0.53
14	TetraTri-Lys-Arg	3.33	3.11	4.3	2.81	3.72
15	TetraTetraTetra	3.05	2.81	2.46	2.66	2.6
16	TetraTriAnh I	0	0	0.36	0	0.28
17	TeraTriAnh II	0	0	0	0	0
18	TetraTetraAnh I	0	0	0.31	0	0.33
19	TeraTetraAnh II	0.32	0.3	0.47	0.28	0.53
20	TetraTetraTetra Anh I	0.9	0.94	0.64	0.99	0.72
21	TetraTetraTetra Anh II	1.42	1.32	1.12	1.34	1.24
1 - 21	all known	95.68	96.54	93.72	96.48	92.99
	monomers (total)	51.79766	53.02465	54.92958	53.30638	52.61856
	monomer di	1.160117	1.087632	1.248399	1.119403	1.290461
	monomer tri	5.194398	4.505904	8.813487	4.539801	8.83966
	monomer tetra	37.34323	39.19619	35.46735	39.78027	34.65964
	monomer tetraGly4	1.849916	1.874871	2.6035	1.751658	2.344338
	dimers (total)	42.7153	41.93081	39.27657	41.6563	41.69266
	dimers(DD)	35.83821	35.15641	29.71618	35.33375	33.39069
	dimers(LD)	3.396739	3.552931	4.972258	3.410033	4.301538
	dimers anhydro	0.334448	0.310752	1.216389	0.290216	1.225938
	Trimers (Total)	5.006271	4.630205	4.097311	4.643449	4.14023
	trimer (anhydro)	0.940635	0.97369	0.682885	1.026119	0.774277
	dipeptides (total)	1.160117	1.087632	1.248399	1.119403	1.290461
	tripeptides (total)	8.732232	7.872385	14.49175	7.837548	14.17536
	tetrapeptides (total)	78.15113	79.23141	69.04965	79.87493	71.89304
	pentapeptides (total)	0.527801	0.45059	0.842936	0.435323	0.602215
	3-3 linkage	1.938754	1.983634	3.3344	1.901949	2.925046
	chain ends (anhydros)	0.480769	0.479939	0.835823	0.487148	0.871061
	degree of cross-linkage	24.69516	24.05221	22.36983	23.92378	23.60648
	% peptides in cross-links	48.20234	46.97535	45.07042	46.69362	47.38144

Appendix 9-22: Distibution of disulfide bonds and cysteine in class A serine β-lactamases. Table 1: Analysis cysteine and/or disulfide bond conservation in class A β-lactamases. Amino acid sequences of all known class A β-lactamases (1378 enzymes) in BLDB are analysed. Ambler numbering is used.

