UCLouvain

LIBST
LOUVAIN INSTITUTE OF BIOMOLECULAR SCIENCE AND TECHNOLOGY

FACULTE DES SCIENCES

Faculté des sciences

Study of the covalent anchoring of proteins in the cell wall of *Streptococcus thermophilus*

Auteur : Martin Desmet

Promoteur : Pr. Patrice Soumillion

Co-promoteur : Pr. Pascal Hols

Lecteurs : Pr. Pierre Morsomme et Pr. Benoît Desguin

Encadrant : Alessandro Ratto

Mémoire en vue de l'obtention du diplôme de master en biochimie et biologie moléculaire et cellulaire

Année académique 2020-2021

Remerciements

Avant toute chose, je tiens à remercier toutes les personnes m'ayant aidé dans la réalisation de ce mémoire ainsi que tout au long de mon parcours universitaire.

Un grand merci à mon promoteur, le Professeur Patrice Soumillion, pour m'avoir accueilli au sein de son laboratoire, pour son soutien et son attention portée envers ce projet ainsi que pour ses remarques judicieuses qui m'ont permis de réaliser ce mémoire.

Merci évidemment à Alessandro pour son encadrement tout au long de ce projet. Merci à lui pour sa patience, sa disponibilité et ses conseils avisés m'ayant permis de réaliser ce mémoire mais aussi de me former au travail en laboratoire.

Je remercie également les Professeurs Pierre Morsomme et Benoît Desguin pour avoir accepté d'être les lecteurs et jury de ce mémoire.

Je tiens aussi à remercier tous les membres du BGM et plus particulièrement Brigitte et Damien pour leurs précieux conseils et leur disponibilité tout au long de mon passage au laboratoire. J'ai particulièrement apprécié l'ambiance de travail au sein de ce laboratoire ainsi que le climat d'entraide fort utile.

Merci également à mes co-mémorants Alexander, Amandine et Denis, avec qui j'ai passé d'agréables moments durant ce mémoire mais aussi à tous les étudiants rencontrés lors de ces 5 dernières années.

Pour terminer, je tiens à remercier ma famille, mes amis et ma copine Clélia pour leurs encouragements et leur soutien qui m'ont permis de réaliser ce parcours universitaire.
Je tiens plus particulièrement à remercier ma sœur Elise Desmet ainsi que Emile Neimry pour leur relecture assidue de ce manuscrit.

Résumé

L'évolution dirigée est un domaine de recherches en pleine expansion, avec de plus en plus d'outils et d'applications développées. Cela consiste à générer des banques contenant des versions mutées aléatoires d'un gène. Les protéines codées par chaque version du gène sont ensuite sélectionnées ou criblées pour une propriété souhaitée et les gènes codant pour la protéine d'intérêt sont récupérés. De nombreuses découvertes ont été faites dans ce domaine pour des applications industrielles, de recherches ou même thérapeutiques. Ces succès ont permis à l'évolution dirigée d'être récompensée en 2018 par le prix Nobel de chimie. La moitié du prix a été attribuée à George Smith et Sir Gregory P. Winter pour leurs travaux sur le phage display. Le phage display est une technologie qui permet l'évolution *in vitro* de protéines de liaison telles que les anticorps en fusionnant génétiquement les protéines de la banque à une protéine de la capside d'un bactériophage. Les protéines affichées sont physiquement liées à leur gène codant à travers la particule de phage et sont disponibles pour une sélection *in vitro* par capture d'affinité. Cette technologie a notamment été déterminante pour le développement des anticorps thérapeutiques.
Les plateformes de display ont été développées avec des entités supramoléculaires, telles que l'affichage de phages mais aussi avec des entités cellulaires, telles que yeast surface display.
La technologie de display est un outil très utile pour l'évolution dirigée. En effet, le criblage ou la sélection des variantes de la protéine qui ont les propriétés souhaitées peuvent être effectués directement sur la protéine fusionnée à la plateforme de display. Cela facilite la récupération et l'amplification des gènes codant pour la protéine sélectionnée.
L'objectif principal de ce mémoire est de développer une nouvelle plateforme de display qui pourrait être utile dans des expériences d'évolution dirigée utilisant *Streptococcus thermophilus*.
S. thermophilus pourrait être un bon candidat pour le display de protéines recombinantes en raison de sa compétence naturelle très efficace qui facilite la construction d'une bibliothèque génétique très diversifiée (jusqu'à 10^9 variantes). En tant que bactérie gram-positive, l'absence de membrane externe devrait également faciliter grandement l'encrage sur la surface extérieure. Afin de construire ce système de display, la souche *Streptococcus thermophilus LMD-9* sera utilisée car l'efficacité de sa compétence naturelle est extrêmement élevée, avec jusqu'à 10% des cellules totales qui peuvent être transformées après une simple incubation des bactéries avec du XIP et de l'ADN linéaire. De plus, la souche LMD-9 exprime un homologue fonctionnel de la sortase A, qui ancre de manière covalente trois protéines différentes sur la paroi cellulaire. L'une de ces protéines est une sérine protéinase appelée PrtS. Cette protéine comprend un peptide signal spécifique pour sa translocation via le sec pathway, et un motif LPNTG pour l'ancrage par sortase au peptidoglycan. Notre stratégie pour afficher des protéines recombinantes à la surface de *S. thermophilus* consiste à remplacer la séquence du gène PrtS par la séquence de la protéine d'intérêt. Le peptide signal et le motif d'ancrage de PrtS sont conservés et fusionnés avec la séquence de la protéine.
Comme preuve de concept, cette plateforme de display sera utilisée pour afficher l'alpha-amylase de *Bacillus licheniformis*, une enzyme naturellement sécrétée.
Un objectif secondaire sera de développer une système de display à partir de laquelle la protéine peut être détachée de la bactérie avec un agent réducteur. À cette fin, un module BIL (bacterial extracellular intein-like) sera inséré entre le signal d'ancrage et la protéine d'intérêt. Ce module peut s'auto-exciser d'une protéine en laissant un lien disulfure entre les deux polypeptides résultant. Un simple traitement des cellules avec un agent réducteur devrait alors entraîner la libération de la protéine du peptidoglycan. Un tel système clivable peut trouver des applications utiles dans des campagnes d'évolution dirigées spécifiques qui tireraient parti de la libération contrôlée de la protéine exposée.

Abstract

Directed evolution is a growing research field with more and more tools and applications being developed. It consists of the generation of a library containing randomly mutated versions of a gene. The proteins encoded by each version of the gene are then selected or screened for a desired property and the genes coding for the protein of interest are recovered. The cycles of genetic diversification and screening/selection may be iterated for progressive improvement of the proteins. Many discoveries have been made in this field for industrial, research or even therapeutic applications. Those succesess have led directed evolution to be awarded in 2018 with the Nobel Prize in Chemistry. Half of the prize was awarded to George Smith and Sir Gregory P. Winter for their work on phage display. Phage display is a technology that allows the *in vitro* evolution of binding proteins such as antibodies by genetically fusing the proteins of the library to a coat protein of a bacteriophage. The displayed proteins are physically linked to their encoding gene through the phage particles and are available for *in vitro* selection by affinity capture. This technology has notably been key for the development of modern therapeutic antibodies.

Display platforms have been developped with supramolecular entities, such as phage display, mRNA display and ribosome display, but also with cellular entities, such as yeast surface display, *Escherichia coli* surface display, etc.

Display technologies are a very useful tool for directed evolution. As a matter of fact, the screening or selection of the protein variants that have the desired properties can be performed directly on the protein fused to the display platform. This makes it easier to retrieve and amplify the genes coding for the selected protein.

The main objective of this master thesis is to develop a new display platform that could be useful in directed evolution experiments using *Streptococcus thermophilus*.

Streptococcus thermophilus could be a good host for the displaying of recombinant proteins because of its highly efficient natural competence that facilitates the construction of highly diverse genetic libraries (up to 10^9 variants). As a gram-positive bacterium, the absence of an outer membrane should also greatly facilitate surface display. In order to build this display platform, the *Streptococcus thermophilus* LMD-9 strain will be used because the efficiency of its natural competence is extremely high, with up to 10% of total cells that can be transformed upon simple incubation of bacteria with XIP and linear DNA. Moreover, the LMD-9 strain encodes for a functional sortase A homolog that covalently anchors three different proteins on the cell wall. One of these proteins is a serine proteinase called PrtS. This protein comprises a sec-specific signal peptide that allows its translocation through the membrane, and an LPXTG motif for sortase-anchoring to the cell-wall.

Our strategy for displaying recombinant proteins on the surface of *Streptococcus thermophilus* is to replace the sequence of the PrtS gene by the sequence of the protein of interest. The signal peptide and the anchoring motif of PrtS are retained and fused in frame with the sequence of the protein. As proof-of-concept, this display platform will be used to display the alpha-amylase from *Bacillus licheniformis*, a naturally secreted enzyme.

A secondary objective will be to develop a display platform from which the protein can be detached from the bacteria with a reductive agent. To this aim, a bacterial extracellular intein-like (BIL) module will be inserted between the anchoring signal and the protein of interest. This module can auto-cleave itself from the precursor protein and leave a disulfide bond between both flanking polypeptides. A simple treatment of the cells with a reducing agent should then result in the release of the protein from the peptidoglycan. Such cleavable system may find useful applications in specific directed evolution campaigns that would take advantage of the controlled release of the displayed protein.

Abbreviations

5'FOA	5-fluoroorotic acid
BET	Ethidium bromide
BIL	Bacterial extracellular intein like
BME	β-mercaptoethanol
bp	base pairs
C°	Celcius
cat	*chloramphenicol acetyltransferase*
cfu	colony forming units
C-terminus	Carboxy-terminus
dNTP	Desoxyribonucleotide triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence activated cell sorting
Ff	F-pilus specific filamentous phage
GCDM	Chemically defined medium supplemented with glucose
GFP	Green fluorescent protein
GM17	M17 medium supplemented with glucose
GRAS	Generally Recognized As Safe (FDA statu)
IPTG	Isopropyl β-d-1-thiogalactopyranoside
kDa	kiloDalton
LB	Luria-Bertani
N-terminus	Amino-terminus
$OD_{(600nm)}$	Optical density at 600nm
POI	Protein of interest
RPM	Rotation per minute
S. thermophilus	*Streptococcus thermophilus*
Sec	Secretion

SRP	Signal recognition particle
Tat	Twin-arginine translocation
tRNAser	transfer RNA serine
WT	Wild type
XIP	ComX inducing peptide

Table of content

1.Introduction

1.1 Directed evolution and display technologies

Directed evolution is a method used in the laboratory to modify or enhance the activity of a biomolecule. It has been used for industrial, research or even therapeutic applications with a lot of success (Porter, Rusli and Ollis, 2016; Arnold, 2018). This technology was awarded the Nobel Prize for Chemistry in 2018 (*Press release: The Nobel Prize in Chemistry 2018 - NobelPrize.org*)

This method imitates the natural evolution processes of mutation and selection but focuses on a single gene/protein with the objective of creating biomolecules with a desired activity. It all starts with the elaboration of a gene library coding for randomly or semi-randomly mutated variants of a protein. This library is then translated *in vivo* or *in vitro* and scientists search for the variants with the desired properties using an appropriate screening or selection method (Figure 1A). The desired property can be a specific enzymatic activity, a specific molecular recognition, an increased stability in specific conditions or even a combination of properties. The fitness of a protein is defined in relation to a desired property. Proteins with the best fitness can be a protein with the best activity or a protein with a more specific activity or even the more stable protein. Mutations that confer a lower fitness than the original gene are discarded, and a new round of directed evolution with the retained variants can be performed for further improvements. These cycles of directed evolution are repeated until the fitness of the protein no longer increases. When the fitness of a protein has reached a maximum, it is however unknown whether the protein could have reached a higher maximum using a different evolutionary trajectory (Figure 1B).

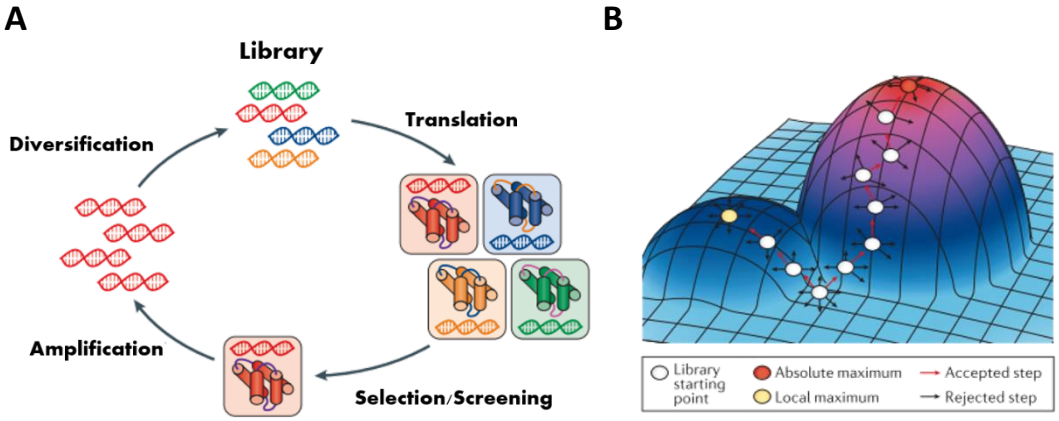

Figure 1. Directed evolution: (A) General steps of directed evolution **(B)** Fitness Landscape of a library. The fitness landscape is a 3D surface that schematizes the multidimensional space of all the possible variants in the x and y axis and the fitness of the library in the z axis.
Adapted from (Packer and Liu, 2015)

Over the past few decades, directed evolution has been a growing research field with more and more tools and applications being developed.
One very useful tool is the development of display technologies. It consists in making a physical link between the protein which is accessible for screening/selection and the gene that encodes it. This facilitates the manipulation, analysis, retrieval, and amplification of the genes coding for the evolving

proteins. These display platforms have been developed *in vivo* with competent cells that can express the protein of interest on its outer membrane/layer. Cell-free *in vitro* technologies also exist.

1.1.1 Supramolecular display

Supramolecular display allows the directed evolution of peptides or proteins that bind to a specific target. The selection is based on simple affinity capture with an immobilised target. Since the protein is linked to its gene via a small entity like a virus or a ribosome, protein capture is associated to the capture of the gene that can be amplified *in vivo* (phage infection) or *in vitro* (PCR or RT-PCR) for fulfilling the directed evolution cycle (Figure 1A).

1.1.1.1 Phage display

Phage display is the most widespread technology and is now commonly used for creating antibodies or alternative proteins endowed with affinity towards a specific target. It takes advantage of the F-pilus specific filamentous phage (Ff) for expressing a recombinant protein fused to one of the coat proteins. The technology was originally described in 1985 by George Smith. Initially, he was only able to display a small peptide (Smith, 1985). Rapidly, improvements were made so that whole proteins could be displayed. In 1990, Sir Gregory P. Winter successfully displayed the single chain variable fragment, scFv (fusion of the heavy (V_H) and the light (V_L) variable domain) of the anti-lysozyme antibody (McCafferty *et al.*, 1990).
Phage display technology has permitted the discovery and the commercialization of 14 antibody-based drugs (Nixon, Sexton and Ladner, 2014; Barderas and Benito-Peña, 2019; Nagano and Tsutsumi, 2021). Those successes led the technology and its two major contributors (George Smith and Sir Gregory P. Winter) to be awarded in 2018 with half of the Nobel Prize in Chemistry, the other half being attributed to Frances H. Arnold for her work in directed evolution of enzymes (*Press release: The Nobel Prize in Chemistry 2018 - NobelPrize.org*).

The filamentous phage is a type of bacteriophage that does not kill its host-cell. After infection, the phage is secreted from the bacterium. The Ff filamentous phage is the most-characterized group. Three major strains of Ff phage have been identified: M13, fd and f1. These strains are specific to *Escherichia coli* (*E. coli*).
The M13 strain is the most commonly used strain for display technology. The capsid of this phage is only made of five different proteins (Kehoe and Kay, 2005)(Figure 2).

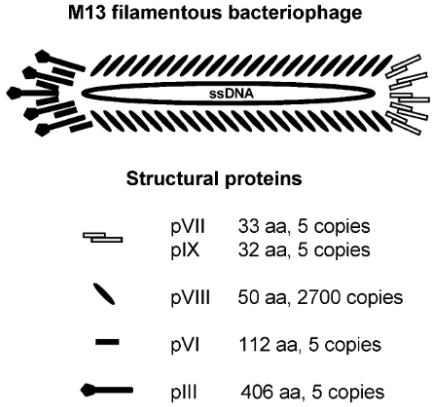

Figure 2. M13 capsid structure: The M13 phage capsid is made of five different proteins, each cap has five copies of two different proteins.
Adapted from (Kehoe and Kay, 2005)

Phage display platforms have been developed on each of the five coat proteins but those with pIII (Figure 3A) and pVIII (Figure 3B) have been the most commonly used.

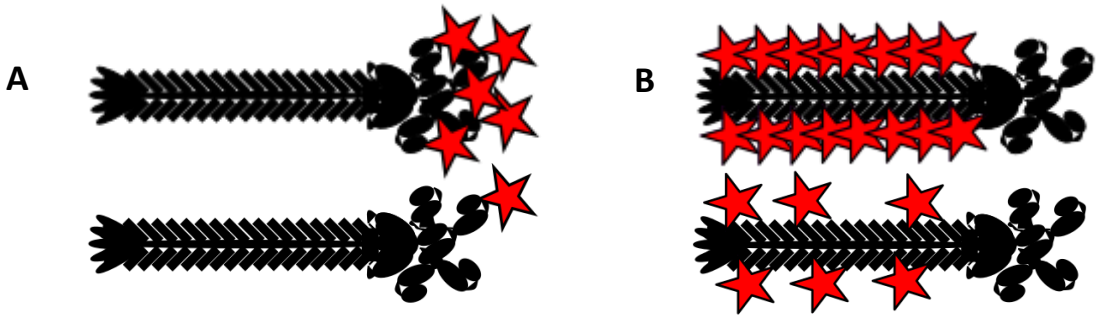

Figure 3. Types of phage display formats: (A) Display fusion with pIII. Using a helper plasmid, the display rate can be modulated **(B)** Display fusion with pVIII; red stars represent the recombinant displayed proteins or peptides
Adapted from *(*Rakonjac *et al.,* 2011*)*

a) pIII display

Each phage particle contains between 3 and 5 pIII proteins at one tip of the filament. The protein pIII has two different functions in the phage life cycle. The N-terminus of the protein contains two different domains, N1 and N2, both essential for host infection (Figure 4). Together, N1 and N2 interact with the F-pilus and TolQRA, an inner membrane complex of *E. coli,* and allow the internalization of the phage.
The C domain of pIII plays a role in the capsid assembly (Marvin, 1998; Rakonjac *et al.*, 2011).

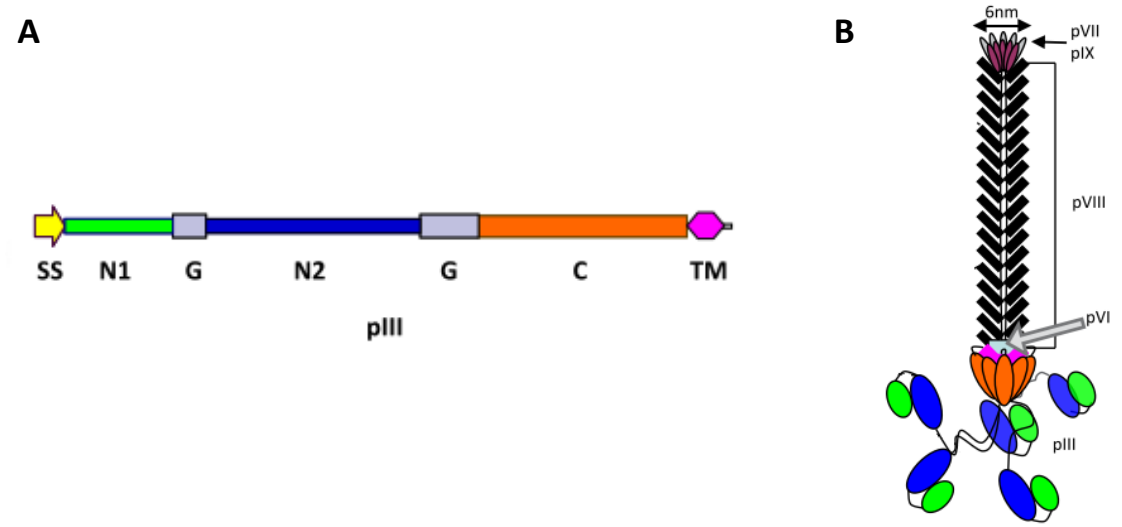

Figure 4. Structure of pIII phage display: (A) Domain organization of pIII. SS, signal sequence. N1, N2, C, Domain of pIII. G, glycine-rich linker. TM, Transmembrane helix **(B)** Structure of pIII in the phage capsid.
Adapted from *(*Rakonjac *et al., 2011)*

If the display platform affects one of those two functions, the resulting phage will not be selected among WT phages and can even not be propagated.
To avoid negative effects on the life cycle of the phage, two major solutions have been developed. The first is to fuse the sequence of the peptide in frame in the glycine-rich linker between the N2 and the C domain of pIII (Figure 4).

Another strategy is to fuse the peptide to a truncated pIII that has only the transmembrane helix and the C domain. This fusion-peptide only keeps the structural function of pIII. In order to develop a fully functional phage, a wild type (WT) pIII must also be expressed by the phage. The result is a phage that has pIII WT and truncated pIII fused to the displayed peptide. In most of the cases, the pIII-fusion protein is cloned into a phagemid vector and the WT pIII is provided by a phage helper vector transformed in *E. coli* and used to trigger the production of phage particles.

b) pVIII display

pVIII has a structural role in the virion capsid and is the major coat protein, expressed at several thousands of copies per phage. The number of copies per capsid is determined by the length of the phagemid encapsulated (Marvin, 1998; Kehoe and Kay, 2005). Polypeptide display is possible at the N-teminus of pVIII.
In order to display large peptides (bigger than 7 residues), a WT pVIII protein must be co-assembled with the fusion display protein, otherwise the phage capsid cannot properly assemble (Iannolo *et al.*, 1995). The construction design ensures that the WT pVIII promotor is more efficient than the promotor of the displayed fusion pVIII so that the virion's capsid is mostly composed of WT pVIII.

1.1.1.2 Ribosome display

Ribosome display is a display technique that can be performed entirely in a cell-free system. This platform produces a ternary complex between the ribosome, the nascent polypeptide and the mRNA that encodes the polypeptide (Figure 5). Therefore, if a protein is captured based on its affinity for a target, the mRNA will be captured as well and can be amplified by RT-PCR, making the whole cycle of directed evolution (Figure 1A) entirely *in vitro*.