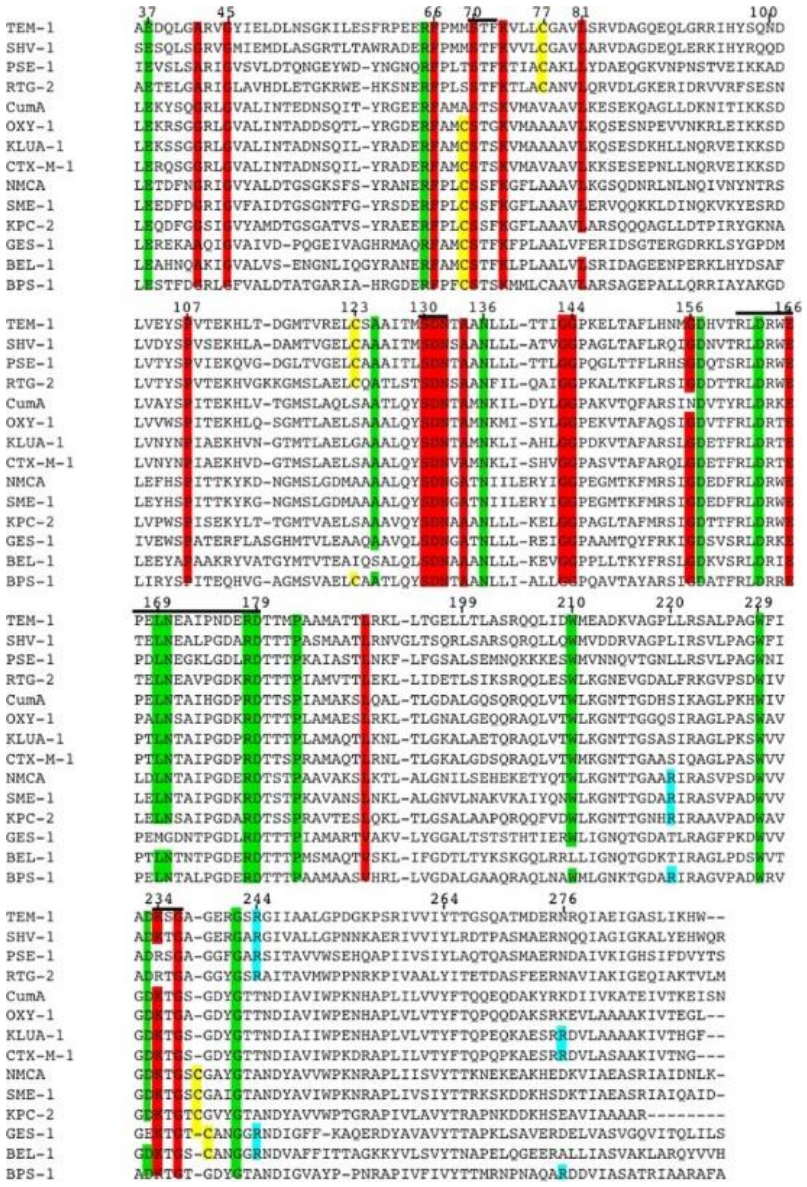
All class A BLs		
	Counts	Percentage
Cys77-Cys123	614	45%
Cys69-Cys238	205	15%
Cys69	402	29%
Cys123	35	3%
Cys77	21	2%
Cys72	1	0%
Cys220	1	0%
Cys135	63.0	5%
No Cysteine	36	3%
sum	1378	100%

Among No S-S bond enzymes		
	Counts	Percentage
Cys69	402	72%
Cys77	21	4%
Cys123	35	6%
Cys72	1	0%
Cys220	1	0%
Cys135	63	11%
No Cysteine	36	6%
sum	559	100%

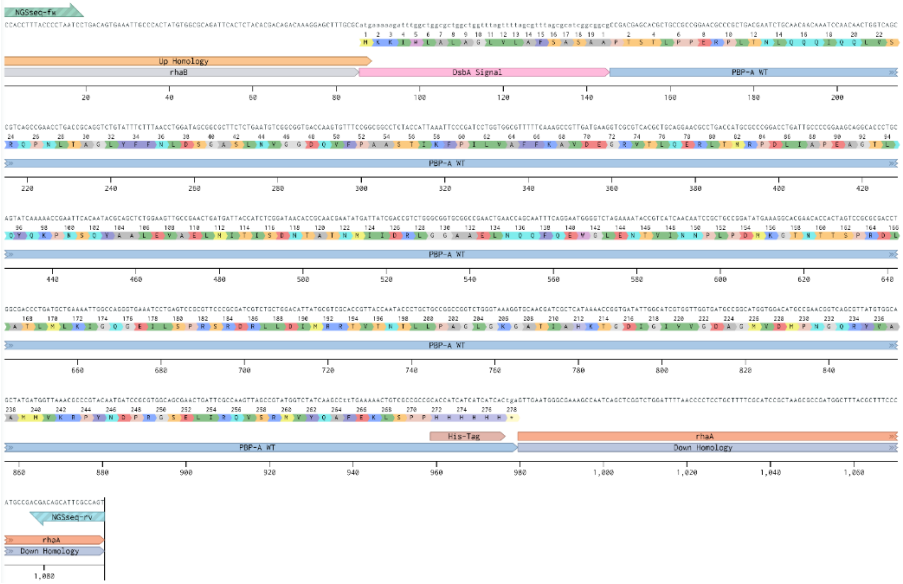
Appendix 9-23: Abundance (%) of all peptides detected in PG isolates of different *E. coli* strains expressing PBP-As with and without Cys68-Cys115 disulfide bond.