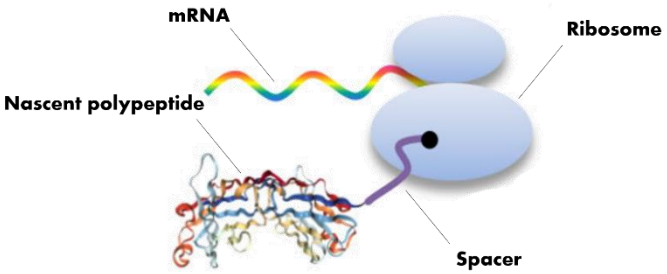

Figure 5. Ternary complex: The ribosome fails to detach from the mRNA because the construct is designed without a stop codon. The spacer allows the nascent polypeptide to correctly fold with minimum interaction from the ribosome
Adapted from *(Li et al., 2019)*

In order to generate a stable complex, the sequence of the gene of interest is fused to a sequence that encodes a polypeptide used as a spacer (Figure 5-6). The whole sequence is designed without stop codon to prevent the release of newly synthetized polypeptide from the ribosome during translation. The spacer must be long enough to enable the complete passage of the protein of interest through the ribosome exit channel so that affinity capture can be performed without

interference between the ribosome and the protein of interest.

Initially, the sequence of the spacer encoded 31 repeated glycine-serine residues (Mattheakis, Bhatt and Dower, 1994). Since then, other studies have used fractions of the gene III of the M13 filamentous phage, or the sequence of the CH3 domain of a human IgM, or even the sequence of the constant chain of a human Igκ as a spacer (Hanes and Plückthun, 1997; He and Taussig, 2002; Zhao *et al.*, 2009).

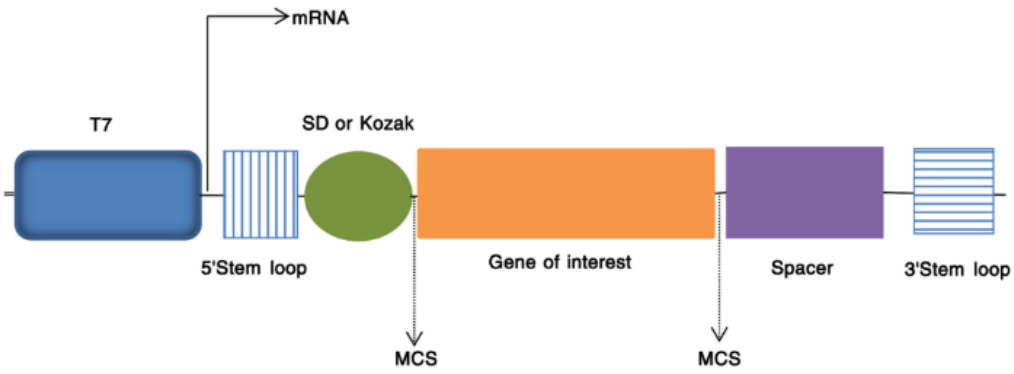

Figure 6. Standard DNA cassette for ribosome display: The displayed gene is placed under the control of a T7 promotor, then a prokaryotic (Shine-Dalgrano) or a eukaryotic (Kozak) translation initiation sequence is integrated in front of the gene and the sequence is fused to a spacer without a stop codon. Two stem loops and two multi-cloning sites (MCS) are also integrated.
Adapted from (Li *et al.,* 2019)

Transcription and translation machinery is required for the display of the protein of interest. In order to use a completely cell-free display platform, the DNA cassette can also be integrated into a cell free translation system. These systems are derived from cell extracts and are commercialized. Different eukaryotic and prokaryotic systems are available, the most suitable system can be chosen depending on the downstream application of the displayed protein (Gregorio, Levine and Oza, 2019).
Using a cell free system makes it possible to get rid of the transformation step, which can limit the size of the displayed library.

Since 1994, ribosome display has been widely used in the development of new specific binders, but the technique has some limitations. The ternary complex between the mRNA, the ribosome and the nascent polypeptide has a low stability. Moreover, the cell-free translation system inevitably contains some impurified proteases and nucleases that can destabilise even more this complex.
To circumvent this degradation, the PURE (Protein synthesis Using Recombinant Elements) cell-free translation system can be used (Kanamori, Fujino and Ueda, 2014). This cell-free translation system was produced with recombinant tagged proteins in order to minimise contaminants such as nucleases or proteases (Shimizu *et al.*, 2001).

1.1.1.3 mRNA display

The central feature of mRNA display is the creation of a covalently bond between the nascent polypeptide and the 3' extremity of the mRNA that encodes it.
To do so, the 3' extremity of the messenger RNA must be ligated to a DNA template fused with puromycin (Figure 7).

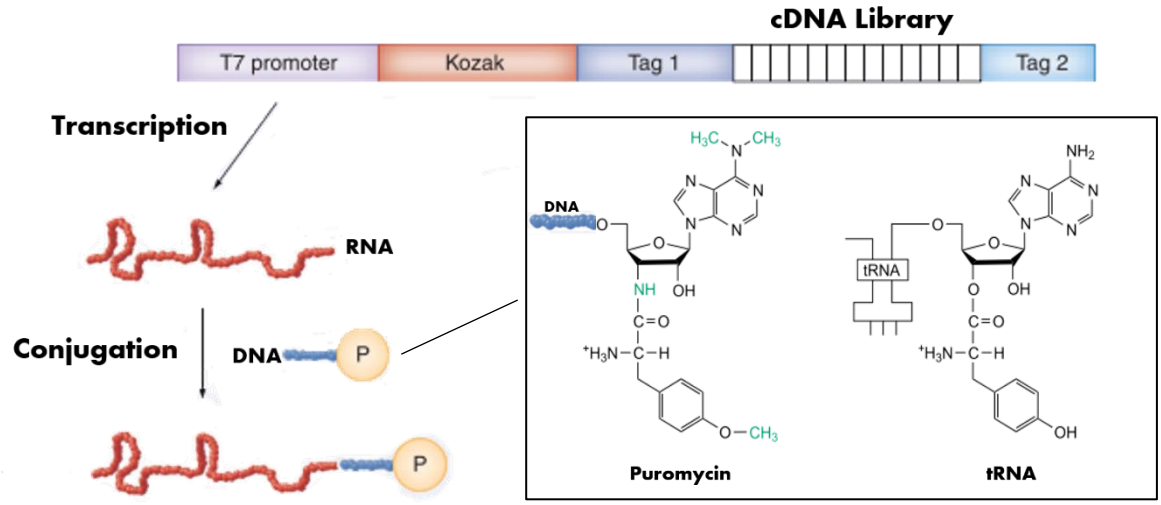

Figure 7. General steps for mRNA display: The standard DNA cassette contains a T7 promotor and a Kozak or a Shine-Dalgrano translation initiation sequence. The 3' extremity of the transcribed RNA is conjugated with a DNA template fused to puromycin. Puromycin is a small molecule analog to the tyrosyl-tRNA; the differences are highlighted in green. Adapted from (Takahashi, Austin and Roberts, 2003; Wang and Liu, 2011)

Puromycin is an antibiotic discovered in *Streptomyces alboniger*. Puromycin acts as an analogue to the tyrosyl-tRNA; it can hijack the aminoacyl-ARNt synthetase to incorporate itself into a growing polypeptide causing the termination of its translation (*Puromycin | C22H29N7O5 - PubChem*).
During the translation of the display cassette, the ribosome pauses at the RNA-DNA junction, and puromycin can enter the ribosome A site where it can form a peptide bond with the nascent polypeptide (Figure 8).

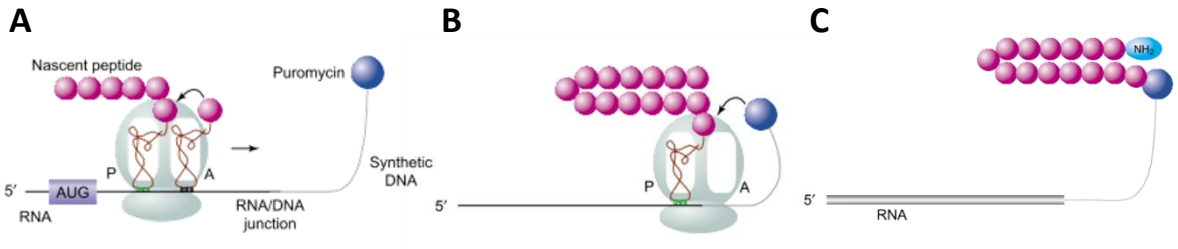

Figure 8. Formation of the protein-mRNA fusion: (A) The ribosome is recruited by the translation initiation sequence and starts the translation of the peptide after the AUG codon. **(B)** At the RNA-DNA junction, the ribosome pauses and puromycin can enter in the site A of the ribosome **(C)** The nascent peptide-puromycin peptidyl bond triggers the end of the translation and the release of the mRNA-polypeptide displayed fusion.
Adapted from (Takahashi, Austin and Roberts, 2003)

As for the ribosome display, mRNA display can be performed entirely *in vitro* by using a cell-free translation system to translate the mRNA-template DNA-puromycin construction.
The PURE cell free translation system can also be used to increase the stability of the mRNA-protein fusion (Josephson, Ricardo and Szostak, 2014) .
The mRNA can also be reversely transcribed into an mRNA-DNA hybrid to increase the stability of the complex. The mRNA-DNA hybrid can minimise the interactions of the mRNA with the displayed polypeptide. These interactions can affect the protein properties and produce biased result (Wang and Liu, 2011).

1.1.2 Cellular display

Supramolecular display platforms (*in vitro* platforms) are powerful because they enable selection within large libraries (up to 10^{10} variants) but they are not adapted to the displaying of complex proteins. As a matter of fact, they cannot perform post-translational modification that are sometimes required for the function of the protein. Moreover, *in vitro* platforms are not adapted to large proteins that require helper proteins to correctly fold and are not adapted for the directed evolution of enzymes since capture-based selection is based on affinity but not catalysis.
Therefore, cellular display technologies were developed to overcome some of these limitations.
The most popular cellular display platform is the yeast surface display, but some bacterial display platforms have also been developed.
In this format, the displayed library is screened by incubating the cell with a fluorescently labelled-target molecule that can interact with the displayed protein. This allows the cell to be differentially labelled depending on the displayed protein properties. The cell can then be selectively extracted from a library using Fluorescence Activated Cell Sorting (FACS). Moreover, unlike phage or ribosome display, a high number of proteins can be expressed at a cell surface therefore increasing the detectability. In particular, cells displaying enzyme variants can be compartmentalised in microdroplets and screened for activity using a microfluidic sorting device (called FADS for Fluorescence Activated Droplet Sorting).

1.1.2.1 Yeast display

Multiple different yeast-based display platforms have been developed since its first proposal in 1997 (Boder and Wittrup, 1997), using different strains or different surface anchor protein fusions (Kondo and Ueda, 2004; Cherf and Cochran, 2015). However, the *Saccharomyces cerevisiae* Aga1-Aga2 display system remains the most popular.
This system uses the mating type specific a-agglutinin complex to display the protein of interest. This complex is specifically expressed by yeast with the *a* mating type and can specifically bind with high affinity α-agglutinin glycoprotein encoded by yeast with the *α* mating type (Zhao *et al.*, 2001). This high affinity allows the irreversible binding of cells of opposite types during their mating.
a-agglutinin is made of two subunits linked through two disulfide bridges: Aga1, a 725 amino acid protein that is anchored to the cell wall through a glycosylphosphatidylinositol (GPI) bond, and a smaller subunit, Aga2, that confers the binding properties to a-agglutinin (Figure 9). Both glycoproteins contain a secretion signal within their sequence (*a-agglutinin | SGD*).

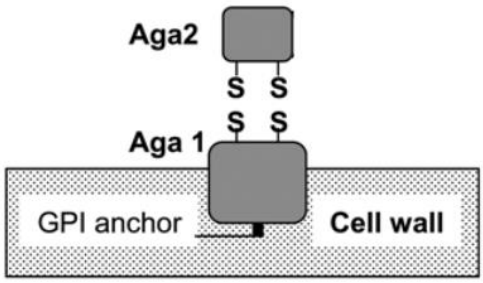

Figure 9. Structure of the a-agglutinin complex: a-agglutinin is made of two subunits, Aga1 and Aga2, linked through two disulfide bridges
Adapted from (Kondo and Ueda, 2004)

This yeast surface display system uses this naturally anchored complex to display the protein of interest. To do so, the N-terminus or the C-terminus of the protein of interest is fused to Aga2. Usually, the terminus farthest from the functional structure of the protein is fused to Aga2 in order to minimise the interaction of the display platform on the properties of the protein of interest.
The displayed protein is also usually flanked with two tags that are recognised by fluorescent antibodies (Figure 10). Those fluorescent antibodies allow to quantify the full-length expression of the displayed protein by flow cytometry (Linciano *et al.*, 2019). This is used to normalise the function of the protein of interest that is measured with fluorescent probes.

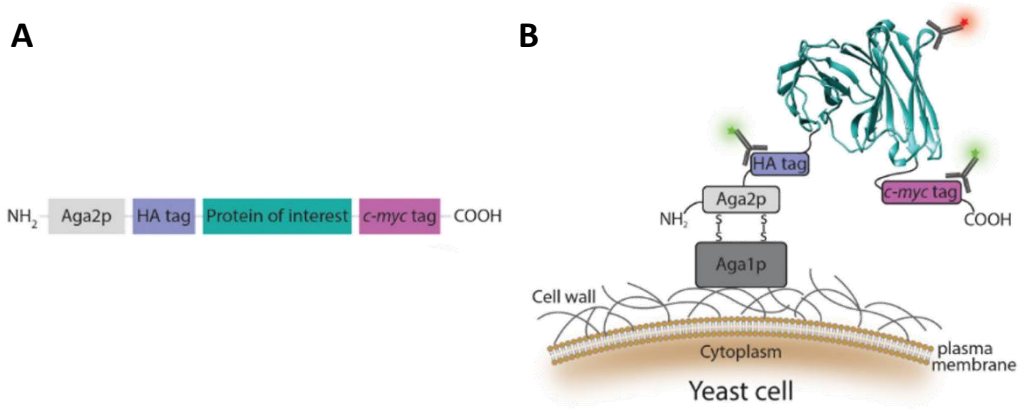

Figure 10. Schematic representation of yeast surface display: (A) Structure of the Aga2 protein of interest fusion **(B)** Yeast display schematic structure
Adapted from (Cherf and Cochran, 2015)

Yeast display is a powerful tool for protein engineering because it allows eukaryotic glycosylation of secreted proteins. Yeast display combined with FACS also permits more control over the function of the protein of interest (Sheehan and Marasco, 2015).
Yet, the technology has some limitations. First the size of the library is limited by the transformation efficiency in yeast. Yeast library has a theoretical limit of 10^7 variants.
Also yeast display permits the expression of 10^4-10^5 heterologous proteins on its surface. With polyvalent targets, the selection by flow cytometry can be driven by avidity for the fluorescent probe rather than its specific affinity.

1.1.2.2 Bacterial display

a) Gram-negative: Escherichia coli surface display

Gram-negative bacteria have a complex envelope cell structure composed of an inner membrane, a periplasmic space and then an outer membrane. In order to develop a display platform on gram-negative bacteria, the protein of interest needs to pass through the inner membrane, the periplasm and the outer membrane in order to be targeted to the external surface.

The most frequently used gram-negative host is *Escherichia coli*. Firstly, because genetic manipulations in *E. coli* are well-known but also because of its rapid growing rate and its efficient transformation rate allowing the display of large libraries.

As shown in Figure 11, a lot of different display strategies have been developed over the years. One of them is to fuse the displayed protein to a carrier protein that is displayed on the outer membrane in *Escherichia coli* or in other gram-negative bacteria (Wentzel *et al.*, 2001). Carrier proteins can also be transmembrane proteins naturally targeted in the outer membrane of *E. coli*. Chimeric fusions such as the Lpp-OmpA display fusion have also been used as a carrier protein (Francisco, Earhart and Georgiou, 1992).
Other display strategies such as fusion of the displayed protein with the flagella or even the pilus of *E. coli* has also been developed as well as plenty of other strategies reviewed elsewhere (Figure 11) (Lee, Choi and Xu, 2003; Daugherty, 2007; van Bloois *et al.*, 2011)

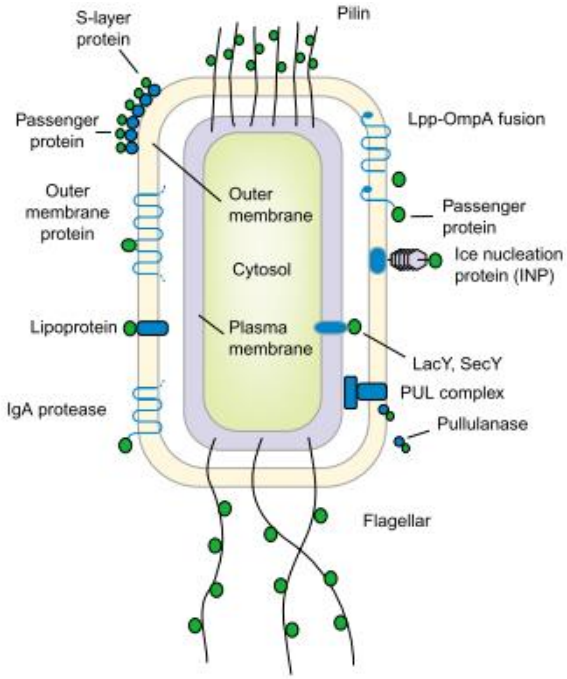

Figure 11. Schematic representation of E. coli display strategies: A lot of different display strategies have been developed on *E. coli*: 1.Fusion of the displayed protein to a carrier protein naturally targeted to the outer membrane or engineered to be so 2.Fusion to external cell structures, such as the pilus or the flagella.
Adapted from (Lee, Choi and Xu, 2003)

So far, *E. coli* displays have been the most commonly used bacterial hosts but the development of new display platforms with gram-positive bacteria can bring other attractive advantages (Löfblom, 2011).

b) Gram-positive: staphylococcal surface display

In 1995, Samuelson *et al* designed a vector for displaying recombinant proteins on the surface of *Staphylococcus carnosus*. This strain was chosen because the strain was well-known and used for over forty years by the food industry. *Staphylococcus carnosus* also had other advantages for the development of a display platform.
Firstly, the bacterium appears to mostly grow as singles or sometimes as pairs. This is an important feature for the selection by flow cytometry. Otherwise, the fluorescence measured by FACS would require another fluorophore expressed constitutively to normalize the number of cells per signal. Moreover, *Staphylococcus carnosus* is classified as generally recognized as safe (GRAS) by the FDA because it does not express any of the virulence factor-associated proteins in pathogenic staphylococci (Gunneriusson *et al.*, 1996).

The vector for *Staphylococcus carnosus* display contains a promotor and a secretion signal from a lipase, a protein naturally displayed on the surface of *Staphylococcus hyicus*. Scientist also added a propeptide after the secretion signal that was shown to be important for the translocation of the lipase. The propeptide was fused to a multicloning site and an albumin binding protein (ABP) that is used as a spacer between the cell wall of *Staphylococcus carnosus* and the displayed protein. Finally, the construction contained an anchoring signal recognized by the sortases of *Staphylococcus carnosus* (see section 1.2.2.2).
They also added a staphylococcal origin of replication and a chloramphenicol acetyl transferase gene for proper replication and expression of the vector in staphylococci.
In order to facilitate the manipulation of the vector, an origin of replication for *E. coli* and a β-lactamase gene was also added to the vector (Figure 12) (Kronqvist *et al.*, 2010).

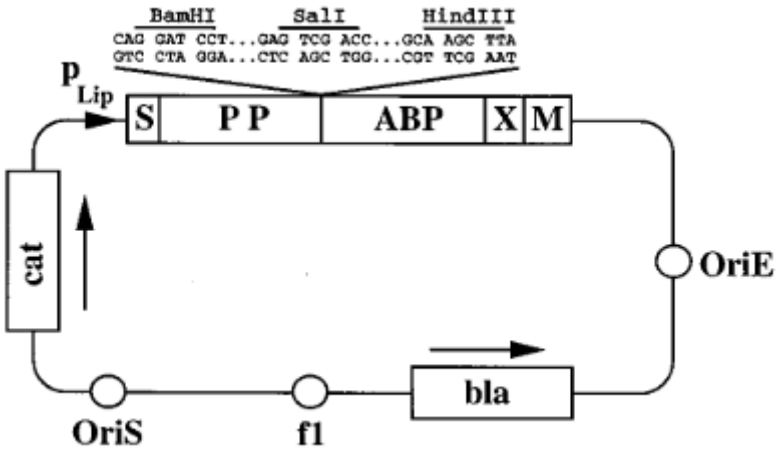

Figure 12. Plasmid for display of recombinant proteins on *Staphylococcus carnosus*: The vector contains the promotor (P_{Lip}), the secretion signal (S) and the propeptide (PP) of a Lipase fused to an albumin binding protein (ABP) and an anchoring signal with the LPXTG (X and M) motif. A staphylococcal origin of replication (OriS) and a chloramphenicol acetyl transferase gene were added to the plasmid as well as a β-lactamase and an origin of replication for *E. coli*
Adapted from (Samuelson *et al.*, 1995)

Gram-positive bacteria are good hosts for the display of recombinant proteins because of their relatively simple organization. The recombinant proteins have just one plasma membrane to pass through in order to be anchored to the cell wall of the bacterium. Moreover, Gram-positive bacteria have a good tolerance to mechanic stress due to their thick peptidoglycan layer. That is an important

advantage because this tolerance reduces the mortality during flow cytometry sorting, a step that causes several physical pressures on the bacteria that can cause the disruption of the membrane.

1.2 Protein sorting

In order to display a recombinant protein to the surface of a competent cell, the DNA cassette must encode an appropriate signal peptide for the translocation of the protein. In gram-positive bacteria display, the protein also needs to be fused to an anchoring motif recognisable by a sortase that can covalently attach the protein to the cell wall.

1.2.1 Translocation of the protein across the plasma membrane

The first step for targeting the protein to the cell wall is the translocation through the cytoplasmic membrane. Two major bacterial protein export pathways are known: the ubiquitous general secretion (sec) and the twin-arginine translocation (tat) pathways. In order to be exported, proteins need a signal peptide at their N-terminus that can be recognised by either a sec-specific translocase or a tat-specific translocase.

1.2.1.1 Sec pathway

The general secretion (sec) pathway is the route most widely used sorting by the exported protein (Tsirigotaki *et al.*, 2017). It allows the transport of high weight unfolded protein precursors across the plasma membrane. The transmembrane SecYEG channel, also called SecYEG translocon, is the central feature of this sorting system. Exported proteins are targeted to the translocon through an amino terminal signal peptide. Depending on the hydrophobicity of the signal peptide, the protein can be translocated in a co-translational or in a post-translational manner (Denks *et al.*, 2014).

If the signal peptide is highly hydrophobic, it can recruit a ribonucleoprotein called signal recognition particle (SRP) right after being translated. Since the signal peptide is located in the N-terminus of the protein, the SRP is recruited before the complete translation of the protein. Upon SRP binding, the translation is stalled, and the ribosome nascent chain complex is targeted to the FtsY membrane receptor. Finally, this receptor allows the transfer of the ribosome nascent chain complex to the SecYEG translocon and translation is resumed. The rest of the protein is directly translated into the translocon pore and then excreted (Figure 13A). In this pathway, the polypeptide never accumulates in the cytosol.