		E. coli Top10 (wild type)	E. coli Top10 pBAD43_PBP-A-WT	E. coli Top10 pBAD43_PBP-A-WT-SS	E. coli Top10 pBAD43_PBP-A-L158E	E. coli Top10 pBAD43_PBP-A-L158E-SS
Peak No.	Muropeptide	% Area	% Area	% Area	% Area	% Area
1	Tri	5.0	8.3	4.8	8.2	5.5
2	TetraGly 4	1.8	2.4	2.0	2.2	1.9
3	Tetra	35.7	33.2	39.6	32.2	38.3
4	Di	1.1	1.2	1.2	1.2	1.3
5	Tri-Lys-Arg	6.0	6.4	6.5	5.1	5.2
6	TetraTri(DAP)Gly 4	0.0	0.4	0.0	0.4	0.0
7	TriTri(DAP)	0.5	1.2	0.5	1.0	0.5
8	TetraTri(DAP)	3.3	4.7	3.9	4.0	3.4
9	TetraTetraGly 4	1.7	2.1	1.5	2.6	1.7
10	TetraTri	2.0	2.8	1.6	3.2	1.9
11	TetraTetra	29.3	20.2	26.9	23.0	28.5
12	TetraPenta	1.0	1.6	1.0	1.1	0.8
13	TetraTetraTri	0.8	0.7	0.8	0.5	0.7
14	TetraTri-Lys-Arg	3.3	4.3	2.9	3.7	2.6
15	TetraTetraTetra	3.1	2.5	2.5	2.6	2.8
16	TetraTriAnh I	0.0	0.4	0.0	0.3	0.0
17	TeraTriAnh II	0.0	0.0	0.0	0.0	0.0
18	TetraTetraAnh I	0.0	0.3	0.0	0.3	0.0
19	TeraTetraAnh II	0.3	0.5	0.3	0.5	0.3
20	TetraTetraTetra Anh I	0.9	0.6	0.9	0.7	1.0
21	TetraTetraTetra Anh II	1.4	1.1	1.2	1.2	1.3
1 - 21	all known	95.7	93.7	96.8	93.0	96.1
	monomers (total)	51.8	54.9	55.9	52.6	54.3
	monomer di	1.2	1.2	1.3	1.3	1.3
	monomer tri	5.2	8.8	4.9	8.8	5.7
	monomer tetra	37.3	35.5	41.0	34.7	39.8
	monomer tetraGly4	1.8	2.6	2.0	2.3	2.0
	dimers (total)	42.7	39.3	39.3	41.7	40.6
	dimers(DD)	35.8	29.7	32.3	33.4	34.4
	dimers(LD)	3.4	5.0	4.0	4.3	3.5
	dimers anhydro	0.3	1.2	0.3	1.2	0.3
	Trimers (Total)	5.0	4.1	4.3	4.1	4.7
	trimer (anhydro)	0.9	0.7	0.9	0.8	1.0
	dipeptides (total)	1.2	1.2	1.3	1.3	1.3
	tripeptides (total)	8.7	14.5	8.5	14.2	9.2
	tetrapeptides (total)	78.2	69.0	78.0	71.9	79.3
	pentapeptides (total)	0.5	0.8	0.5	0.6	0.4
	3-3 linkage	1.9	3.3	2.3	2.9	2.0
	chain ends (anhydros)	0.5	0.8	0.5	0.9	0.5
	degree of cross-linkage	24.7	22.4	22.5	23.6	23.4
	% peptides in cross-links	48.2	45.1	44.1	47.4	45.7