Sec post-translational secretion is a pathway followed by proteins that have a less hydrophobic signal peptide. Those proteins are not recognised by SRP and are entirely translated into the cytosol. There, those unfolded proteins can interact with post-translationally interacting proteins (PiP's), such as folding chaperones or post-translational modification machinery. The signal peptide is then recognised by the protein SecA, which can prevent the aggregation of the proteins but also maintain them unfolded. The complex is then targeted to the SecYEG translocon, where the ATPase subunit of SecA can push the protein through the pore of the translocon. The SecYEG translocon complex is also associated with SecD and SecF, which can also exert a pulling force on the excreted polypeptide (Figure 13B). In this pathway, besides its key role in the targeting of the protein to the translocon, the

signal peptide also influences the properties of the protein such as folding kinetics, biosynthesis and stability (Rusch and Kendall, 2007).

In both pathways, the signal peptide is cleaved off after the translocation by a signal peptidase.

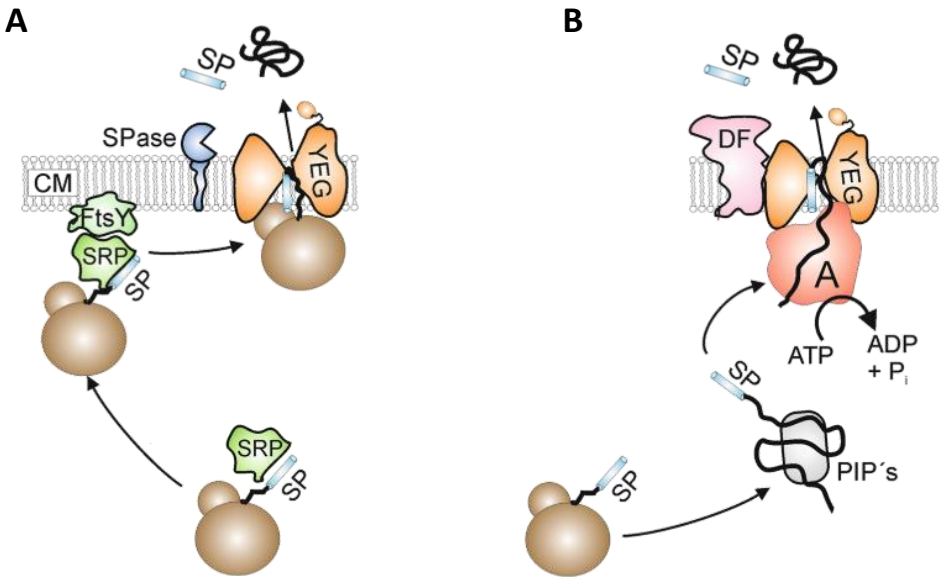

Figure 13. Schematic representation of Sec pathway: (A) Co-translational pathway **(B)** Post-translational pathway
Adapted from (Freudl, 2018)

Bioinformatic tools have been developed in order to predict if a protein contains a Sec signal peptide (Phobius http://phobius.sbc.su.se/ ; Signal IP http://www.cbs.dtu.dk/services/SignalP/). Those tools can predict the size of the signal peptide and also its cleaving site.

It is difficult to predict which signal peptide has the best properties for a given recombinant protein. Proteins whose folding cannot be prevented by the signal peptide such as many small cytosolic proteins may only be secreted via the co-translational pathway. However, higher fitness cost is associated with this pathway compared to the post-translational pathway (unpublished observation). To develop the best system, scientists should therefore screen for several different signal peptides.

1.2.1.2 Twin-arginine translocation pathway

The twin-arginine translocation (Tat) pathway is another secretion route that is present in almost all bacteria. It has the particularities to secrete fully folded proteins or even proteins bound to their cofactor (Freudl, 2018). This pathway allows the secretion of a few proteins that feature a conserved twin-arginine motif in their signal peptide. The translocation complex of this pathway is the multimeric TatABC complex.

The transport cycle begins with the binding of the signal peptide to the TatBC complex. The TatC protein recognises the twin-arginine motif of the signal peptide and confers the specificity to the secretion route. The formation of the TatBC-substrate complex triggers the recruitment of TatA that oligomerises and creates a pore through the membrane in an energy-dependent manner. The Tat system is energized by transmembrane proton motive force (PMF) (Palmer and Berks, 2012).

The substrate can then be excreted through the pore. Finally, the signal peptide is cleaved off by a signal peptidase associated to the membrane (Figure14).

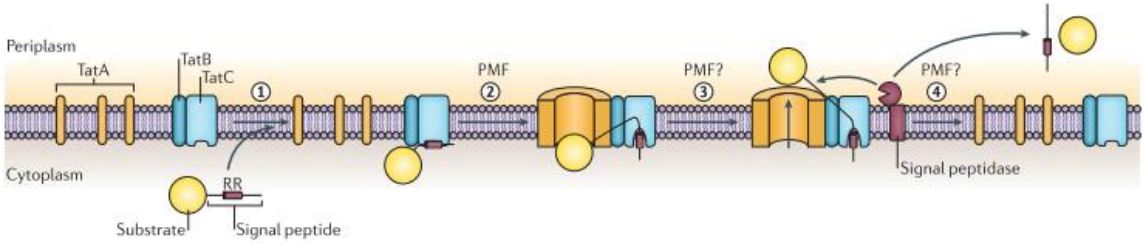

Figure 14. Schematic representation of Tat-twin pathway: The signal peptide of the substrate binds itself to the TatBC complex. The TatBC-substrate complex triggers the polymerization of the TatA subunits and creates a pore through the plasma membrane. The substrate is then exported, and the signal peptide is cleaved off by the signal peptidase.
Adapted from (Palmer and Berks, 2012)

As for the sec pathway, the Tat-specific signal peptide can be predicted using bioinformatic tools (Tatfind server http://signalfind.org/tatfind.html) (Natale, Brüser and Driessen, 2008).

1.2.2 Anchoring of the protein to the cell wall

Once the recombinant protein is exported to the periplasmic space, the second step to build a display platform is the formation of a covalent bond between the protein and the cell wall.
The bacterial cell wall is predominantly composed of a peptidoglycan backbone.
Some proteins are naturally anchored to the cell wall of Gram-positive bacteria by a family of enzymes called sortases. Those proteins carry a specific sorting motif in their sequence in order to be recognised by the sortase.

1.2.2.1 The bacterial cell wall

The bacterial cell wall is a rigid exoskeletal structure with several essential functions: it creates a physical barrier protecting the bacteria from the hostile environment, it can also protect the integrity of the bacteria preventing membrane rupture upon physical or osmotic stress. In gram-positive bacteria, the cell wall also plays another essential role: it can serve as a scaffold to display exported proteins in the external surface of the bacterium. Those proteins can promote bacterial adhesion, nutrient acquisition or even suppression of the host immune response.

The major structure of the bacterial cell wall is the peptidoglycan. Although its composition can vary between species, several structures remain conserved among all bacteria. First the glycan chains are composed of a repetition of the disaccharide N-acetylglucosamine (NAG) linked to N-acetylmuramic acid (NAM) by a β1-4 bond. Those glycan chains are crosslinked by peptides giving a three-dimensional matrix (Figure 15). The composition of the peptide bridge can vary from species to species.

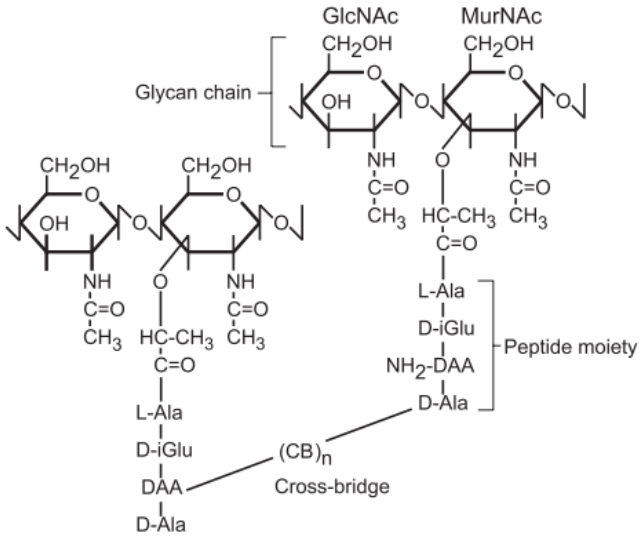

Figure 15. Structure of the gram-positive bacterial peptidoglycan: The glycan chains are cross-linked by a peptidic structure giving a three-dimensional matrix. This structure is composed of two peptide moieties connected by cross-bridges $(CB)_n$. The peptidic moieties are composed of L-alanine (L-Ala), D-isoglutamete (D-iGlu), a diamino acid (DAA) such as L-lysine or m-diaminopimelic, and D-alanine (D-Ala). The composition of the cross-bridge may vary between species. $(CB)_n$ = pentaglycine $(Glu)_5$ for *Staphylococcus aureus*; $(CB)_n$ = nothing for *E. coli* where the DAA is directly linked to the D-Ala
Adapted from (Ton-That, Marraffini and Schneewind, 2004)

Peptidoglycan synthesis begins in the cytoplasm. The peptidoglycan precursors are incorporated into the plasma membrane, where the cross-bridge between the glycan chains are catalysed by penicillin binding proteins. Finally, the mature peptidoglycan is released and creates the rigid cell-wall (Vollmer, Blanot and de Pedro, 2007; Rohde, 2019).

1.2.2.2 Sortase and sorting signal for cell wall protein anchoring

Sortase enzymes are used by bacteria to anchor proteins to the cell wall and to assemble pili. Those enzymes specifically target proteins that carry an LPXTG motif in the C-terminal region. This motif is followed by a hydrophobic domain and a positively charged tail that retains the protein within the plasma membrane during the translocation and before the sortase-mediated anchoring.

The most characterised sortase is the sortase A from *Staphylococcus aureus* (SaSrtA) (Mazmanian *et al.*, 1999). The structure function of this protein has been studied in details and SaSrtA serves as a model for the family.

The active catalytic site of the SaSrtA is composed of a cysteine positioned close to an arginine and a histidine (*RCSB PDB - 1T2P: Crystal structure of Sortase A from Staphylococcus aureus*; Zong *et al.*, 2004). The LPXTG motif is positioned near the cysteine whose nucleophilicity is activated by the arginine. The cysteine attacks the peptide bond between the threonine and the glycine. This generates a tetrahedral intermediate which can resolve itself by cutting the link between the two amino acids. A substrate-sortase complex is then created through a thioester bond. SaSrtA will then recognise a peptidoglycan precursor integrated to the plasma membrane and catalyse the nucleophilic attack of the primary amine group of the free pentaglycine linker on the thioester. Another tetrahedral intermediate is created and can be released by cutting the bond between the

LPXT substrate and the sortase, releasing the substrate covalently linked to a glycan chain that can be integrated to the cell wall (Figure 16) (Zong *et al.*, 2004).

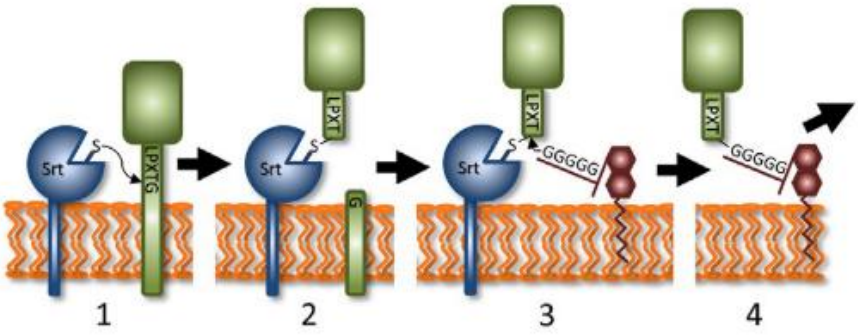

Figure 16. Mechanism of cell wall protein anchoring: (1) The active cysteine attacks the peptide bond between Thr and Gly. **(2)** Resolving of the tetrahedral intermediate by the cleavage of the peptide bond. **(3)** The peptidoglycan precursor (pentaglycine) attacks the acyl-enzyme. **(4)** Resolving of the tetrahedral intermediate by the cleavage of the thioesther bond and release of the protein covalently linked to the peptidoglycan precursor. Adapted from *(*Jacobitz *et al.,* 2017*)*

SaSrtA homologs have been identified in all known gram-positive bacteria, and most of the time more than one sortase per bacteria strain have been characterized. Among all of those sortases, the cysteine, the arginine and the histidine of the active catalytic site are conserved, and the folding of the catalytic site remains similar. An LXTC motif has been reported on each catalytic domain containing the active cysteine (Ton-That, Marraffini and Schneewind, 2004).
Variation of the structure of other domains of the sortases have been reported. In fact, the sortase family has been classified in five different classes (A-E) depending on these structural differences. Small modifications on the sortase structure allows to recognise a slightly different anchoring motif (Jacobitz *et al.*, 2017).

1.3 *Streptococcus thermophilus*

Streptococcus thermophilus is widely used in the dairy industry for its high acidification ability. This gram-positive anaerobic-facultative bacterium belongs to the streptococcus genus including 28 other species divided into seven different groups (Shanker *et al.*, 2016). Several of these streptococcus species carry human pathogenic factors but *Streptococcus thermophilus* has a GRAS status. As a matter of fact, studies of several strains of *Streptococcus thermophilus* demonstrate that most of the streptococcal pathogenic factors are absent from the genome or in a pseudogene form (Bolotin *et al.*, 2004).

Streptococcus thermophilus could be a good host for the display of recombinant proteins. Indeed, the study of its genome revealed that *Streptococcus thermophilus* only has one housekeeping protease, htrA, able to degrade exported proteins. This gene could be deleted to decrease the degradation of the displayed protein.
Moreover, the analysis of the *Streptococcus thermophilus* genome predicts only three sortase-dependent proteins with the LPXTG motif. Those three proteins are: PrtS, a cell envelope serine proteinase, Mub, a mucus-binding protein and CpdB, a cyclo-nucleotide phosphodiesterase (Goh *et al.*, 2011).

Another key advantage of *Streptococcus thermophilus* as a host for a display platform is its highly efficient natural competence and the mechanisms that allow the introduction of recombinant DNA into its chromosome by homologous recombination. The activation of the natural competence can be performed in the lab facilitating the construction of genetic libraries. Furthermore, the competence efficiency of some strains such as LMD-9 allows the construction of libraries with much higher diversities (up to 10^9) compared to classical hosts, such as *Saccharomyces cerevisiae* or *Escherichia coli* (Fontaine *et al.*, 2010).

1.3.1 Natural competence

Natural competence is a trait well conserved among gram-positive and gram-negative bacteria. It allows the integration of exogenous DNA which then could be integrated into the genome of the bacterium. This trait plays an important role in the plasticity of the bacterium, allowing it, for example, to develop escape mechanisms from the immune system of its host in case of pathogenic strains. It can also help the spread of antibiotics resistance genes.

The integration of exogenous DNA is a highly energetic process. Bacteria therefore need a regulatory mechanism that can upregulate or downregulate the transcription of the genes involved in this process, depending on environmental conditions.

In streptococci, a competence regulation mechanism called ComCDE was characterized in 1995 (Håvarstein, Coomaraswamy and Morrison, 1995). This system uses a protein ComE to activate the transcription of *comX* gene. ComX is a transcription factor that upregulates the transcription of the genes required for activating natural competence. The ComE protein is activated when a competence stimulating peptide (CSP) derived from an excreted protein, called ComC, binds a transmembrane histidine kinase. In response ComC auto-phosphorylates and then transfers its phosphoryl group to the cytosolic protein ComE, making it able to upregulate the transcription of ComX gene (Figure 17A).

Other Streptococci strains including *Streptococcus thermophilus* lack ComCDE orthologs in their genomes. Instead, these bacteria use the ComRS system that was discovered in *Streptococcus thermophilus* in 2009 while its general mechanism was characterised in 2010 (Gardan *et al.*, 2009; Fontaine *et al.*, 2010).
The ComRS system also activates the transcription of the ComX gene. The system is activated when XIP (ComX Inducing Peptide), a peptide derived from the excreted ComS protein, is reintegrated into the bacteria through the Ami/Opp transporter. In the cytosol, XIP forms a complex with the cytosolic ComR protein. This complex serves as a transcription factor for the competence-activating *comX* gene (Figure 17B) (Fontaine *et al.*, 2013).

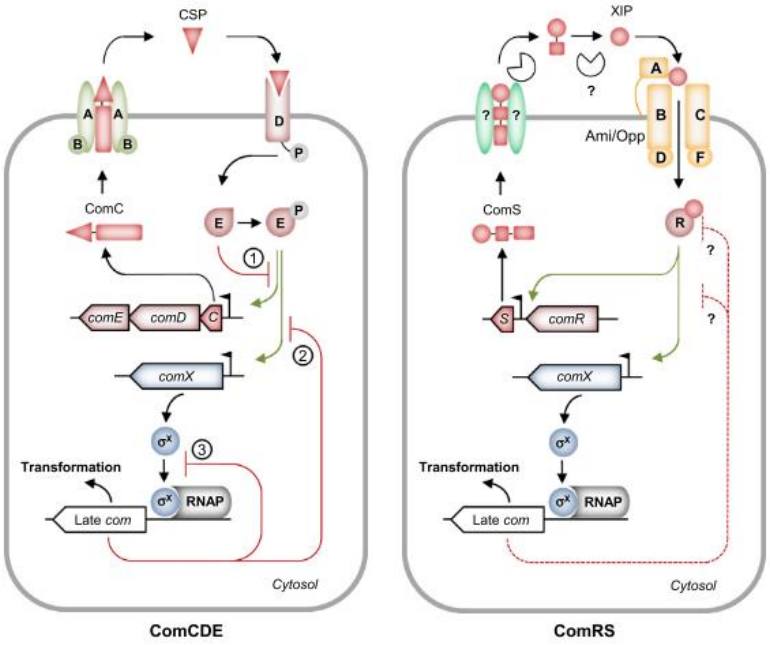

Figure 17. Mechanism of competence activation: (A) ComCDE system **(B)** ComRS system
(Fontaine *et al.*, 2015)

Once the exogenous DNA is integrated into *Streptococcus thermophilus*, it can easily integrate its genome by double-cross homologous recombination, provided that the DNA sequence has homology with the genome. For LMD-9 strain, the efficiency of DNA capture and integration in the chromosome is extremely high with up to 10% of total cells that can be transformed upon simple incubation of bacteria with XIP and linear DNA.

1.3.2 Peptidoglycan anchoring

As mentioned before, the peptidoglycan is an important structure in gram-positive bacteria, such as *Streptococcus thermophilus*. The peptidoglycan is composed of glycan strands made of repeated successions of the disaccharide N-acetylglucosamine (NAG) linked to N-acetylmuramic acid (NAM) by a β1-4 bond. The glycan strands are linked together through peptide bridges in order to give a three-dimensional matrix. Those peptide bridges are composed successively of a L-alanine, D-glutamate, L-lysine and two D-alanine. In *Streptococcus thermophilus*, the peptide bridges are cross-linked with two or three L-alanine between the L-lysine and the first D-alanine (Heinz and Kandler, 1972; Hols *et al.*, 2005).

Several *S. thermophilus* strains have been the subject of comparative genomic studies to characterize their specific features but also to highlight conserved mechanisms with the research of homologous genes. In the field of protein anchoring, only one complete sortase A homolog has been found in the LMD-9 strain. A non-functional homologous pseudogene has also been highlighted in two other tested strains (LMG18311 and CNRZ1066) (Goh *et al.*, 2011).
Further analysis in the LMD-9 genome highlights three proteins containing the LPXTG motif: Prts, Mub and CpdB (Fernandez-Espla *et al.*, 2000; Hols *et al.*, 2005).

2. Objectives and strategies

The objective of this master thesis is to develop and evaluate a functional display platform with *Streptococcus thermophilus* as a host.

Streptococcus thermophilus has several properties that could be used in order to build an advantageous display platform for directed evolution.
Firstly, the natural competence of *S. thermophilus* allows the construction of large genetic libraries (up to 10^9 variants) with moderate efforts.
Secondly, *S. thermophilus* is a gram-positive bacterium with all the advantages discussed earlier such as the non-complex cell envelope structure and the high resistance to physical pressure allowing high viability in flow cytometry experiments.
However, a disadvantage of *Streptococcus thermophilus* is that it grows as chains of cells. This property is problematic in FACS since the signal intensity will vary with the size of the chain passing in the laser beam.
Therefore, another objective of this master thesis will be to overcome this drawback.
To that end, a DNA cassette containing the sequence of a green fluorescent protein (GFP) optimised for expression in *S. thermophilus* and a constitutive p32 promotor will be inserted into a permissive locus of the genome. In flow cytometry experiments, this mutant should allow the normalisation of the fluorescent signal per cell since its intensity should be proportional to the number of cells per chain.
In addition, we also plan to evaluate how sonication may help in reducing the size of the cell chains.

Our display platform will use the LMD-9 strain because of its highly efficient natural competence and its functional sortase A homolog allowing the anchoring of proteins with the LPXTG motif.

In order to display recombinant proteins on the surface of *S. thermophilus*, the chosen strategy is to use the signal peptide and the anchoring motif of a naturally anchored protein.
The protein PrtS was chosen because its function is well characterised and does not seem to be essential for the viability.
The sequence of PrtS will be replaced by the sequence of the recombinant protein, while the signal peptide and the anchoring motif of PrtS are kept and fused in frame respectively with the N-terminus and C-terminus of the recombinant protein.

As proofs-of-concept, two different constructs with the same protein of interest (POI) will be tested (Figure 18). The α-amylase from *Bacillus licheniformis,* that we called AmyL*,* was chosen as POI because it is a monomeric and naturally secreted enzyme from a gram-positive bacterium and its activity can be easily monitored.
The first construct is composed of AmyL fused to the signal peptide and the anchoring motif of the PrtS gene. Respectively, 8 and 5 residues of PrtS will be left after the signal peptide and before the anchoring motif in order to keep a similar chemical background for proper secretion and sorting.

The second construct is the same than the first one without the end of the PrtS gene and its anchoring motif. It will be used as a negative control since we expected AmyL to be excreted in the culture media.

A secondary objective is to develop an alternative display platform affording to release the α-amylase with a simple treatment of the cells with a reducing agent.

To this end, a third construct containing the sequence of the α-amylase from *Bacillus licheniformis,* fused to the signal peptide and the anchoring motif of the PrtS gene will be developed. However, a bacterial extracellular intein-like module (BIL) will be inserted between the end of the alpha-amylase and the anchoring motif. Unpublished results of the laboratory revealed that BIL can auto-cleave itself from the peptide and leave a disulfide bond between both flanking polypeptides. By combining BIL with sortase anchoring, we expect AmyL to end up linked to the peptidoglycan through a disulfide bridge.