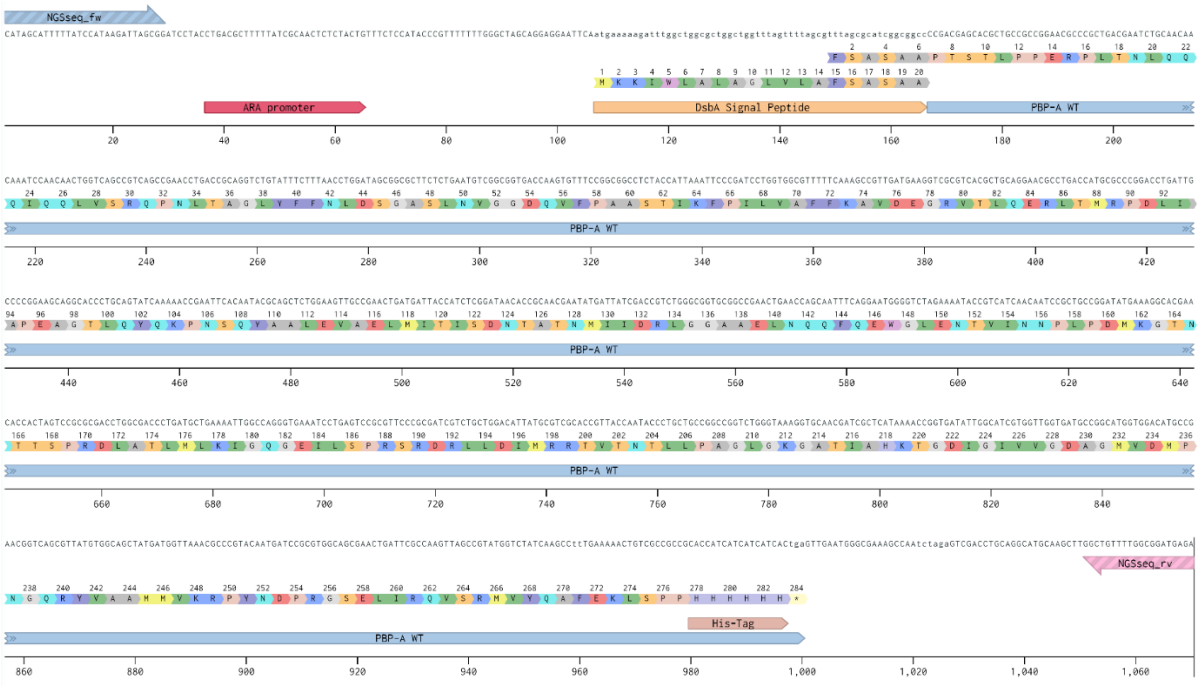
Appendix 9-24: Amino acid sequence alignment of 15 class A β-lactamases. Cysteine residues are highlighted in yellow, which show conservation of 77-123 or 69-238 disulfide bonds in many class A β-lactamases. However, some member of this class has no disulfide bond, such as CumA enzyme. Figure is taken from Philippon *et al.,* 2016.



Appendix 9-25: Sequence of the amplicon used in NGS sequencing of *php-a* libraries in single-copy (chromosomal) context. A window of 1070 bp, was amplified that contains extra bases upstream (106 bp) and downstream (70 bp) of *DsbA_php-a_His* to avoid the target gene in extremities of amplicons.



Appendix 9-26: Sequence of the amplicon used in NGS sequencing of *pbp-a* libraries in low-copy (plasmid) context. A window of 1070 bp, was amplified that contains extra bases upstream (106 bp) and downstream (70 bp) of *DsbA_pbp-a_His* to avoid the target gene in extremities of amplicons.



Appendix 9-27: Distribution of relative fitness (W_{rel}) of nonsynonymous substitutions on the PBP-A-L158E protein under different conditions of the drift. N-terminal *DsbA* signal peptide, which was not subjected to epPCR is indicated as SP. Dashed lines represent the thresholds for beneficial (W_{rel} >1.3) and deleterious (W_{rel} < 0.6) fitness values. Color code on the graph represent the types of substituted amino acids.