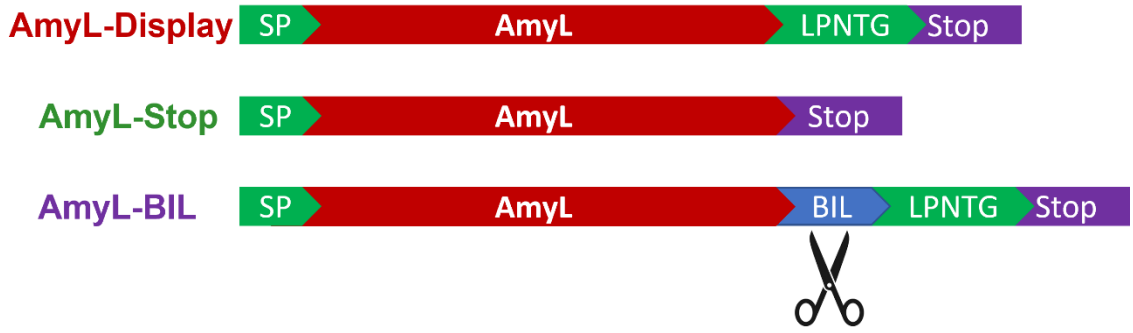

Figure 18. Schematic representation of the three constructs: The AmyL-Display construct will be used to make the display mutant. The sequence of the α-amylase is fused in frame with the signal peptide (SP) in N-terminus and the anchoring motif (LPNTG) in C-terminus from PrtS. Respectively , 8 and 5 residues of PrtS will be left after the signal peptide and before the anchoring motif. The AmyL-Stop construct is identical to the first one without the anchoring motif. This construct will be used to develop a mutant that excretes the α-amylase. It will be used as a negative control of the display. The final construct called AmyL-BIL contains the α-amylase fused in frame with the SP and the MA but with a small module called BIL inserted between the sequence of the α-amylase and the anchoring motif. This construct will be used to develop a controllable display platform

3. Results and discussion

3.1 Constitutive expression of GFP in *S. thermophilus*

One of the objectives of this master thesis was to overcome the difficulties encountered by the cell chains of *Streptococcus thermophilus*. In fact, this mode of growth can cause problems in the selection of mutant by FACS.
To this aim, our first strategy was to create a *Streptococcus* mutant that constitutively expresses GFP.

Our hypothesis was that the intrinsic fluorescence of this mutant could be used to normalise a signal that would be proportional to the number of cells per chain in flow cytometry experiments.

Since *Streptococcus thermophilus* grows in anaerobic conditions, the GFP maturation step could be disrupted. We should check for GFP fluorescent signal and toxicity of the GFP expression to the cell.

3.1.1 Transformation of *S. thermophilus* with p32-GFP cassette

Using the natural competence of *Streptococcus thermophilus,* we wanted to insert a DNA cassette of 1024 base pairs (bp) into the serine tRNA (tRNAser) locus of the chromosome. This specific locus was chosen because it is considered as permissive for insertion and has already been used in the laboratory for protein expression with success.
The cassette is containing a p32 promotor, a GFP and a modified T7 terminator. This cassette was optimized for the constitutive expression of GFP in *Streptococcus salivarius*, another strain of streptococcus. The insertion of this cassette should generate mutants that express constitutively the GFP at a level that should be detected by flow cytometry experiments.
Since we wanted to develop a GFP mutant that behave as close as possible to the WT, the inserted cassette is not carrying selection markers. As a matter of fact, the antibiotic selection pressure applied on the mutant has deleterious effects even on the bacterium that carry the resistance gene. Moreover, the production of the protein that confers the resistance can also alter the growth of the bacterium.
In order to select exclusively for mutants that have integrated the DNA cassette in their genome, we used a secondary cassette called CAT-oroP. This secondary cassette is used to develop an intermediate mutant. It contains the genes encoding *chloramphenicol acetyltransferase* (*cat*) and an orotate transporter (*oroP),* both with their promotor. The *cat* gene is a resistance gene to chloramphenicol used as a positive control of the integration of the CAT-oroP cassette. Only mutants that have integrated the cassette are able to grow in media supplemented with chloramphenicol.
The *oroP* gene encodes for a transmembrane transporter able to transport orotate but also 5-fluoroorotic acid (5'FOA), a toxic molecule to the *S. thermophilus* cells. The gene is used as a negative selection marker that is used in the second step when the CAT-oroP cassette is substituted with the cassette of interest because only mutants that don't express oroP are able to grow on media supplemented with 5'FOA.
The general process for the integration of a DNA cassette with the CAT-oroP selection system is shown in Figure 19.

By activating the natural competence of *S. thermophilus* (see Material and methods), we allows the

internalisation of the CAT-oroP single stranded DNA cassette. The cassette must be flanked by homology regions (min 1000 bp) allowing homologous recombination at the target locus.

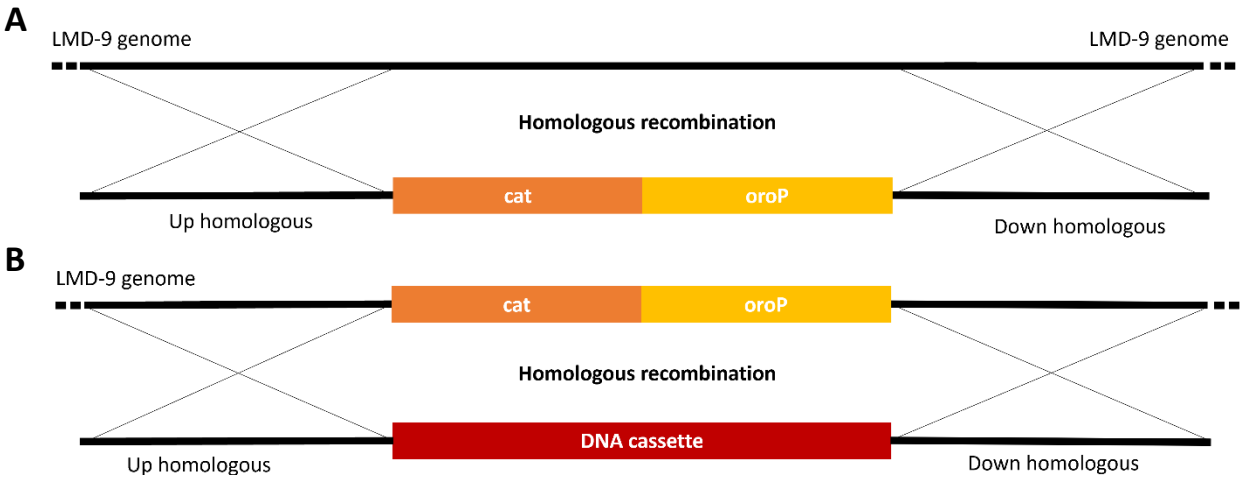

Figure 19. Schematic representation of the integration of a markerless DNA cassette in the chromosome of *S. thermophilus* using the CAT-oroP selection system: (A) By activating the natural competence of *S. thermophilus*, we can integrate a single stranded DNA fragment containing the CAT-oroP cassette flanked by homology regions. The DNA cassette is integrated in the targeted locus by homologous recombination and the intermediate mutants are selected with media supplemented with chloramphenicol. **(B)** Using the same protocol, we integrate the cassette of interest flanked by the same homology regions. The CAT-oroP cassette is substituted with the cassette of interest by homologous recombination and only the mutants substituted are able to grow on media supplemented with 5'FOA.

Since we used a locus already used for POIs sequence integration, the intermediate mutant containing the CAT-oroP cassette was already available.
We started from this intermediate mutant to develop our GFP mutant. As displayed in Figure 19B, we needed to construct a linear DNA fragment containing the p32-GFP cassette flanked by homology regions from the tRNAser locus of *S. thermophilus.*
To this aim, we amplified the up homologous fragment, the p32-GFP cassette and the down homologous fragment. The up and down homologous fragments were both amplified with one primer containing a floating tail of respectively 33 and 32 bp in order to create homology regions with the p32-GFP cassette for fusion by Gibson assembly (see Material and methods). The size of the three fragments was verified with agarose gel electrophoresis (Figure 20A).

Once the fragment were produced and cleaned up using a commercial kit (see Material and methods), we performed the assembly of the three fragments by Gibson reaction. Two different DNA concentrations were tested: a high concentration (280ng in 20µl) and a low concentration (90ng in 20µl). The finished reaction was cleaned up and eluted in 15µl. 5µl of each reaction was kept for the gel in Figure 20B and 1µl was introduced to a PCR mix with primer for the amplification of the whole cassette (from up to down homologous).

In Figure 20B, we can see the result of an agarose gel electrophoresis with the two cleaned-up Gibson and the PCR reactions. Our first observation from this gel was that we almost couldn't see any band in the cleaned-up Gibson wells. Indeed, only a faint band at ~1500bp corresponding to the homologous fragments is visible for the high DNA concentration Gibson.
This observation is in fact logic since, unlike PCR, Gibson assembly is not amplifying any DNA. The technique allows the fusion of fragment with homologous ends by the simultaneous action of a DNA exonuclease and a DNA polymerase. Moreover, the Gibson mixes were cleaned up before being charged in the wells. The clean-up step, as optimised possible, is still not able to recover the total

quantity of the input DNA. It means that the total quantity of DNA charged in the wells is inferior to the quantity contained in 5µl of the Gibson mixes (<70ng for high concentration and <22.5 for low concentration). This quantity of DNA is hardly detected by ethidium bromide (BET) agarose gel revelation.
This result doesn't mean that the Gibson assembly didn't work. In fact, the wells of the PCR product of the cleaned-up Gibson show that there was indeed the assembled fragment in the Gibson mixes. Indeed, in both low and high DNA concentration we can see a faint band at approximately 4000bp that should correspond to the amplification of the whole linear DNA. While the band, and so the quantity of DNA, is less important for the PCR of the low DNA concentration, it is the purest sample because there is no smear visible. This sample was then used to continue the manipulation.

Since the PCR from the cleaned-up Gibson contains a visible contamination with a fragment of approximately 3000bp, we decided to perform a gel band extraction protocol using a commercial kit (see Material and methods). Once the DNA was retrieved from the agarose gel band, we performed a last PCR with the primer to amplify the whole assembly. In Figure 20C, we can see the final agarose gel electrophoresis with the PCR product.

The final PCR product contained a pure band at approximatively 4000bp that corresponds to the length of the whole assembly (4024bp). We can also see an unexplained band at ~500bp. This product was cleaned up and used as donor DNA. Since the contaminant visible in the gel was only 400bp, it should not integrate in the genome of *Streptococcus thermophilus*. Indeed the integration of DNA is made by homologous recombination and needs homologous arms of 1000bp minimum.

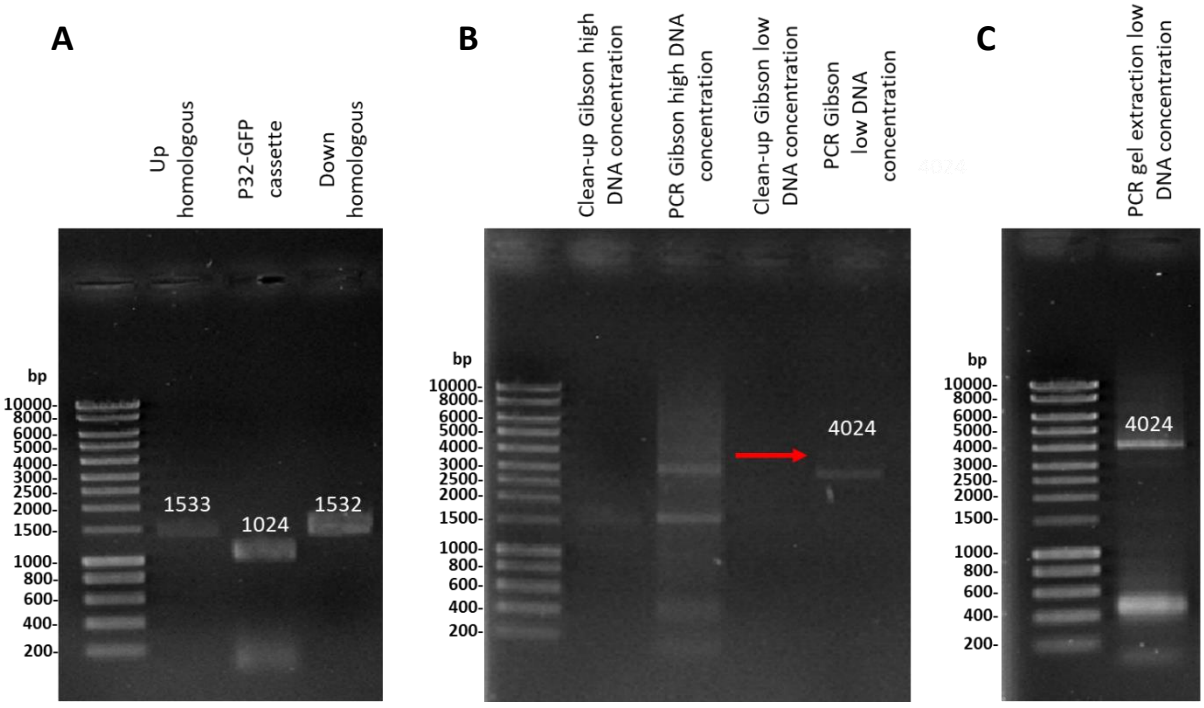

Figure 20. Agarose gel electrophoresis for the p32-GFP linear donor DNA: (A) Agarose gels revealed with BET containing the up and down homologous fragments and the p32-GFP cassette for Gibson assembly. **(B)** Agarose gels revealed with BET containing the cleaned up Gibson assembly and the PCR of the whole assembly from the mixes. The PCR performed from the low DNA concentration Gibson assembly has a fainter band at ~4000bp but is less contaminated. The section framed in red has been extracted out of the gel. **(C)** Agarose gels revealed with BET containing the PCR product of the whole assembly from the DNA retrieved in B. The PCR product was cleaned up and the elution was used as donor DNA for the GFP mutant

Using the protocol to activates the natural competence of *S. thermophilus* (see Material and methods), we were able to integrate the donor DNA in the cells as single stranded DNA. The transformants were selected on media containing 5'FOA. A couple of candidates were selected and the integration of the p32-GFP cassette was verified by Sanger sequencing (Microsynth AG).

3.1.2 Evaluation of the fitness of the mutant

Once the mutant was generated, we wanted to check whether the genetic modification could have a negative effect on the growth or the fitness of the bacterium.
Figure 21 is showing growth curves for *Streptococcus thermophilus* WT and the GFP mutant.

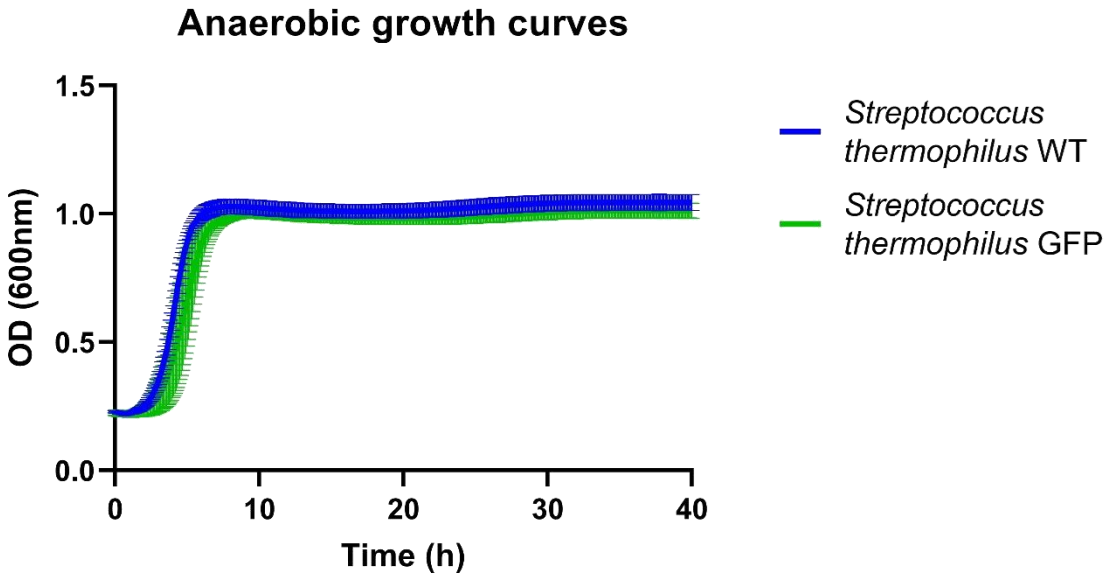

Figure 21. Comparative growth curve of the GFP mutant with *Streptococcus thermophilus* WT:
The same volume of overnight culture was re-inoculated in 200µl of culture media. The bacteria were incubated at 37°C in an anaerobic environment (see Material and methods). The $OD_{(600nm)}$ was calculated every 5 min by a micro-plate reader (TECAN). Each growth curve was performed in biological triplicates.

From the two growth curves, we can see a small delay in the growth and a slightly lower final $OD_{(600nm)}$ for the GFP mutant since p32 promotor is quite strong. This could be due to the GFP mutant using more energy to produce the GFP. Nevertheless, this is small fitness cost and we consider that the constitutive expression of GFP is not toxic for *Streptococcus thermophilus*. This should not cause any problem in future directed evolution experiments since potential revertants will be rapidly purged out during FACS sorting.

3.1.3 Fluorescent microscopy experiments

Once we saw the growth curve of our GFP-expressing strain, we wanted to check if the GFP was expressed at a sufficient level for detecting a fluorescent signal. Our first tests with naked eyes and with a micro-plate reader (HIDEX) failed to detect any fluorescent signal in culture of GFP-expressing strain. We then checked the GFP signal by fluorescence microscopy. In Figure 22, we can see the pictures taken with the microscope of a WT culture, a GFP culture and an equi-volume mix of the two cultures.

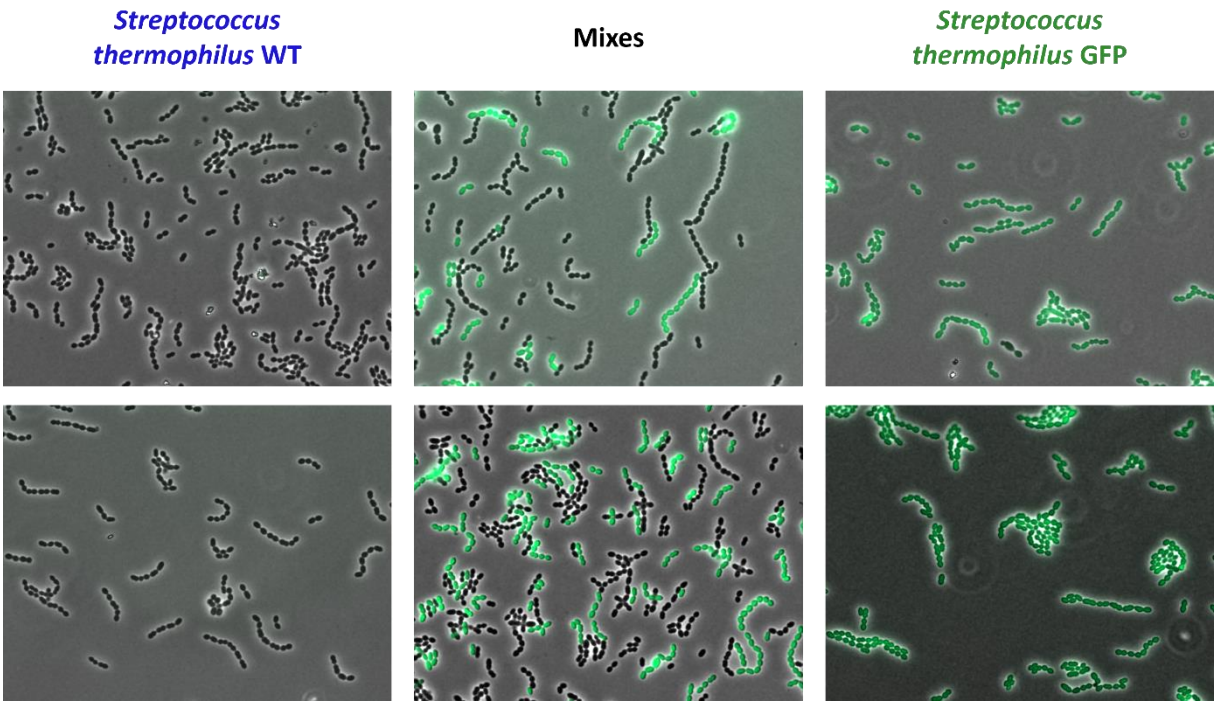

Figure 22. Fluorescent microscopy with *Streptococcus thermophilus* WT and GFP: Using a fluorescent microscope, we were able to highlight the GFP production of our mutants. The images were obtained by superposition of brightfield transmission and the fluorescent signal.

Those pictures prove that our GFP mutants are indeed producing GFP at a sufficient level to be detected by fluorescent microscopy. A better view of the distribution of GFP expression is obtained with just fluorescent images. In Figure 23, we can see that almost all the cells have a similar homogeneous fluorescent signal. Only a few are brighter and it could be explained by differences in the GFP maturation, or cells in a high metabolism state (natural competence, stress,…).

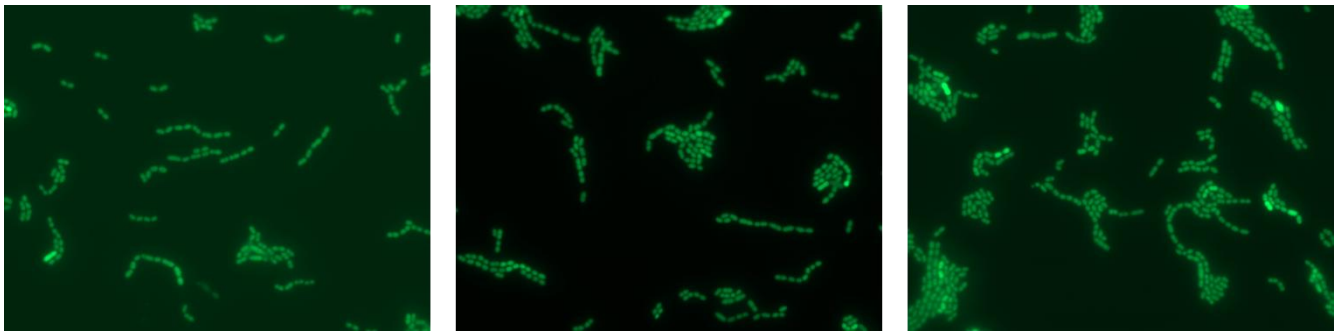

Figure 23. Fluorescent microscopy with *Streptococcus thermophilus* GFP green channel: By isolating the green channel, we can see that the level of expression seems constitutive in each cell.

As expected, the fluorescent microscopy experiment shows that our GFP mutant is indeed expressing GFP at a constitutive level. The small fraction of brighter cells may generate biases in FACS sorting experiments and reduce the screening efficiency. Filtering out false positive backgrounds may require multiple cycles of growth and sorting.

3.1.4 Flow cytometry experiments

For the flow cytometry experiments, we decided to use *Escherichia coli BL21 DE3* as a positive control. We developed an *E. coli* mutant that can express the GFP when induced with isopropyl-β-D-thiogalactopyranoside (IPTG). However, after the creation of the mutant, we noticed that the inducible promoter was leaky and that we could detect with naked eye GFP expression in cell culture without the IPTG induction step.
Nonetheless this strain was used as positive control for GFP expression for the flow cytometry experiments (without any induction step).

The first goal of those tests is to see if the level of GFP expression in *Streptococcus thermophilus* is sufficient enough to be detected by flow cytometry.
Secondly, the main goal of those experiments is to evaluate the impact of chain-length heterogeneity on the FACS read-out by comparing to *E.coli*, which does not form cell chains.

In Figure 24, we can see two graphs comparing the fluorescent signal of the *E. coli* and *Streptococcus thermophilus* WT and GFP. As predicted, the fluorescent signal of *E. coli* GFP is more important and its population is slightly more distinct from the WT population than with *S. thermophilus.*
Nevertheless, the expression level of GFP in the *S. thermophilus* mutant is sufficient enough to be detected and distinguished from the intrinsic fluorescence of the WT. Unlike *E. coli*, the *S. thermophilus* data displays multiple peaks indicative of the cell chains phenotype.
Also, it seems that a small subpopulation of the *S. thermophilus* GFP mutant has a smaller fluorescent signal and is mixed up with the WT population. This subpopulation has not been detected with the fluorescent microscopy experiments. The apparition of this effect in the flow cytometry experiments could be explained by dead cells or small residue that are detected by the flow cytometer. Other runs should be performed to analyse if this effect is reproductible.

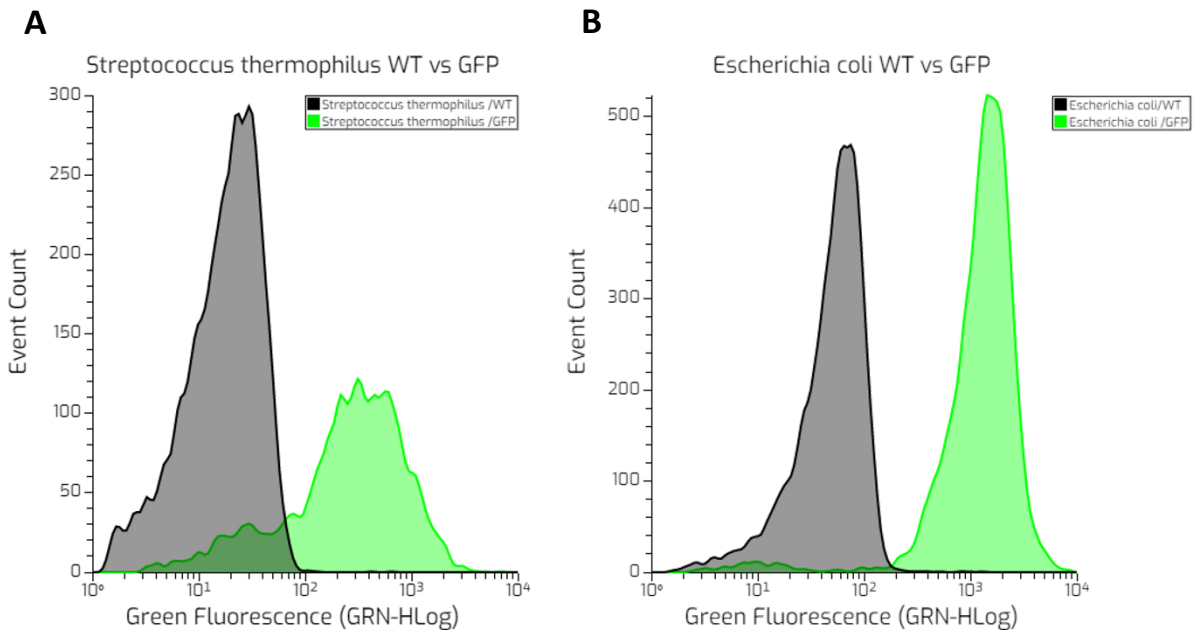

Figure 24. Comparative graphs of the green fluorescent signal detected by the flow cytometer for the WT and GFP mutant of *Streptococcus thermophilus* and *Escherichia coli*: (A) The intrinsic green fluorescence of *Streptococcus* WT can be seen in black while the fluorescence of the GFP mutant is in green. The GFP mutant is expressing the fluorescent protein at a level detectable by the flow cytometer, and the WT and the GFP population are distinguishable. **(B)** As for the *Streptococcus thermophilus* graph, the WT is in black and the GFP mutant in green. As expected, the fluorescent signal of the *E. coli* GFP mutant is higher than with *Streptococcus*.

Since we saw that our GFP mutant is detected as expected by the flow cytometer, we wanted to see if it could be used to normalise the signal per number of cells in chains.
To evaluate this, in Figure 25, we first looked at the scatter plot of the side scatter vs the green fluorescent signal (in a logarithmic scale). The side scatter provides information on the granularity of the object (Jahan-Tigh *et al.*, 2012). In the case of cell chains, we can imagine that the granularity of the chains rises with the number of cells composing it. The side scatter data could then be used as information regarding the number of cells per chain.

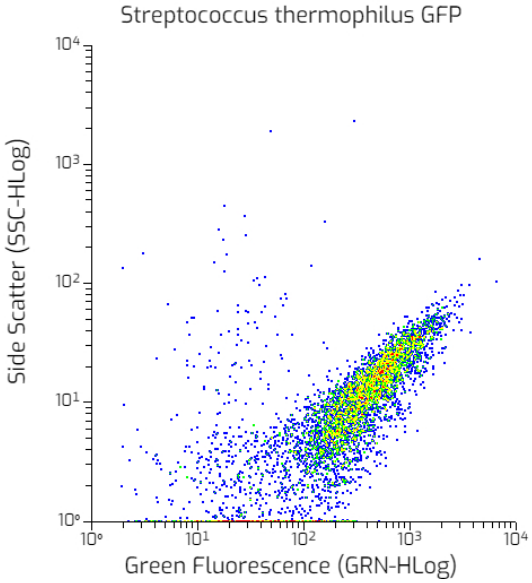

Figure 25. Scatter plot displaying the green fluorescent signal (x axis) and the side scatter (y axis) of the *Streptococcus thermophilus* GFP population: We can observe a linear correlation between the side scatter and the green fluorescent signal. This is confirming the theory that the side scatter could be used to evaluate the number of cells per chain

In Figure 25, we can see that the events of low fluorescence have a highly dispersed side scatter distribution, confirming that the subpopulation detected in Figure 24A are probably cellular debris or dust particles.
Also, we observe a linear correlation between the side scatter and the green fluorescent signal. This is confirming the hypothesis that side scatter could be an alternative to GFP for normalising the signal per number of cells in chains.

3.1.5 Sonication effect on the cell chains

We also evaluated the effect of sonication on the chain lengths. The aim of the sonication was to disrupt the chains of cells without breaking the cells.
Various different protocols of low sonication were tested without success. We still could see the cell chains with a microscope.
We then decided to perform sonication with the same intensity as in protein extraction protocols but with shorter exposition time. This solution is more harmful to the cells and was probably destroying some of them but it was the only sonication protocol that had a significant effect in the dissociation of the cell chains. The effect of the sonication has been evaluated with fluorescence microscopy and flow cytometry.

3.1.5.1 Fluorescence microscopy on the sonicated cells

Once we develop a sonication protocol able to dissociate the cell chains, we took some pictures with the fluorescent microscope of sonicated and unsonicated cultures to see the effects of sonication on the chain cells.
In Figure 26, we can see the comparison of a WT and a GFP cultures sonicated and unsonicated as well

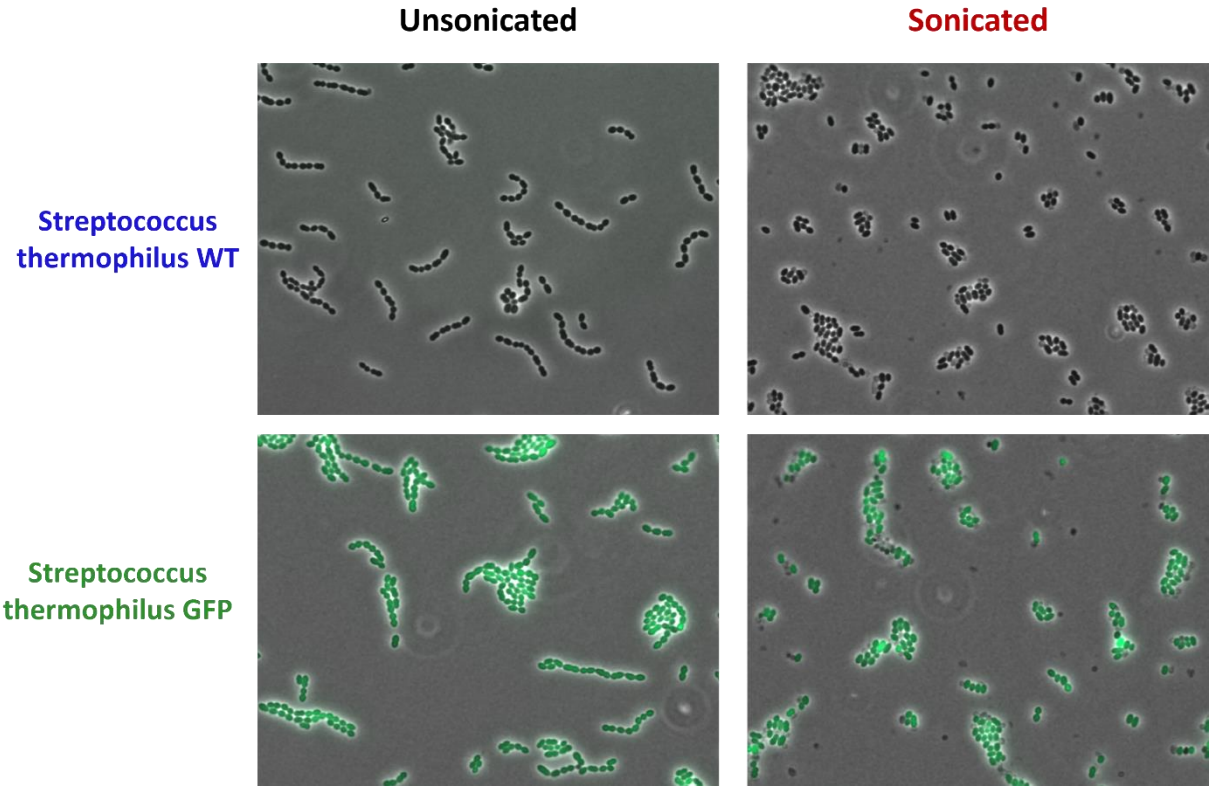

Figure 26. Florescence microscopy on sonicated and unsonicated cultures: Those picture show the effects of sonication on the cell chains.

In the picture, we can see the direct effect of sonication on the cell chains. The sonicated cultures seemed to have a lot of cells completely free and some small chains up to 4 cells.
Therefore, we concluded that the sonication protocol is effective for the dissociation of the cell chains.
During the microscopy experiment, we did notice some cellular debris that could be a consequence of cell destruction by sonication. However, we didn't perform a viability assay of the sonicated culture and so we can't evaluate if the sonication has a strong effect on the cell death.

We also counted the total number of cells in each chain in 6 microscope pictures from sonicated cultures to directly measure the distribution of chain lengths. A total number of 1435 chains of cells or isolated cells were counted in total to make the graph in Figure 27.

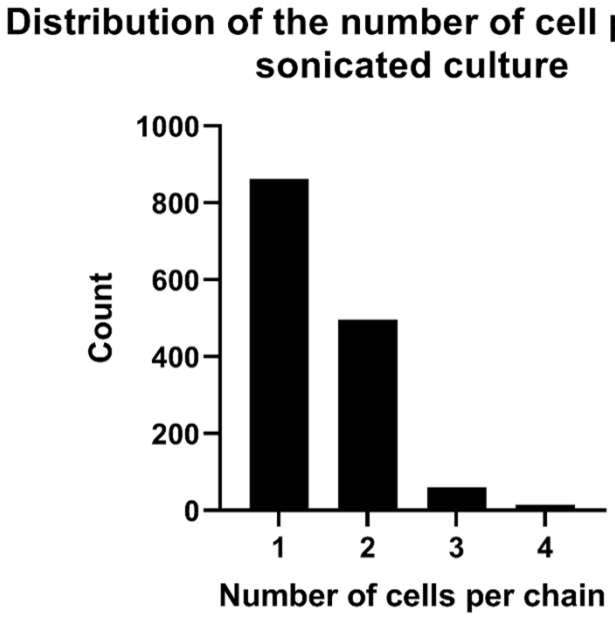

Figure 27. Histogram displaying the distribution of the number of cells per chain in sonicated cultures: Using 6 microscope pictures, we counted the total number of cells per chain in order to display the distribution in sonicated cultures. A total number of 1435 chains were counted to develop this graph.

This distribution shows us that after sonication, most of the cells are isolated or in pairs. The longest chains detected were containing 4 cells.

3.1.5.2 Flow cytometry experiments on the sonicated cultures

Subsequently, we tried flow cytometry experiments on the sonicated cultures to highlight the effects of the sonication.
In Figure 28, we can see a graph comparing the side scatter of sonicated and unsonicated cultures. The graph highlights the direct effect of the sonication on the side scatter distribution. Indeed, the peak of the distribution of the sonicated culture is shifted to the left in agreement with the microscope images.

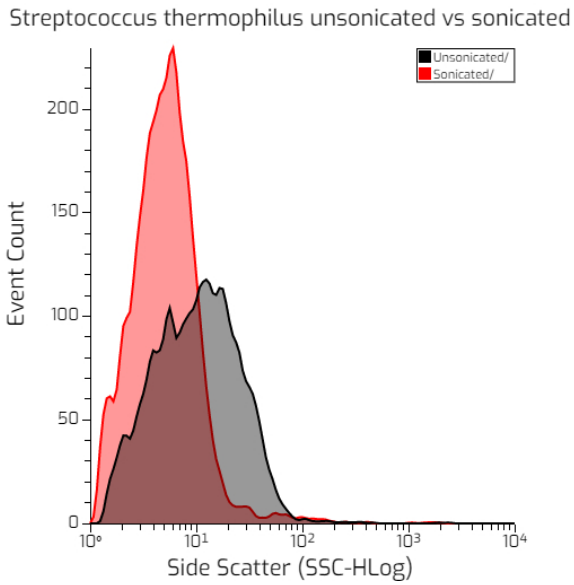

Figure 28. Graph comparing the side scatter of sonicated and unsonicated cultures: We can see a direct effect of the sonication on the side scatter distribution. The peak of sonicated culture is shifted to the left and narrows.

Another effect that can be seen in the graph is the narrowing of the peak of the sonicated cultures. Indeed, as seen in Figure 27, the dispersity of the number of cells per chain is decreasing when the cultures are sonicated. We only saw chains of cells ranging from 1 to 4 cells per chain. The narrowing of the peak is the direct consequence of this effect.
Since we saw that the sonication had a strong effect on the side scatter, we made the scatter plot displaying the green fluorescence vs the side scatter for the sonicated culture. With unsonicated cultures this graph showed a linear correlation between the two factors.

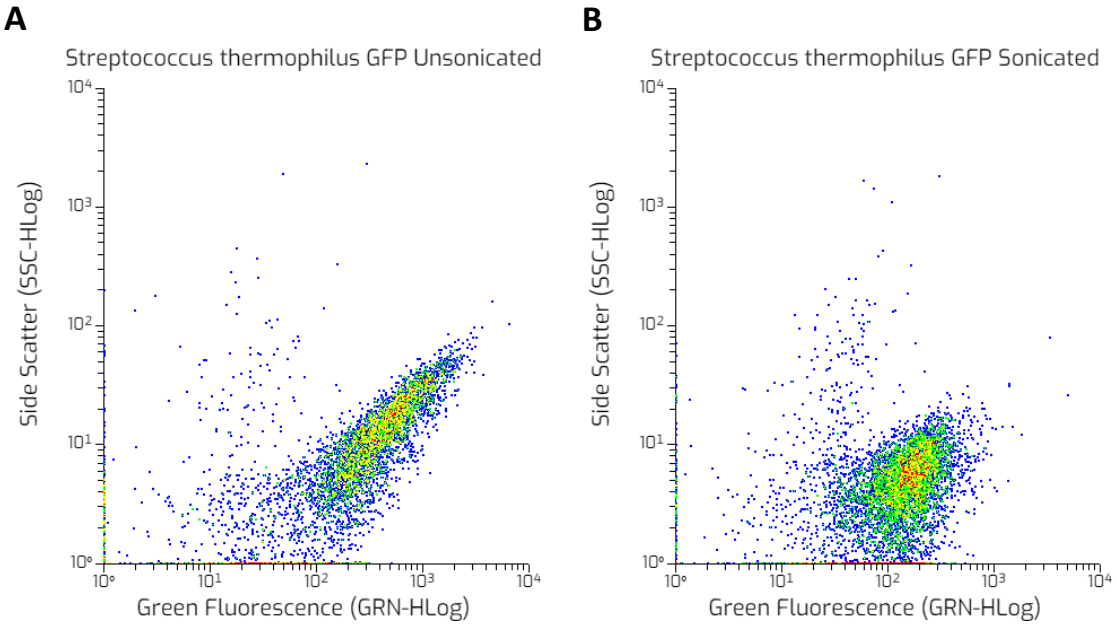

Figure 29. Scatter plot displaying the green fluorescence (x axis) and the side scatter (y axis) of the *Streptococcus thermophilus* GFP population sonicated and unsonicated: (A) As mentioned in the last chapter, we can see a linear correlation between the side scatter and the green fluorescent signal. **(B)** We can see that the correlation disappears with sonicated culture.

As expected, the decrease in side scatter dispersity is correlated with a decrease in GFP fluorescence dispersity (Figure 29). According to the distribution calculated from microscope pictures (Figure 27), the events in the graph are corresponding to the chains of cells ranging from 1 to 4 cells. As for the unsonicated culture, we can see cellular debris or other residue characterized by low fluorescent signal and a highly dispersed side scatter.
The proportion of those event seems to be even between the two graph suggesting that the

sonication protocol could not have a powerful effect on the cell death. In order to confirm this hypothesis, we should perform a viability assay on the sonicated culture. For example, we could measure the colony forming unit (cfu) of a culture before and after sonication. If the effect of sonication is not harmful to the cells, we should notice an increase of the cfu after sonication.

3.2 *Streptococcus thermophilus* display: α-amylase display

The second and main objective of this master thesis was to develop a display platform with *Streptococcus thermophilus*.
To do so, we used the features of PrtS, a naturally anchored proteins in *S. thermophilus*.
Firstly, in the N-terminus, the sequence of PrtS harbours a signal peptide that is recognised by the sec pathway (translocation pathway).
Secondly, in the C-terminus, the motif LPNTG is recognized by a sortase that will covalently link the protein to the peptidoglycan through a transpeptidation reaction.
For designing a display platform, we replaced the gene encoding PrtS with our gene of interest while keeping the promotor and the sequences encoding the signal peptide and the LPNTG motif.
As a first test of the display strategy, we used the α-amylase from *Bacillus licheniformis*.

Two mutants were designed to assess the efficiency of the display (Figure 18).
The first one, called Amyl-Display, is the mutant for the display of the α-amylase.
The second mutant only has the signal peptide of PrtS. The Amyl-Stop mutant will produce the α-amylase but the enzyme will be secreted due to the lack of the LPNTG motif.

A secondary objective was to design a controllable display platform.
To this aim, a third mutant called AmyL-BIL, contains the sequence of the alpha-amylase fused to the signal peptide and the anchoring motif of the PrtS gene. Moreover, a bacterial extracellular intein-like module (BIL) is inserted between the end of the alpha-amylase and the anchoring motif. This third mutant will be used as a controllable display platform. Indeed, since the protein of interest is linked through a disulfide bridge, it could easily be detached by using a reducing agent such as dithiothreitol (DTT).

3.2.1 Development of the mutants

The mutants were generated using the same protocol of double transformation with an intermediate mutant with the CAT-oroP cassette as discussed in section 3.1.1.
The linear donor DNA for each mutant was created with Gibson assembly just as in section 3.1.1 and the integration in the genome of *S. thermophilus* was performed by activation of its natural competence (see Material and methods). Constructs were validated by Sanger sequencing.

3.2.2 Verification of the fitness of the mutants

Once the mutants were generated using the natural competence of *Streptococcus thermophilus* (see Material and Methods), we generated the growth curves for each mutant with the same protocol used in the section 3.1.2.

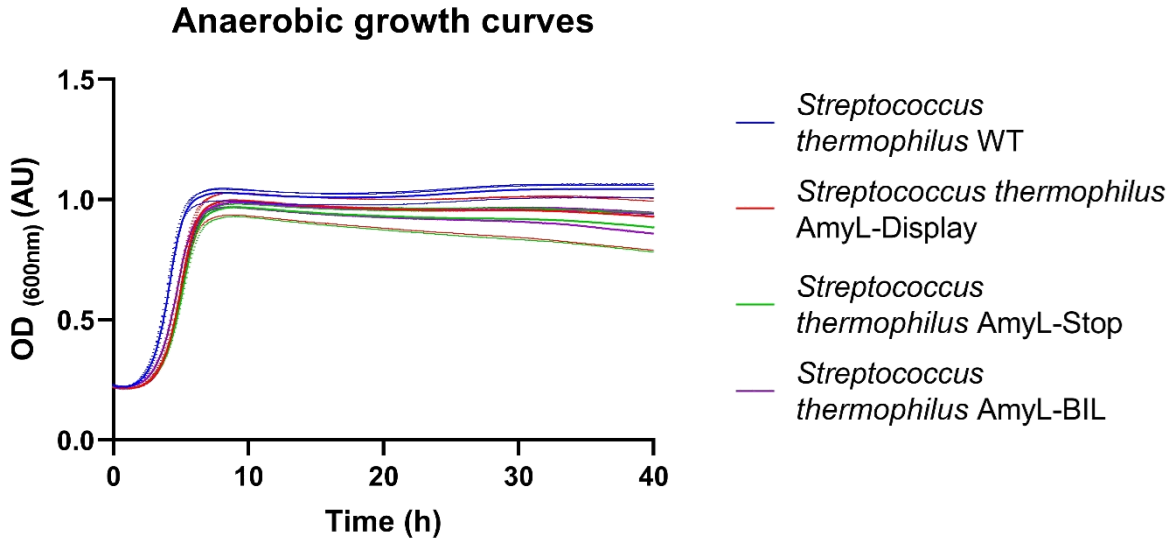

Figure 30. Comparative growth curve of the display mutants: The same volume of O/N culture was re-inoculated in 200µl of culture media. The bacteria were incubated at 37°C in an anaerobic environment. The $OD_{(600nm)}$ was calculated every 5 min by a micro-plate reader (TECAN). Each growth curve was performed in biological triplicates.

Again, the growth curves in Figure 30 show a small delay in the growth and a lower final $OD_{(600nm)}$ for the mutants. This could be due to the deletion of the gene encoding for PrtS. Somehow, the absence of this protein could have a negative impact on the growth and the fitness of the bacterium. Nonetheless, the impact seems to be moderate, and the production of α-amylase does not seem to be toxic for the cells.

3.2.3 Preliminary amylase activity tests

One of the reasons for which we chose an α-amylase as the first displayed enzyme, is the fact that the amylase activity can be easily tested with a simple experiment.
The test relies on the dark blue complex formed between starch and iodine. The activity of the α-amylase is the digestion of the starch. Therefore, if a mutant expressing the α-amylase grows on a plate containing starch, a colourless halo will form around the colonies where starch is consumed.
As a positive control, we tried this test on *E. coli BL21 DE3* transformed with an expression plasmid carrying the α-amylase gene.

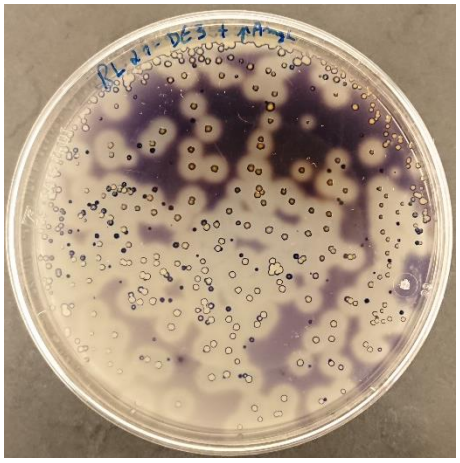

Figure 31. Iodine-starch test on *E. coli* BL21 DE3 excreting the α-amylase: The plate complemented with 0.2% of starch was exposed to iodine. Starch and iodine form a dark blue complex. Since the bacteria are excreting α-amylase, a halo can be seen around the colonies due to the digestion of the starch by the enzyme. In this plate, we can see a contamination with some colonies in black that are not expressing the α-amylase.

In Figure 31, we can see a positive control of this iodine-starch test. As expected, a colourless halo can be seen around the colonies due to their α-amylase excretion.
Moreover, this test highlights a contamination with some colonies in black that don't seem to express the α-amylase.
The same test was then performed on our three mutants to test their amylase activities (Figure 32).

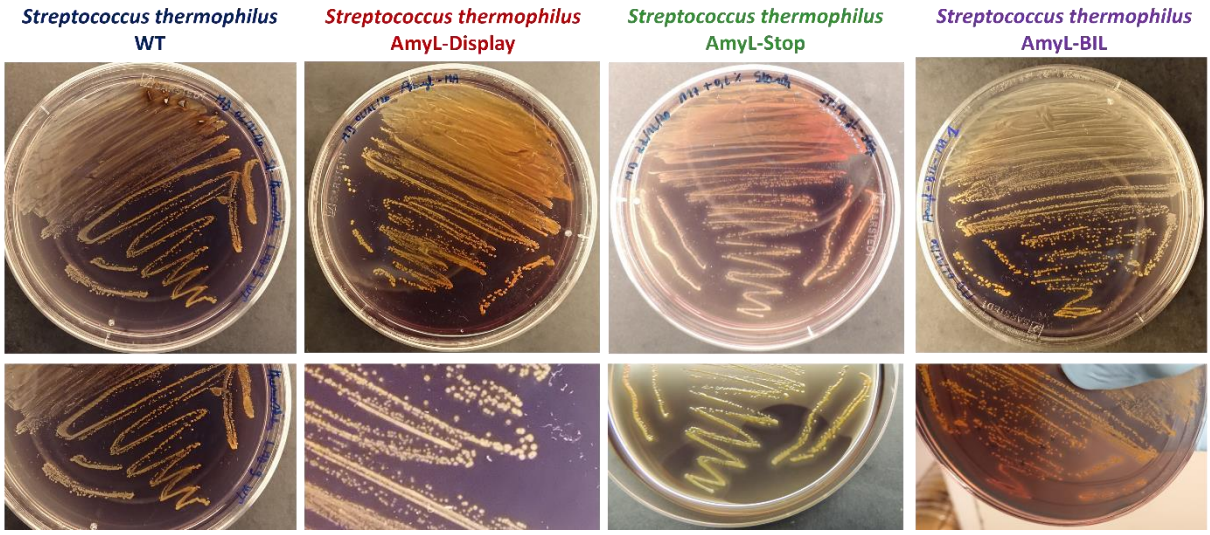

Figure 32. Iodin-starch test on the 3 display mutants: The test on the three mutants reveals no halo for the WT and AmyL-Display, a faint halo for AmyL-BIL and a bigger halo for AmyL-Stop.

In Figure 32, we can see that the iodine-starch preliminary test showed no sign of α-amylase activity for the AmyL-Display mutant, a faint activity for the AmyL-BIL mutant and a higher activity for the AmyL-Stop mutant. Indeed, even if it is difficult to see it in the picture, we observe a small halo around the colonies in the AmyL-BIL mutant while the AmyL-Stop mutant revealed a bigger halo.

After seeing those results, our hypothesis is that this test allows to see the amylase activity only for excreted α-amylase. As a matter of fact, in order to see a halo, the α-amylase needs to diffuse on the agar plate.
This hypothesis explains why we don't see α-amylase activity for our AmyL-Display mutant. Because with this mutant, the α-amylase is not excreted but directly anchored to the peptidoglycan. It could

also explain why we can see a small halo for the BIL mutant. Indeed, in this mutant, the α-amylase is linked to the peptidoglycan through a disulfide bridge. This bond is more sensitive to cleavage by reduction than a covalent bond and we can easily assume that some could break, liberating the α-amylase to diffuse and digest starch around the colonies.

3.2.4 Amylase activity assay

In order to confirm those first good results, we performed a more precise amylase activity assay. Using a kit bought from Sigma-Aldrich©, we were able to calculate the amylase activity of our mutants. We performed the tests separately on the culture supernatants and the washed cells (see Material and Methods).
The amylase activity is determined using a coupled enzymatic assay. The chain of reactions is summarized in Figure 33.

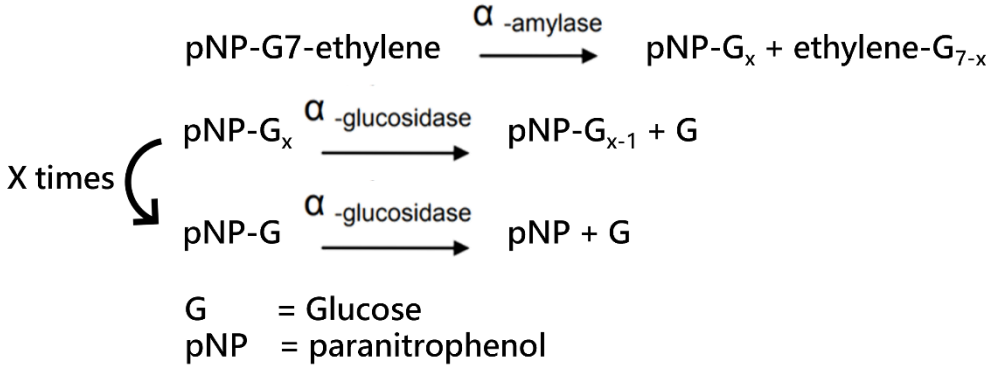

Figure 33. Chain of reactions of the α-amylase activity assay: The substrate called pNP-G7- ethylidene is a maltoheptaoside derivative cleaved by the α-amylase. One of the products is then completely hydrolyzed by the α-glucosidase already mixed up with the substrate in the kit generating glucose and the chromogenic paranitrophenol (405 nm). The quantity of free paranitrophenol produced is proportional with the α-amylase activity

The action of the α-amylase coupled with the α-glucosidase (already mixed up with the substrate in the kit) lead to the production of free paranitrophenol, a colorimetric product (405nm).

The α-amylase activity was calculated as indicated by the manufacturer's protocol by calculating the amount of paranitrophenol generated. The quantity of paranitrophenol produced was evaluated by measuring the absorbance at 405nm of the sample every 5 min. The maximal velocity at steady-state was then measured as a quantitative indicator of AmyL activity.

3.2.4.1 Evaluation of wall-associated alpha-amylase activity

The amylase activity calculated with the kit is measured in nmol/ml/min. The kit is used to measure a relative activity in order to compare it between different samples. We used *E. coli BL21 DE3* transformed with the plasmid carrying the α-amylase gene (pAmyL) from the section 3.2.2 as a positive control to have a comparison with a surexpression host.

For the first assays, we tested our positive control for expression (*E. coli* transformed with pAmyL), *E. coli* and *Streptococcus thermophilus* WT, our negative control for the display platform (AmyL-Stop) and the mutant for the display (AmyL-Display).
For each mutant, an overnight culture was re-inoculated and incubated until the $OD_{(600nm)}$ reaches ~0.7. Afterwards, the supernatant and the cells were separated by centrifugation. Both supernatant

and resuspended cells were prepared for the amylase activity assay (see Material and Method). The activity assay was performed in triplicates for each mutant. Since the OD_{600nm} of the culture were heterogenous, the final activity was normalised for a starting cultures at $OD_{(600nm)}$=0.7.

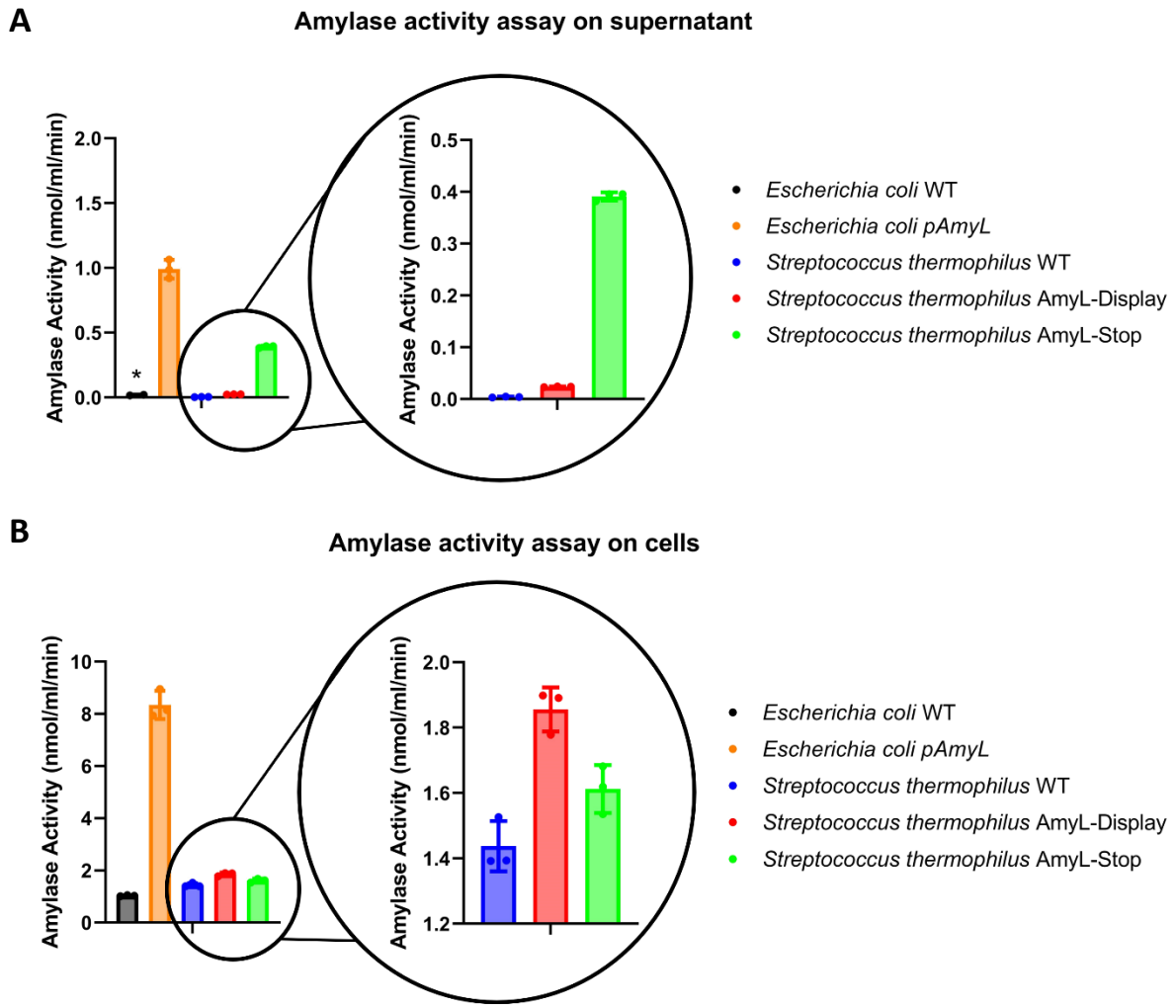

Figure 34. Amylase activity assay on supernatant and pelleted cells: The final activity was normalized for a starting culture at $OD_{(600nm)}$=0.7. The activity assay was performed in triplicates for all mutants excepted those with a star on top of the histogram that are duplicates. **(A)** Relative activity of the α-amylase in the supernatant. We can see that the supernatant of *E. coli* pAmyL mutant contains the highest amylase activity, more than twice the activity of the best *Streptococcus thermophilus* mutant (AmyL-Stop). **(B)** Relative activity of the α-amylase with the cell suspensions. The first observation is that *E. coli* and *S. thermophilus* WT both display a residual α-amylase activity associated with the bacteria. As for the supernatant, the activity with *E. coli* pAmyL is again the strongest. The *Streptococcus* mutant that has the strongest activity is the AmyL-Display.

In Figure 34A, we can see the amylase activity calculated for the supernatant of the culture. As expected, *E. coli* pAmyL has the strongest activity suggesting that the production of the enzyme is more important with *E. coli.*
The second best activity detected is in the supernatant of *Streptococcus thermophilus* AmyL-Stop. This result is in total concordance with the preliminary tests.
The supernatant of the AmyL-Display liquid culture reported almost no amylase activity, suggesting that if the mutant is producing the enzyme, it is either in the cytoplasm or directly anchored to the cells.

Figure 35B shows the amylase activity measured with the cell suspensions.
The first observation is that *E. coli* and *Streptococcus thermophilus* WT cells both have an intrinsic

amylase activity. For *S. thermophilus*, the WT activity is more important than the activity calculated for the enzyme (1.44 for WT and 0.41 calculated for AmyL-Display mutant). This activity could be due to other enzyme naturally produced by the bacteria.
Secondly, as for the supernatant, the strongest activity is with the *E. coli* pAmyL. This result was not expected since the α-amylase produced by *E. coli* should not be excreted and so localised in the cytoplasm. It could be explained by the lysis of some cells during the washing steps that could liberate the enzyme in the media.
The AmyL-Stop cells also has an amylase activity. This mutant possesses only the signal peptide that allows the translocation of the enzyme through the plasma membrane. The activity detected with the resuspended cells could be due to some enzyme restrained in the peptidoglycan mesh.
Finally, the *streptococcus* mutant that has the strongest activity for the cells is AmyL-Display, the display mutant.

The results of this test is very important because it is the first one that shows amylase activity with the AmyL-Display mutant. The fact that the amylase activity can only be seen in the assay on the cells and almost not in the supernatant is very encouraging for the display platform. Indeed, the enzyme seems to not be excreted but strongly associated with the cell wall.

The presence or the absence of the LPNTG anchoring motif is the only difference between AmyL-Display and AmyL-Stop. The fact that the mutant that possesses the anchoring motif is also the one that has the best amylase activity for the cells shows a proof of concept of our display platform. From a quantitative point of view, if we sum up the activity in supernatants and cell suspensions, and subtract the negative control backgrounds, we obtain 0,56 and 0,44 nmol/ml/min for AmyL-Stop and AmyL-Display respectively, indicating similar total levels of expression.

3.2.4.2 Evaluation of the controllable display platform

The second amylase assay performed was focussed on the display platform with the BIL module. Here the α-amylase should be anchored to the peptidoglycan through a disulfide bridge. In contrast to the regular display platform, the enzyme is linked through a covalent bond that should easily be cleaved using a reductor agent.

In this set of experiment, AmyL-Display, AmyL-Stop and AmyL-BIL mutants were treated with 10mM of DTT (reducing agent) before separating supernatants and cells by centrifugation. The same protocol as the last experiment has been used except for the DTT treatment (See Material and methods).

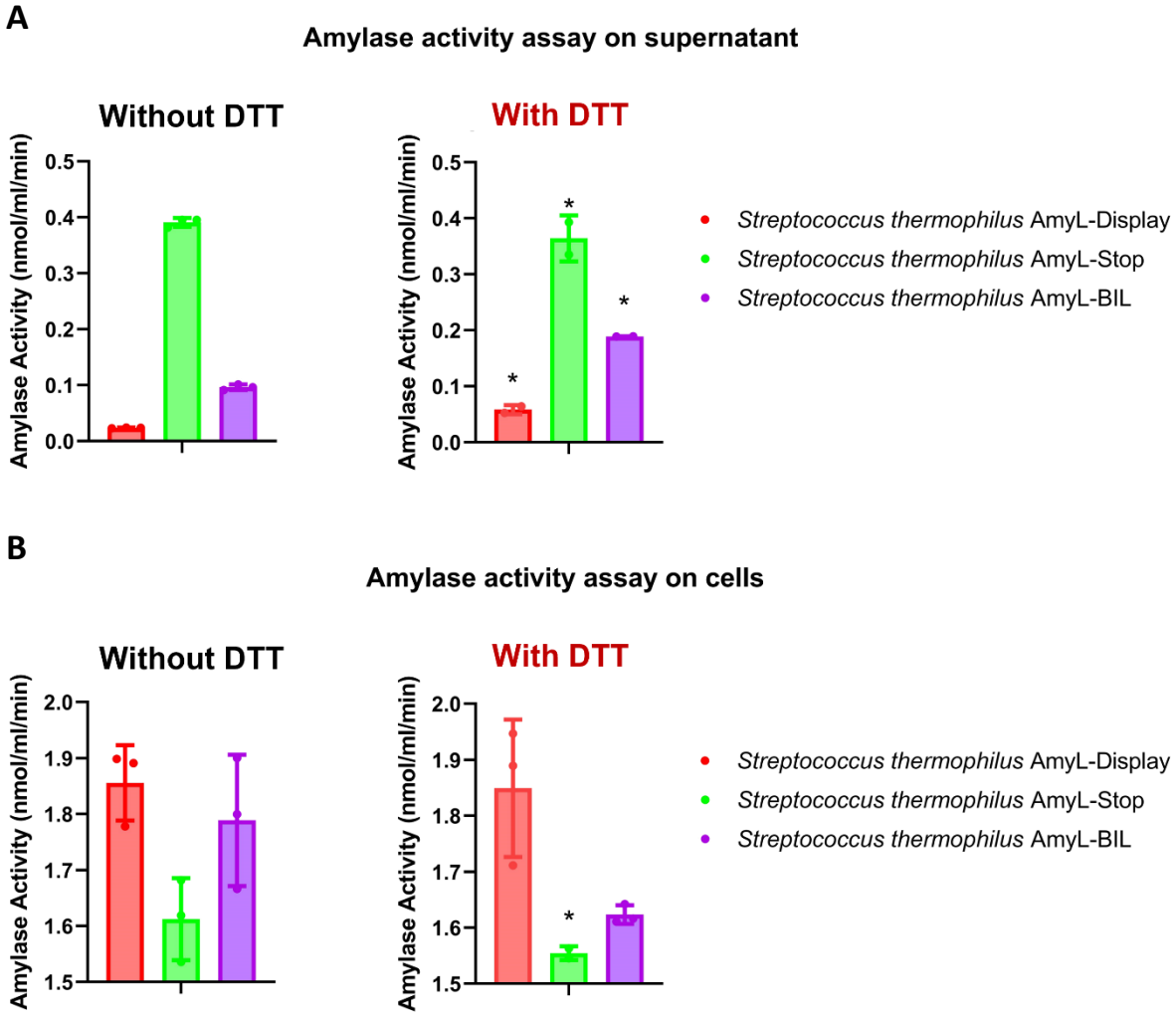

Figure 35. Amylase activity assay on supernatant and pelleted cells with or without DTT treatment: The final activity was normalized for a starting culture at $OD_{(600nm)}$=0.7. The activity assay was performed in triplicates for all mutants excepted those with a star on top of the histogram that are duplicates. **(A)** Relative activity of the α-amylase excreted in the supernatant by the bacteria. Each mutant was tested with or without a DTT treatment before the centrifugation. We can see that the supernatant of a culture of AmyL-BIL mutants has a small amylase activity. After the DTT treatment, we can see almost a 2-fold increase of the activity on the supernatant. **(B)** Relative activity of the α-amylase on the cells. Each mutant was tested with or without a DTT treatment before the centrifugation. Firstly, the amylase activity on the untreated AmyL-BIL cells is stronger than on the AmyL-Stop cells. After the DTT treatment, this activity is reduced.

In Figure 35, we can see that, without the DTT treatment, the AmyL-BIL mutant seems to have the same activity pattern as the AmyL-Display mutant. Indeed, the activity is mainly associated with the cells, contrary to the AmyL-Stop mutant. Again, those results are showing that with the anchoring motif fused in frame with the amylase, the enzyme is expressed and strongly associated to the cell. Although, as suspected with the preliminary iodine-starch test, the activity of its supernatant is stronger than the AmyL-Display mutant. Two reasons could explain this difference. First, the disulfide bond that linked the enzyme to the peptidoglycan is more susceptible to be cleaved than the covalent bond of AmyL-Display. Secondly, the activity of the BIL module could interfere with the anchoring motif, disrupting its recognition by the sortase.

When the cultures are incubated with DTT before the amylase assay, the relative activity of each mutant is altered. While for AmyL-Display and AmyL-Stop those changes are moderate, we can see a

stronger effect of the DTT treatment for the AmyL-BIL mutant. Indeed, the DTT incubation causes the augmentation of the amylase activity in the supernatant (2-fold) and the decrease of the activity of the cells. This effect is highlighted in Figure 36.

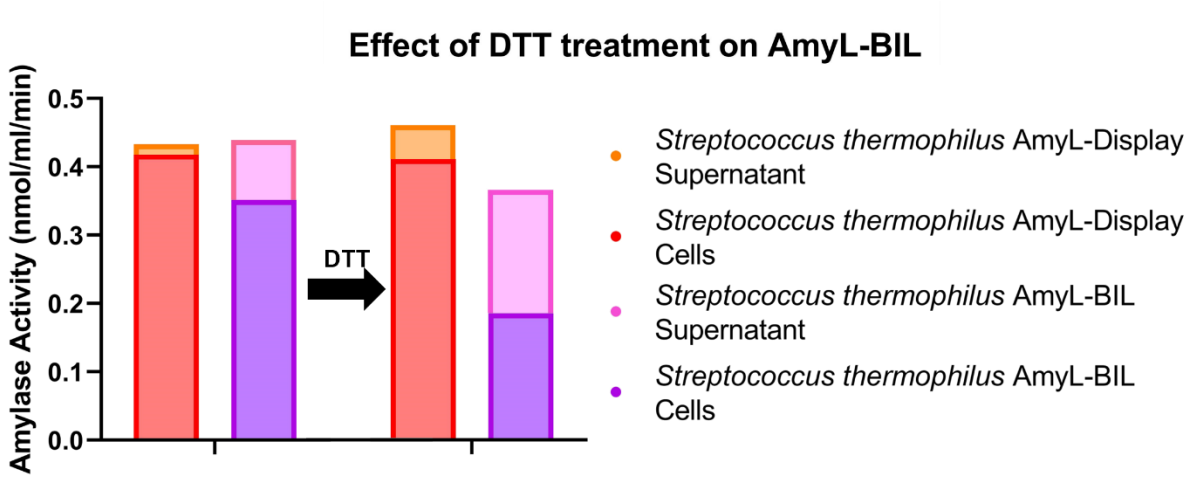

Figure 36. Histogram comparing the AmyL-Display display with the AmyL-BIL display: Without DTT treatment, the two display platforms are relatively similar. The relative expression of the α-amylase is identical in both mutants. The only difference is that the AmyL-BIL display is releasing more α-amylase in the supernatant. When the AmyL-BIL mutant is incubated with DTT, even more amylase is released, increasing the activity in the supernatant at the expense of the cell activity.

Figure 36 is comparing the total AmyL activity for both display systems showing that the sum soluble and anchored activity is more or less constant and independent of DTT treatment. The main difference is that about 40% of the activity of AmyL-BIL construct can be released in the medium by DTT treatment. The release is not quantitative, suggesting that part of the enzyme remains trapped in the bacterial wall. This may be due to non-covalent physical entrapment as it is also observed with the AmyL-Stop construct, and/or to inefficient BIL processing. Overall, these results validate the cleavable system based on BIL processing.

4. Conclusions and perspectives

4.1 Normalisation of signal from cell chains

One of the objectives of this master thesis was to overcome the drawback of the cell chains of *Streptococcus thermophilus* for the screening of cell chains with FACS.

To do so, we developed a mutant that constitutively expresses GFP. The strategy was to use the GFP fluorescent signal to normalise the future read-out signal per number of cells in chains.
The first test with the GFP mutant shows that we were able to detect the fluorescent signal with flow cytometry and that it was correlated with the number of cells per chain.
However, we highlighted some heterogeneity in the intracellular expression that could causes the introduction of biases in the normalisation. As a matter of fact, we highlighted a small fraction of brighter cells that could reduce the screening efficiency by discarding mutants with good activity due to an excessive normalisation by the GFP signal.
In addition, the use of GFP for the normalisation of the signal from cell chains reduces the possibilities of fluorophore to be used for FACS screenings of mutants since their emission and excitation wavelength cannot overlap.

The results gathered in this section led us to think that the side scatter measured by flow cytometry could be an alternative to the GFP for the normalisation. The advantage of the side scatter is that we would no longer need to use a mutated strain expressing GFP.

Finally, we saw that it was possible to reduce the size of the chain cells by sonication. This could be useful to reduce the variance in the cell chains population but we would need to perform a proper viability assay on the sonicated culture first. If the effect of sonication on the cell death is not too important, it could be used before the screening of the mutant.

4.2 Development of the display platform with *Streptococcus thermophilus*

The second and main objective of this master thesis was to develop a display platform with *Streptococcus thermophilus* as host.

To do so, we used the features of a naturally displayed protein called PrtS to display an α-amylase. In order to evaluate the display strategy, two mutants were constructed. The display mutant (AmyL-Display) and a negative control lacking the LPNTG anchoring motif (AmyL-Stop).
Using an amylase activity assay kit, we tested the activity of the supernatant and the pelleted cells from a liquid culture of each mutant.

The results with the amylase assay kit shows a proof of concept of the display technology with *Streptococcus thermophilus*. Indeed, it appears that when the LPNTG motif is fused to the C-terminus of an excreted enzyme, that enzyme becomes closely associated with cells. This effect is probably due, as expected, to the anchoring of the enzyme to the peptidoglycan by the sortase. However, we

don’t have the concrete proof of this covalent anchoring. The enzyme could also be trapped in the peptidoglycan mesh.

In order to prove covalent anchoring, we could first start by displaying a tag (his-, flag- or myc-tag). Afterwards, we could use a mix of mutanolysin and lysozyme to digest the peptidoglycan layer of mutants displaying the tag. After affinity capture based on the tag, analysis by mass spectrometry should allow to identify tag-PG fragments.

Even though we collected interesting data with the α-amylase display, the fact that the cells had an intrinsic amylase activity, more important than the activity of the displayed enzyme, was problematic for the assay. The display platform should be tested with another enzyme with a substrate that is not reacting with the WT strain. For example, we could test the display platform with a completely artificial retro-aldolase engineered in silico and through directed evolution by the groups of D. Baker and D. Hilvert (Giger *et al.*, 2013). As for the α-amylase, the activity of this enzyme can be easily evaluated by the release of a colorimetric product after interaction with a substrate called methodol. With this assay, we should not detect an intrinsic activity of the WT cells.

Also, we could optimize the display platform to increase the expression and so the display rate. In this master thesis, we used the promotor and the signal peptide of PrtS for the expression and the translocation of our protein of interest. Other promotors and/or signal peptides could be used to increase the expression and the display rate. By anchoring a tag recognised by an antibody fused to a fluorophore such as a Flag-tag, we could easily evaluate the display rate of various mutants and select for the best. With the anchorage of a tag we could even try to evolve the promotor and the signal peptide by directed evolution.

Once our display platform is fully confirmed and optimized, we will be able to apply it for directed evolution campaigns. For example, we could try to evolve the sortase of *S. thermophilus* itself. The ability of this enzyme to ligate biomolecules together through a covalent bond is a useful biotechnological tool such as for the peptide synthesis (Jacobitz *et al.*, 2017).
Using our display platform, we could evolve the sortase toward new anchoring motif recognition. Indeed, we could express a Flag-tag fused to a signal peptide and a modified anchoring motif in mutants transformed with a library of sortase. The mutants would be screened by labelling the cells with antibodies anti-tag fused to a fluorophore and we could sort the best mutants by FACS.
At first, we could try with simple mutations that already exist in sortase homologs to test our system and afterwards we could go for more challenging specificity changes. Such study would also help understanding structure-function relationships of the sortase.

Another interesting experiment would be to evolve the sortase to ligate non-natural peptides.

4.3 Development of a controllable display platform

A third objective of this master thesis was to develop a controllable display platform.
To do so, we used a small module called BIL fused in frame between the anchoring motif and the sequence of the amylase. This module auto-catalytically cleaves itself from a polypeptide while reconnecting both flanking sequences through a disulfide bridge. The new mutant called AmyL-BIL is displaying the α-amylase but instead of being linked to the peptidoglycan through a covalent bond, the enzyme is linked through a disulfide bridge.

This mutant was tested for amylase activity with the same kit as for the regular display mutants but with and without an incubation with DTT (reducing agent). We wanted to check whether the DTT treatment could detach the enzyme from the cells.

The mutants that weren't incubated with DTT had the same pattern of activity as the AmyL-Display mutant, proving that the enzyme was indeed associated to the cells. Nonetheless, the supernatant activity of the AmyL-BIL mutant was stronger than the AmyL-Display. This effect could be due to a partial BIL processing leading to peptide cleavage without disulphide formation. Such reaction would liberate the enzyme in the supernatant. Another hypothesis is that the BIL module could interact or have steric effect on the anchoring motif, preventing the anchoring by the sortase. The fact that a disulfide bridge is more sensitive to cleavage by reduction than a covalent bond could also explained this effect.
The incubation with DTT led to a 2-fold increase of the amylase activity in the supernatant at the expense of the cells activity.

Those result showed that the display platform with AmyL-BIL is promising and that the BIL display seems to work, at least partially. Since reducing agent such as DTT could be toxic to the growth of the cells, we should test lower concentration and also weaker reducing agent such as beta-mercaptoethanol (BME).

The interest of a controllable display is the fact that the mutant could be used for screening but also for the production of soluble enzymes for subsequent characterization without the need of re-cloning the gene.

The BIL display system could later be used to anchor adhesion protein such as hemagglutinin. With the controllable display technology, we could try to control the adhesion of the cells and/or the biofilm formation by incubating the cells with a reducing agent that could detach the hemagglutinin.

5. Material and methods

5.1 Material

5.1.1 Bacterial strains and growth conditions

The strains of bacteria used in this master thesis are listed in Table 1.

Name	Genotype/Plasmids	Usage
***Streptococcus thermophilus* LMD-9 mutants**		
Streptococcus thermophilus WT	Wild type strain regularly used in the laboratory	Starting point of all mutants and negative control
Streptococcus thermophilus tRNAser::p32-GFP	Strain with the insertion of a p32-GFP cassette in the tRNAser locus	Constitutive GFP expression for cell chains normalisation by flow cytometry
Streptococcus thermophilus tRNAser::p32-GFP Δ*PrtS*::Amyl-Display	Strain where *PrtS* is knockout and substituted by *α-amylase* with the signal peptide and the membrane anchoring motif of *PrtS* retained	Expression and anchoring of the α-amylase on the peptidoglycan
Streptococcus thermophilus tRNAser::p32-GFP Δ*PrtS*::Amyl-Stop	Strain where PrtS is knockout and substituted by α-amylase with the signal peptide of *PrtS* retained	Expression and secretion of the α-amylase
Streptococcus thermophilus tRNAser::p32-GFP Δ*PrtS*::Amyl-BIL	Strain where *PrtS* is knockout and substituted by *α-amylase* with the signal peptide and the membrane anchoring motif of *PrtS* retained. A small module called BIL is inserted between the *α-amylase* and the anchoring motif	Expression and anchoring of the α-amylase on the peptidoglycan through a disulfide bridge
***Escherichia coli* BL21 DE3**		
Escherichia coli WT	Wild type strain from commercial stocks	Starting point of all mutants and negative control
Escherichia coli pET26b GFP-KanR	Strain transformed with a modified pET26b plasmid carrying the sequence of GFP_{UV}	Positive control for microscopy and flow cytometry experiments
Escherichia coli pUC18 AmyL-AmpR	Strain transformed with a modified pUC18 plasmid carrying the sequence of α-amylase	Positive control for amylase activity assays

Table 1. Table of strains used during this master thesis

The *Streptococcus thermophilus* mutants were cultivated at 37°C in anaerobic conditions without agitation. Depending on the experiment, the bacteria were cultivated in liquid media M17 or CDM supplemented with 1% glucose. Solid media was prepared by adding 1.5% of bacteriologic agar. When required, the media were supplemented with chloramphenicol or 5-fluoroorotic acid (5'FOA) at a final concentration of 5µg/ml and 0.5mg/ml. The anaerobic conditions were generated by using hermetic Eppendorf filled up to the top for liquid culture and by using an anaerobic chamber and the GasPack™ EZ for the plates.
The *Escherichia coli* mutants were cultivated at 37°C in aerobic conditions with agitation (200RMP). The bacteria were cultivated in liquid media LB. Solid media was prepared by adding 1.5% of bacteriologic agar. Depending on the experiment, the media were supplemented with ampicillin or kanamycin at a final concentration of 100µg/ml and 30µg/ml.

5.1.2 Culture media

M17 supplemented with glucose (GM17)

The M17 culture media is optimized for the lactic acid bacteria growth. In this master thesis, we were using the M17 broth from Oxoid (ThermoFisher Scientific). We followed the manufacturer recipe and added 37.25g of the media in 1L of demineralised water. For this master thesis, the M17 media was supplemented with 1% glucose (GM17).
The media composition given by the manufacturer can be seen in table 2.

Component	Concentration (g/L)
Tryptone	5
Soya peptone	5
"Lab-Lemco" powder	5
Yeast extract	2.5
Ascorbic acid	0.5
Magnesium sulphate	0.25
Di-sodium-glycerophosphate	19
	pH 6.9 ± 0.2 at 25°C

Table 2. Composition of M17 broth from Oxoid (ThermoFisher Scientific)

CDM supplemented with glucose (GCDM)

The chemically defined medium is a very optimized medium used in the induction of the natural transformation of *Streptococcus thermophilus.* This medium needs to be totally prepared in the lab. The final medium is obtained by mixing 5 different solutions in this proportion:

- 860 mL of growth buffer
- 10 mL of vitamin solution 100X
- 10 mL of metal solution 100X
- 20 mL of DNA precursors 50X
- 100 mL of amino acid solution 10X

The composition of those 5 solutions is displayed in Table 3.

Component	Concentration (g/L)
Growth buffer (860mL)	
Glucose monohydrate ($C_6H_{12}O_6$)	10
Sodium acetate (CH_3COONa)	1
Triammonium citrate ($C_6H_{17}N_3O_7$)	0,6
Potassium phosphate (KH_2PO_4)	3
Potassium hydrogen phosphate (K_2HPO_4)	2,5
Urea (CH_4N_2O)	0,24
L-Tyrosine ($C_9H_{11}NO_3$)	0,29
L-ascorbic acid ($C_6H_8O_6$)	0,5
Glutamine ($C_5H_{10}N_2O_3$)	0,39
Glutamate ($C_5H_9NO_4$)	0,398
Vitamin solution 100x (1 L)	
Pyridoxamine-HCl	0,5
Nicotinic acid	0,1
Riboflavine	0,1
Calcium pantothenate	0,1
Thiamine-HCl	0,1
Pyridoxine-HCl	0,2
Para-aminobenzoic acid salt	1
Biotine	1
Folic acid	0,1
B12 vitamin	0,1
Oritic acid	0,5
Thymidine	0,5
Inosine	0,5
Lipoic acid	0,25
Metals solution 100x (1 L)	
$MgCl_2.(H_2O)_6$	20
$CaCl_2.(H_2O)_2$	5
$FeCl_2.(H_2O)_4$	0,5
$ZnSO_4.(H_2O)_7$	0,5
$CuSO_4.(H_2O)_5$	0,01
$CoCl_2.(H_2O)_6$	0,25
$MnSO_4.H_2O$	2,8
DNA precursors 50x (20 ml)	
Adenine	0,01
Guanine	0,01
Xanthine	0,01
Uracil	0,01
Amino acids solution (1 L)	
L-arginine-HCl	3,5
L-cysteine	2,5
L-histidine	1,5

L-proline	6,75
L-alanine	2,4
L-asparagine	3,5
L-acide aspartique	4,55
L-glycine	1,75
Isoleucine	2,5
L-leucine	4,75
L-lysine-HCl	4,4
L-methionine	1,25
L-phenylalanine	2,75
L-serine	2,25
L-threonine	1,75
L-tryptophane	0,5
L-valine	3,25

Table 3. Composition of 5 solutions needed to prepare CDM

The final solution is homogenized and set at pH 6.6. The solution is filtered sterilized and stored at -20°C. In this master thesis, CDM was supplemented with 1% glucose (GCDM).

Luria-Bertani (LB)

The Luria-Bertani (LB) medium was optimized for the culture of the Enterobacteriaceae. In this master thesis, we were using LB dehydrated media. The composition of the media can be seen in Table 4.

Component	Concentration (g/L)
Tryptone	10
Yeast extract	5
Sodium chloride	5

Table 4. Composition of LB medium

5.1.3 Antibiotics and selection molecules preparation

Chloramphenicol stock solution of 35mg/ml was prepared in ethanol > 96%, ampicillin stock solution of 100mg/ml was prepared in water as well as kanamycin stock solution of 30 mg/µl. 5-Fluoroorotic acid (5'FOA) stock solution of 100mg/ml was prepared in DMSO. All stock solutions were stored at -20°C

5.1.4 Primers

All the primers used for this master thesis are listed in Table 5.

Name	Sequence 5'-> 3'	Usage
Streptococcus thermophilus tRNAser::p32-GFPUV		
tRNAser upfl fwd	AACTATCATCCGTCATCCCAG	Amplification of the up homologous arm tRNAser
tRNAser upfl rev	CAGAAATAGACAAAGTGTCCCTTCC	Amplification of the up homologous arm tRNAser
tRNAser:P32-GFP fwd	GGAAGGGACACTTTGTCTATTTCTGTGAAAAATTGAGAGGTATAAATCAAATTAATAG	Amplification of the p32-GFP cassette with floating tail for Gibson assembly
tRNAser:P32-GFP rev	CCTCTCCTTTTCAGAAATAGACAAATTTAAATAATAAAAAAGCCGGATTAATAATCTGG	Amplification of the p32-GFP cassette with floating tail for Gibson assembly
tRNAser dwfl fwd	AATTTGTCTATTTCTGAAAAGGAGAGG	Amplification of the down homologous arm tRNAser
tRNAser dwfl rev	ACAACACCAATCTTGGCAG	Amplification of the down homologous arm tRNAser
Streptococcus thermophilus ΔPrtS::cat-oroP		
PrtS upfl fwd	TCTTTTCAGCGATTTCC	Amplification of the up homologous arm PrtS for all mutants
PrtS upfl-CAT-oroP rev	CTTATGGGATTTATCTTCCTTATTAAGTTGAGGTAGCCTC	Amplification of the up homologous arm PrtS with floating tail for Gibson assembly
CAT-oroP-PrtS fwd	GAGGCTACCTCAACTTAATAAGGAAGATAAATCCCATAAG	Amplification of the CAT-oroP cassette with floating tail for Gibson assembly
CAT-oroP-PrtS rev	GTGCCTGTACATCAACTTATAAAAATTTAATGAATATTATTCCGGT	Amplification of the CAT-oroP cassette with floating tail for Gibson assembly
PrtS dwfl-CAT-oroP fwd	ACCGGAATAATATTCATTAAATTTTTATAAGTTGATGTACAGGCAC	Amplification of the down homologous arm PrtS with floating tail for Gibson assembly
PrtS dwfl rev	CTGTACACCATTTCTGATC	Amplification of the down homologous arm PrtS for all mutants
PrtS upfl external fwd	CTAATCTATTGGGTGGC	Sequencing primer for all mutant
CAT-oroP internal rev	GAATGATGTACCTGTAAAGATAG	Sequencing primer
***Streptococcus thermophilus ΔPrtS::Amyl*-Display**		
PrtS upfl SP-AmyL rev	CGTCCCATTAAGATTTGCTTCATCTGCAGCTACCG	Amplification of the up homologous arm PrtS with floating tail for Gibson assembly

AmyL-PrtS fwd	CGGTAGCTGCAGATGAAGCAAATCTTAATGGGACG	Amplification of AmyL with floating tail for Gibson assembly
AmyL-MA-PrtS rev	GTAGTTGAGTCACCTGTTTACCACCTCTTTTGAACATAAATTGAAACCG	Amplification of AmyL with floating tail for Gibson assembly
PrtS dwfl-AmyL-MA fwd	CGGTTTCAATTTATGTTCAAAGAGGTGGTAAACAGGTGACTCAACTAC	Amplification of the down homologous arm PrtS with floating tail for Gibson assembly
AmyL internal fwd	TCGGTTGCAAATTCAGGTTTGG	Sequencing primer
AmyL internal rev	TCATAAAGGTCGTAAGCACCGTAG	Sequencing primer
***Streptococcus thermophilus ΔPrtS::Amyl*-STOP**		
AmyL-Stop-PrtS rev	GTGATAAAGCTAGTGTTGGTTACTATCTTTTGAACATAAATTGAAACCG	Amplification of AmyL with floating tail for Gibson assembly
PrtS dwfl-AmyL-STOP fwd	CGGTTTCAATTTATGTTCAAAGATAGTAACCAACACTAGCTTTATCAC	Amplification of the down homologous arm PrtS with floating tail for Gibson assembly
***Streptococcus thermophilus ΔPrtS::Amyl*-BIL**		
AmyL-BIL rev	GCCAGCACCACCTCTTTGAACATAAATTGAAACCG	Amplification of AmyL with floating tail for Gibson assembly
BIL-AmyL fwd	CGGTTTCAATTTATGTTCAAAGAGGTGGTGCTGGC	Amplification of BIL with floating tail for Gibson assembly
BIL-MA PrtS Rev	GTAGTTGAGTCACCTGTTTACCACCTAGAGCATAGCCTTCCG	Amplification of BIL with floating tail for Gibson assembly
PrtS dwfl-BIL-MA fwd	CGGAAGGCTATGCTCTAGGTGGTAAACAGGTGACTCAACTAC	Amplification of BIL with floating tail for Gibson assembly

Table 5. List of the primer used in this master thesis. Floating tails are highlighted in red.

5.2 Methods

5.2.1 DNA manipulations

Polymerase chain reaction (PCR)

PCR is an *in vitro* method that allows the amplification of a wanted DNA sequence from a template (genomic DNA, plasmid, linear DNA). The common reaction is a cycle divided in 3 general steps. The first step is the denaturation. The solution is heated at high temperature to separate the double strand DNA into single strands. The second step is the hybridization. In this step the temperature is decreasing to a determined temperature to allow the annealing of the primers. The temperature is determined by the length and the composition of the primer. The last step is the extension. In this step, a DNA polymerise will extent the primers in the 5'-> 3' sense. Regularly, we set up the reaction to 30 cycles.

Therefore, to perform a PCR, we mixed in a sterile PCR tube 10µl of the 5X Q5 buffer (NEB), 2,5µl of the forward and the reverse primer (stock concentration 10µM), 2µl of dNTP mixes (stock concentration 5mM), 0.5µl of template DNA, 0.5µl of Q5 polymerase (NEB) and 32µl of sterile DNase and RNase free water.

The PCR mixes were incubated in a thermocycler with the parameter shown in Table 6.

Step	Temperature (°C)	Cycles	Time (seconds)
Initial denaturation	98	1	30
Denaturation	98	30	10
Hybridization	Calculated for each primer		15
Elongation	72		30s/kb
Final Elongation	72	1	300
Conservation	16	1	∞

Table 6: Thermocycler parameter for PCR reaction

Gibson assembly

In a PCR tube we mixed 15µl of homemade Gibson mixes with 5µl containing a mix of each fragment (250 fmole for each fragment). The homemade Gibson mix was prepared by imitating the composition of the Gibson mixes from NEB (Tris-HCl pH 7.5 100mM, MgCl2 10mM, 4 dNTPs 0.2mM each, NAD+ 1mM, PEG-8000 15%, T5 exonuclease (2 U/mL), Phusion DNA polymerase (33 U/ mL), Taq DNA ligase (1666 U/mL)).
The assemblies were performed at 50°C during 30 min.

S. thermophilus genomic extraction

The genome of *S. thermophilus* was extracted from a culture using a Fastprep. 500µl of an overnight culture was centrifuged at 16 000g for 1 minute to pellet the cells. The supernatant was discarded and the cells were washed twice with TE buffer (Tris base 1mM, EDTA 0.1mM at pH 8) and mixed with 0.17-0.18mm plastic beads (Sartorius). The mix was put in a Fastprep at maximum intensity for 1 minute. The resulting solution was centrifuged at 16 000g to pellet the cell debris and the beads and the supernatant was recovered as genomic DNA.

DNA purification

The DNA fragment were purified using the Monarch PCR & DNA cleanup kit (5 µG) (NEB T10130). Following the manufacturer protocol, the samples were diluted with a binding buffer and charged into a purification column. After the first centrifugation, the DNA was bound to the column and we washed it twice by charging the column with awashing buffer containing ethanol. Once the sample was washed, we eluted the DNA with sterile DNase- and RNase-free water warmed up to 70°C.
The concentration of DNA in ng/µl was measured using a spectrophotometer (Nanodrop).

Agarose gel electrophoresis

Agarose gel electrophoresis is a method to verify the size of the DNA fragments amplified by PCR. To do so, a warm agarose solution (0.8%) was poured into a holder with a comb to make the wells. The solution was cooled at room temperature to polymerise the agarose and solidify the gel.
Afterwards, the gel was incubated in a migration tank filled with TAE 1X (Tris 40mM, acetic acid 20mM, EDTA 1Mm at pH 8.5). We charged 5µl of PCR mixed with 1µl of loading buffer 6X (NEB) in the wells. In order to determine the molecular weight of each we also charged a smart ladder in (Eurogentec) in the wells. We fixed the voltage at 120 V for small gel and 140 V for big gel and the samples migrated during approximatively 30min.
Once the migration was completed, the gel were incubated in a BET bath (concentration 0.5µg/ml) during 20min and the DNA bands were revealed under UV.

Gel band extraction

DNA was purified from agarose gel band using the Monarch® DNA Gel Extraction Kit (NEB #T1020). Following the manufacturer protocol, the agarose was melted in a dissolving buffer at 50°C for 15min. The resulting solution was charged into a purification column. After the first centrifugation, the DNA was bound to the column and we washed it twice by charging the column with a washing buffer containing ethanol. Once the sample was washed, we eluted the DNA with sterile DNase- and RNase-free water warmed up to 70°C.
The concentration of DNA in ng/µl was measured using a spectrophotometer (Nanodrop).

5.2.2 Bacterial manipulations

Induction of the natural competence of *S. thermophilus*

The transformation of *Streptococcus thermophilus* was performed using a protocol optimized in the lab.

Firstly, the strain that was going to be transformed was grown overnight in CDMG at 37°C in a sealed Eppendorf filled to the top to mimic anaerobic conditions. The next day, 50µl of the culture was reinoculated in 950µl of semi-skimmed pasteurized milk for 1h15.
Afterwards we added a small peptide called XIP (ComX inducing peptide) at a final concentration of 1µM. The cultures were splitted in two 500µl Eppendorf and one of them was mixed with the donor DNA ~1µg. The other was used as a negative control of the transformation.
Next, the two culture were incubated at 37°C for 3h30 and then dilutions were plated on solid media containing either chloramphenicol for the integration of the CAT-oroP cassette or 5'FOA for the substitution of the CAT-oroP cassette with the cassette of interest. The plates were incubated overnight at 37°C in an anaerobic chamber with GaspakTM EZ.
The next day, we selected transformants in dilutions where we could not see colonies for the negative control. The transformant were streaked on agar plates containing chloramphenicol or 5'FOA to disrupt the chain cells and recover colonies formed by single cells. The streaked plates were again incubated overnight in anaerobic conditions.
The final day, isolated colonies were selected and the integration of the cassette was checked by genomic extraction, PCR control and sequencing.

Anaerobic growth curves

In order to check for the fitness of the mutants, we realised anaerobic growth curves.
To do so, we used a microplate reader (TECAN) and measured the OD_{600nm} of triplicates of each mutant every 5 min.
Since *S. thermophilus* grows in anaerobic conditions, we used modified microplates to incubates the culture. As it can be seen in Figure 37, we inserted the powder used in the GaspakTM EZ system into the microplate and sealed it with airtight paste to prevent gas exchange between the plate and the air. The microplate was also sealed from the outside using gasproof tape. The anaerobic conditions into the plate was checked using an anaerobic indicator colored strip.

Figure 37: Anaerobic modified microplate: The plate was filled with powder from the GaspakTM EZ system and sealed using airtight paste and gasproof tape. The anaerobic conditions were verified using an anaerobic indicator strip.

Fluorescence microscopy experiments

The culture were prepared for fluorescence microscopy as following. Overnight cultures were reinoculated and incubated at 37°C in anaerobic conditions.
Once the culture reached an OD_{600nm} of ~0.6-0.7, the cells were pelleted and washed twice with 500µl of PBS (NaCl 137mM, KCl 2,7mM, KH_2PO_4 1.8mM, $Na_2HPO_4.H_2O$ 1mM) and resuspended in 50µl PBS.
Using a PBS solution supplemented with agarose (1.5%), we prepared glass slides with thin PBS matrices to drop the 1.5µl of the resuspended cells.

The microscope used was an Axio I inverted microscope (Zeiss) equipped with a Plan-Apochromat objective (100 ×/1.46 Oil DIC M27) (Zeiss), a HXP 120 C lighting unit (Zeiss) and a C10600 ORCA-R2 camera (Hamamatsu). We used the 100X objective which requires the application of immersoil on the cover of the glass slide.

The pictures were obtained by using the phase contrast mode and the eGFP filter. The exposure time for the phase filter was 70ms and for the eGFP filter it was 300ms. The image were processed using MicrobJ, a plugin of ImageJ.

Flow cytometry experiments

The cultures were prepared for flow cytometry as following. Overnight cultures were reinoculated and incubated at 37°C in anaerobic conditions.
Once the cultures reached an OD_{600nm} of ~0.6-0.7, the cells were pelleted and resuspended in 1ml of PBS. *E. coli* cultures were diluted 1:4000 and *S. thermo* cultures were diluted 1:2000.

The flow cytometer used was an Guava® easyCyte 5 HPL Benchtop Flow Cytometer. The instrument was calibrated and washed following the manufacturer protocol and the GFP channel was calibrated for the detection of fluorescence with an *E. coli* GFP strain.

The data collected with the flow cytometer were analysed using Floreada.io an online resource (https://floreada.io)

Amylase activity assay

The cultures were prepared for the amylase assay with the Sigma-Aldrich© Kit MAK009 as following. Overnight cultures were reinoculated and incubated at 37°C in anaerobic conditions.
Once the cultures reached an OD_{600nm} of ~0.7, the culture were centrifugated at 16 000g for 5min.

a) Preparation of the supernatant sample

500µl of the clear supernatant was transferred into an Amicon Ultra-0.5 Centrifugal Filter Devices with 30 KDa cutoff. The Amicon was centrifugated for 15 minutes at 14,000 g. The residual volume in the Amicon was washed with 500µl of fresh PBS and the resulting concentrate was used as a sample for the assay (~20 µl)

b) Preparation of the cells sample

The rest of the supernatant was discarded and the cells were washed twice with 500µl of PBS. The cells were finally resuspended in 200µl of PBS and, 20µl of the resuspension was used as a sample for the assay.

c) DTT treatment

For the assay with the DTT treatment, we prepared the supernatant and the cells sample as following.
First, as before, 500µl of the supernatant was transferred into an Amicon Ultra-0.5 Centrifugal Filter Devices with 30 KDa cutoff. The Amicon was centrifugated for 15 minutes at 14000g.
The cells were resuspended in 1ml of PBS supplemented with DTT (10mM) and incubated at room temperature for 20 minutes.
Afterwards, the cells were pelleted by centrifugation and 500µl of the supernatant was also transferred into the Amicon. The Amicon was again centrifugated for 15 minutes at 14000g. The other 500µl of the supernatant was discarded.
The pelleted cells were again washed in 500µl of PBS and re-peletted by centrifugation at 16000g for 5min.
250µl of the supernatant mixed with 250µl of fresh PBS was used to wash the concentrate in the Amicon and the residual volume was used as the supernatant sample for the DTT treatment (~20µl).
The rest of the supernatant was discarded and the pelleted cells were resuspended in 200µl of fresh PBS. 20µl of the resuspended cells was used for the assay.

d) Amylase assay

Using the samples prepared with the protocol above, the assay was performed as indicated by the manufacturer, with the exception that we added 3µl of spectinomycin in the assay to avoid any potential growth or contamination.

The α-amylase activity was calculated as indicated by the manufacturer's protocol by calculating the amount of paranitrophenol generated. The quantity of paranitrophenol produced was evaluated by measuring the $OD_{(405nm)}$ of the sample each 5 min. Only the datapoints within the linear portion of the sampling curve, occurring strictly inside the first 3 hours, were considered, in a way that guaranteed $R^2 > 0.99$ for each sample. In other words, datapoints outside the linear range were simply discarded due to bubbles, pre-steady state, condensation, detector saturation, substrate depletion,…

6. Bibliography

a-agglutinin | SGD (no date). Available at: https://www.yeastgenome.org/complex/CPX-1838 (Accessed: May 19, 2020).

Arnold, F. H. (2018) "Directed Evolution: Bringing New Chemistry to Life," *Angewandte Chemie - International Edition*, 57(16), pp. 4143–4148. doi: 10.1002/anie.201708408.

Barderas, R. and Benito-Peña, E. (2019) "The 2018 Nobel Prize in Chemistry: phage display of peptides and antibodies," *Analytical and Bioanalytical Chemistry*, 411(12), pp. 2475–2479. doi: 10.1007/s00216-019-01714-4.

van Bloois, E. *et al.* (2011) "Decorating microbes: Surface display of proteins on Escherichia coli," *Trends in Biotechnology*. Elsevier, pp. 79–86. doi: 10.1016/j.tibtech.2010.11.003.

Boder, E. T. and Wittrup, K. D. (1997) "Yeast surface display for screening combinatorial polypeptide libraries," *Nature Biotechnology*, 15(6), pp. 553–557. doi: 10.1038/nbt0697-553.

Bolotin, A. *et al.* (2004) "Complete sequence and comparative genome analysis of the dairy bacterium Streptococcus thermophilus," *Nature Biotechnology*, 22(12), pp. 1554–1558. doi: 10.1038/nbt1034.

Cherf, G. M. and Cochran, J. R. (2015) "Applications of yeast surface display for protein engineering," *Methods in Molecular Biology*, 1319, pp. 155–175. doi: 10.1007/978-1-4939-2748-7_8.

Daugherty, P. S. (2007) "Protein engineering with bacterial display," *Current Opinion in Structural Biology*. Elsevier Current Trends, pp. 474–480. doi: 10.1016/j.sbi.2007.07.004.

Denks, K. *et al.* (2014) "The Sec translocon mediated protein transport in prokaryotes and eukaryotes," *Molecular Membrane Biology*, 31(2–3), pp. 58–84. doi: 10.3109/09687688.2014.907455.

Fernandez-Espla, M. D. *et al.* (2000) "Streptococcus thermophilus cell wall-anchored proteinase: Release, purification, and biochemical and genetic characterization," *Applied and Environmental Microbiology*, 66(11), pp. 4772–4778. doi: 10.1128/AEM.66.11.4772-4778.2000.

Fontaine, L. *et al.* (2010) "Development of a versatile procedure based on natural transformation for marker-free targeted genetic modification in Streptococcus thermophilus," *Applied and Environmental Microbiology*, 76(23), pp. 7870–7877. doi: 10.1128/AEM.01671-10.

Fontaine, L. *et al.* (2013) "Mechanism of competence activation by the ComRS signalling system in streptococci," *Molecular Microbiology*, 87(6), pp. 1113–1132. doi: 10.1111/mmi.12157.

Fontaine, L. *et al.* (2015) "Regulation of competence for natural transformation in streptococci," *Infection, Genetics and Evolution*, 33, pp. 343–360. doi: 10.1016/j.meegid.2014.09.010.

Francisco, J. A., Earhart, C. F. and Georgiou, G. (1992) "Transport and anchoring of β-lactamase to the external surface of Escherichia coli," *Proceedings of the National Academy of Sciences of the United States of America*, 89(7), pp. 2713–2717. doi: 10.1073/pnas.89.7.2713.

Freudl, R. (2018) "Signal peptides for recombinant protein secretion in bacterial expression systems," *Microbial Cell Factories*. BioMed Central Ltd., p. 52. doi: 10.1186/s12934-018-0901-3.

Gardan, R. *et al.* (2009) “The oligopeptide transport system is essential for the development of natural competence in Streptococcus thermophilus strain LMD-9,” *Journal of Bacteriology*, 191(14), pp. 4647–4655. doi: 10.1128/JB.00257-09.

Giger, L. *et al.* (2013) “Evolution of a designed retro-aldolase leads to complete active site remodeling,” *Nature chemical biology*, 9(8), p. 494. doi: 10.1038/NCHEMBIO.1276.

Goh, Y. J. *et al.* (2011) “Specialized adaptation of a lactic acid bacterium to the milk environment: The comparative genomics of Streptococcus thermophilus LMD-9,” *Microbial Cell Factories*, 10(SUPPL. 1), p. S22. doi: 10.1186/1475-2859-10-S1-S22.

Gregorio, N. E., Levine, M. Z. and Oza, J. P. (2019) “A User’s Guide to Cell-Free Protein Synthesis,” *Methods and Protocols*, 2(1), p. 24. doi: 10.3390/mps2010024.

Gunneriusson, E. *et al.* (1996) “Surface display of a functional single-chain Fv antibody on staphylococci,” *Journal of Bacteriology*, 178(5), pp. 1341–1346. doi: 10.1128/jb.178.5.1341-1346.1996.

Hanes, J. and Plückthun, A. (1997) “In vitro selection and evolution of functional proteins by using ribosome display,” *Proceedings of the National Academy of Sciences of the United States of America*, 94(10), pp. 4937–4942. doi: 10.1073/pnas.94.10.4937.

Håvarstein, L. S., Coomaraswamy, G. and Morrison, D. A. (1995) “An unmodified heptadecapeptide pheromone induces competence for genetic transformation in Streptococcus pneumoniae,” *Proceedings of the National Academy of Sciences of the United States of America*, 92(24), pp. 11140–11144. doi: 10.1073/pnas.92.24.11140.

He, M. and Taussig, M. J. (2002) “Ribosome display: cell-free protein display technology.,” *Briefings in functional genomics & proteomics*, 1(2), pp. 204–12. doi: 10.1093/bfgp/1.2.204.

Heinz, K. and Kandler, O. (1972) *Peptidoglycan .Types of Bacterial Cell Walls and their Taxonomic Implications*, *BACTERIOLOGICAL REVIEWS*.

Hols, P. *et al.* (2005) “New insights in the molecular biology and physiology of Streptococcus thermophilus revealed by comparative genomics,” *FEMS Microbiology Reviews*, 29(3), pp. 435–463. doi: 10.1016/j.fmrre.2005.04.008.

Iannolo, G. *et al.* (1995) “Modifying filamentous phage capsid: Limits in the size of the major capsid protein,” *Journal of Molecular Biology*, 248(4), pp. 835–844. doi: 10.1006/jmbi.1995.0264.

Jacobitz, A. W. *et al.* (2017) “Sortase Transpeptidases: Structural Biology and Catalytic Mechanism,” in *Advances in Protein Chemistry and Structural Biology*. Academic Press Inc., pp. 223–264. doi: 10.1016/bs.apcsb.2017.04.008.

Jahan-Tigh, R. R. *et al.* (2012) “Flow Cytometry,” *Journal of Investigative Dermatology*, p. 1. doi: 10.1038/jid.2012.282.

Josephson, K., Ricardo, A. and Szostak, J. W. (2014) “MRNA display: From basic principles to macrocycle drug discovery,” *Drug Discovery Today*. Elsevier Ltd, pp. 388–399. doi: 10.1016/j.drudis.2013.10.011.

Kanamori, T., Fujino, Y. and Ueda, T. (2014) “PURE ribosome display and its application in antibody technology,” *Biochimica et Biophysica Acta - Proteins and Proteomics*. Elsevier B.V., pp. 1925–1932. doi: 10.1016/j.bbapap.2014.04.007.

Kehoe, J. W. and Kay, B. K. (2005) “Filamentous phage display in the new millennium,” *Chemical Reviews*. American Chemical Society, pp. 4056–4072. doi: 10.1021/cr000261r.

Kondo, A. and Ueda, M. (2004) “Yeast cell-surface display - Applications of molecular display,” *Applied Microbiology and Biotechnology*. Springer, pp. 28–40. doi: 10.1007/s00253-003-1492-3.

Kronqvist, N. *et al.* (2010) “Staphylococcal Surface Display in Combinatorial Protein Engineering and Epitope Mapping of Antibodies,” *Recent Patents on Biotechnology*, 4(3), pp. 171–182. doi: 10.2174/187220810793611536.

Lee, S. Y., Choi, J. H. and Xu, Z. (2003) “Microbial cell-surface display,” *Trends in Biotechnology*. Elsevier, pp. 45–52. doi: 10.1016/S0167-7799(02)00006-9.

Li, R. *et al.* (2019) “Ribosome Display: A Potent Display Technology used for Selecting and Evolving Specific Binders with Desired Properties,” *Molecular Biotechnology*. Humana Press Inc., pp. 60–71. doi: 10.1007/s12033-018-0133-0.

Linciano, S. *et al.* (2019) “Molecular evolution of peptides by yeast surface display technology,” *MedChemComm*. Royal Society of Chemistry, pp. 1569–1580. doi: 10.1039/c9md00252a.

Löfblom, J. (2011) “Bacterial display in combinatorial protein engineering,” *Biotechnology Journal*, 6(9), pp. 1115–1129. doi: 10.1002/biot.201100129.

Marvin, D. (1998) “Filamentous phage structure, infection and assembly,” *Current Opinion in Structural Biology*, 8(2), pp. 150–158. doi: 10.1016/S0959-440X(98)80032-8.

Mattheakis, L. C., Bhatt, R. R. and Dower, W. J. (1994) “An in vitro polysome display system for identifying ligands from very large peptide libraries,” *Proceedings of the National Academy of Sciences of the United States of America*, 91(19), pp. 9022–9026. doi: 10.1073/pnas.91.19.9022.

Mazmanian, S. K. *et al.* (1999) “Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall,” *Science*, 285(5428), pp. 760–763. doi: 10.1126/science.285.5428.760.

McCafferty, J. *et al.* (1990) “Phage antibodies: filamentous phage displaying antibody variable domains,” *Nature*, 348(6301), pp. 552–554. doi: 10.1038/348552a0.

Nagano, K. and Tsutsumi, Y. (2021) “Phage Display Technology as a Powerful Platform for Antibody Drug Discovery,” *Viruses*, 13(2). doi: 10.3390/V13020178.

Natale, P., Brüser, T. and Driessen, A. J. M. (2008) “Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-Distinct translocases and mechanisms,” *Biochimica et Biophysica Acta - Biomembranes*. Elsevier, pp. 1735–1756. doi: 10.1016/j.bbamem.2007.07.015.

Nixon, A. E., Sexton, D. J. and Ladner, R. C. (2014) “Drugs derived from phage display from candidate identification to clinical practice,” *mAbs*, 6(1), pp. 73–85. doi: 10.4161/mabs.27240.

Packer, M. S. and Liu, D. R. (2015) “Methods for the directed evolution of proteins,” *Nature Reviews Genetics*. Nature Publishing Group, pp. 379–394. doi: 10.1038/nrg3927.

Palmer, T. and Berks, B. C. (2012) “The twin-arginine translocation (Tat) protein export pathway,” *Nature Reviews Microbiology*. Nat Rev Microbiol, pp. 483–496. doi: 10.1038/nrmicro2814.

Porter, J. L., Rusli, R. A. and Ollis, D. L. (2016) “Directed Evolution of Enzymes for Industrial Biocatalysis,” *ChemBioChem*. Wiley-VCH Verlag, pp. 197–203. doi: 10.1002/cbic.201500280.

Press release: The Nobel Prize in Chemistry 2018 - NobelPrize.org (no date). Available at: https://www.nobelprize.org/prizes/chemistry/2018/press-release/ (Accessed: April 21, 2020).

Puromycin | C22H29N7O5 - PubChem (no date). Available at: https://pubchem.ncbi.nlm.nih.gov/compound/puromycin (Accessed: May 16, 2020).

Rakonjac, J. *et al.* (2011) "Filamentous bacteriophage: biology, phage display and nanotechnology applications.," *Current issues in molecular biology*, pp. 51–76. doi: 10.21775/cimb.013.051.

RCSB PDB - 1T2P: Crystal structure of Sortase A from Staphylococcus aureus (no date). Available at: https://www.rcsb.org/structure/1t2p (Accessed: May 27, 2020).

Rohde, M. (2019) "The Gram-Positive Bacterial Cell Wall," *Microbiology Spectrum*, 7(3). doi: 10.1128/microbiolspec.gpp3-0044-2018.

Rusch, S. L. and Kendall, D. A. (2007) "Interactions that drive Sec-dependent bacterial protein transport," *Biochemistry*. NIH Public Access, pp. 9665–9673. doi: 10.1021/bi7010064.

Samuelson, P. *et al.* (1995) "Cell surface display of recombinant proteins on Staphylococcus carnosus," *Journal of Bacteriology*, 177(6), pp. 1470–1476. doi: 10.1128/jb.177.6.1470-1476.1995.

Shanker, E. *et al.* (2016) "Pheromone Recognition and Selectivity by ComR Proteins among Streptococcus Species," *PLOS Pathogens*. Edited by G. Zhang, 12(12), p. e1005979. doi: 10.1371/journal.ppat.1005979.

Sheehan, J. and Marasco, W. A. (2015) "Phage and Yeast Display," *Microbiology Spectrum*, 3(1). doi: 10.1128/microbiolspec.aid-0028-2014.

Shimizu, Y. *et al.* (2001) "Cell-free translation reconstituted with purified components," *Nature Biotechnology*, 19(8), pp. 751–755. doi: 10.1038/90802.

Smith, G. P. (1985) "Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface," *Science*, 228(4705), pp. 1315–1317. doi: 10.1126/science.4001944.

Takahashi, T. T., Austin, R. J. and Roberts, R. W. (2003) "mRNA display: Ligand discovery, interaction analysis and beyond," *Trends in Biochemical Sciences*. Elsevier Ltd, pp. 159–165. doi: 10.1016/S0968-0004(03)00036-7.

Ton-That, H., Marraffini, L. A. and Schneewind, O. (2004) "Protein sorting to the cell wall envelope of Gram-positive bacteria," *Biochimica et Biophysica Acta - Molecular Cell Research*. Elsevier, pp. 269–278. doi: 10.1016/j.bbamcr.2004.04.014.

Tsirigotaki, A. *et al.* (2017) "Protein export through the bacterial Sec pathway," *Nature Reviews Microbiology*. Nature Publishing Group, pp. 21–36. doi: 10.1038/nrmicro.2016.161.

Vollmer, W., Blanot, D. and de Pedro, M. A. (2007) "Peptidoglycan structure and architecture." doi: 10.1111/j.1574-6976.2007.00094.x.

Wang, H. and Liu, R. (2011) "Advantages of mRNA display selections over other selection techniques for investigation of protein-protein interactions," *Expert Review of Proteomics*. Taylor & Francis, pp. 335–346. doi: 10.1586/epr.11.15.

Wentzel, A. *et al.* (2001) "Display of passenger proteins on the surface of Escherichia coli K-12 by the enterohemorrhagic E. coli intimin EaeA," *Journal of Bacteriology*, 183(24), pp. 7273–7284. doi: 10.1128/JB.183.24.7273-7284.2001.

Zhao, H. *et al.* (2001) “Interaction of α-agglutinin and a-agglutinin, Saccharomyces cerevisiae sexual cell adhesion molecules,” *Journal of Bacteriology*, 183(9), pp. 2874–2880. doi: 10.1128/JB.183.9.2874-2880.2001.

Zhao, X. L. *et al.* (2009) “Selection and affinity maturation of human antibodies against rabies virus from a scFv gene library using ribosome display,” *Journal of Biotechnology*, 144(4), pp. 253–258. doi: 10.1016/j.jbiotec.2009.09.022.

Zong, Y. *et al.* (2004) “Crystal structures of Staphylococcus aureus Sortase A and its substrate complex,” *Journal of Biological Chemistry*, 279(30), pp. 31383–31389. doi: 10.1074/jbc.M401374200.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des sciences
Place des sciences, 2 bte L6.06.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/sc**