UNIVERSIDADE DE LISBOA

FACULDADE DE FARMÁCIA

DEPARTAMENTO DE MICROBIOLOGIA E IMUNOLOGIA

Construction of Ms6 derivative mutants as tools for interaction studies with mycobacteria

Sofia Pires de Oliveira Pombo

Dissertação de Mestrado

MESTRADO EM CIÊNCIAS BIOFARMACÊUTICAS

2014

UNIVERSIDADE DE LISBOA

FACULDADE DE FARMÁCIA

DEPARTAMENTO DE MICROBIOLOGIA E IMUNOLOGIA

Construction of Ms6 derivative mutants as tools for interaction studies with mycobacteria

Sofia Pires de Oliveira Pombo

Dissertação de Mestrado
Orientada pela Professora Doutora Madalena Pimentel

MESTRADO EM CIÊNCIAS BIOFARMACÊUTICAS

2014

Agradecimentos

Os meus primeiros votos de gratidão são dirigidos à Professora Madalena Pimentel, sem a qual a realização desta dissertação não seria de todo a mesma. Agradeço-lhe ter aceitado ser minha orientadora, e por ter apostado em mim, sempre depositando confiança nas minhas capacidades, o que decerto foi e sempre será um estímulo. Muito obrigado pelos ensinamentos que me transmitiu, e que contribuíram para fomentar ainda mais o meu gosto pela investigação! Agradeço-lhe mais uma vez, toda a atenção disponibilizada em tudo, e em particular na correcção desta dissertação.

Gostaria também de agradecer ao Professor José Moniz-Pereira, pelo entusiamo e interesse que mostrou durante todo o meu percurso do estágio laboratorial.

A todos os membros do laboratório do Professor Carlos São-José, que são como colegas do mesmo laboratório, ao Adriano, ao Francisco e à Vanessa: Muito Obrigado por tudo! Este agradecimento engloba obviamente toda a experiência e ajuda prestada no laboratório, mas também todos os bons momentos. Obrigado à Sofia pelas longas conversas. E um especial obrigado ao Francisco e ao Adriano, colegas desde o primeiro dia!

Aos membros laboratório do Professor João Gonçalves, e em particular à Sofia, pela partilha de bons momentos e companhia de lanches.

À Juliana, pelas correcções do Inglês: obrigado!

Ao Luís, o meu Muito Obrigado por tudo! Pelo apoio e atenção, amizade, carinho, pelas críticas construtivas e claro, pela paciência! Obrigado pelas correcções e formatações da tese!

Aos Meus Pais e irmã, o meu Muito Obrigado! Por todo o apoio que me dão nas decisões que tomo, pelo incentivo em todos os projectos, pela confiança, amizade e por estarem sempre presentes.

Por último, tenho de agradecer à Sílvia, à minha Salvadora de última hora! Muito obrigado pela ajuda, e pela amizade demonstrada ao longo destes 2 anos!

Abstract

Mycobacteriophage Ms6 is a temperate phage isolated in 1989 from a spontaneously induced culture of *Mycobacterium smegmatis*. Although several studies regarding lysis and integration have been published, the host range of Ms6 is not known. In this work we observed that incubation of this phage with fast- and slow-growing mycobacteria did not result in plaque or halos formation, as usually happens in infection plaque assay. However, an absence of clearing in a bacterial lawn does not mean that Ms6 in unable to infect the tested mycobacteria apart from its natural host, *M. smegmatis*. Taking this into consideration, our main goal was to construct Ms6 derivative mutants with reporter genes, and use them as tools to follow the mycobacteriophage replication inside other mycobacteria.

Using the Bacteriophage Recombineering of Electroporated DNA (BRED) technique, we successfully constructed two Ms6 derivative reporter phages, Ms6*lysB*::*egfp* and Ms6*pin*::gfp_m^{2+}, by gene replacement. We successfully showed that their reporter proteins, improved green fluorescent proteins (GFP), are respectively expressed in Ms6 lytic and lysogenic life cycle, on *M. smegmatis*. Consequently, these new mutant phages are very important and useful tools to be used in future studies since they permit to follow a lytic or lysogenic cycle. With Ms6*lysB*::*egfp* and Ms6*pin*::gfp_m^{2+}, we expect to determine with more accuracy the ability of Ms6 to infect/replicate in other mycobacteria.

Key words: Recombineering; Mycobacteriophage Ms6; Ms6*lysB*::*egfp;* Ms6*pin*::gfp_m^{2+}*;* Mycobacteria; GFP.

Resumo

O micobacteriófago Ms6 é um fago temperado, isolado em 1989 a partir de uma cultura de *Mycobacterium smegmatis* induzida espontaneamente. Apesar de vários estudos sobre o sistema de lise e integração terem já sido publicados, a gama de hospedeiros do Ms6 não é ainda conhecida. Neste trabalho verificou-se que a incubação deste fago com micobactérias quer de crescimento rápido quer de crescimento lento, não resultou na formação de placas fágicas nem de halos, tal como acontece num ensaio clássico de infecção em placa. No entanto, a ausência de placas fágicas não significa que Ms6 não tenha capacidade de infectar outras micobactérias para além de *M. smegmatis*, o seu hospedeiro natural. Tendo este facto em consideração, o principal objectivo deste trabalho foi construir fagos Ms6 mutantes com genes repórter, de modo a usá-los como ferramentas para testar a capacidade do Ms6 se replicar noutras micobactérias.

Através da metodologia *Bacteriophage Recombineering of Electroporated DNA* (BRED), foi possível proceder a uma troca de genes do fago Ms6 por genes repórter e construir, com sucesso, dois fagos recombinantes designados Ms6*lysB*::*egfp* e Ms6*pin*::gfp_m^{2+}, capazes de produzir fluorescência quando infectam um hospedeiro sensível. Foi com sucesso que mostrámos que as proteínas verde fluorescentes modificadas (GFP), são expressas em *M. smegmatis* no ciclo lítico e lisogénico. Consequentemente, estes novos fagos mutantes constituem importantes ferramentas que poderão ser usados em estudos futuros, tendo em conta que permitem seguir o ciclo lítico e lisogénico.

Esperamos assim, através dos fagos construídos Ms6*lysB*::*egfp* and Ms6*pin*::gfp_m^{2+}, conseguir determinar de uma forma mais precisa, a capacidade de Ms6 infectar/replicar-se em outras micobactérias.

Palavras-chave: Recombinação; Micobacteriófago Ms6; Ms6*lysB*::*egfp*; Ms6*pin*::gfp_m^{2+}; Micobactérias; GFP.

Table of contents

Abbreviations

aa	Amino acid
Ala	Alanine
ATCC	American Type Culture Collection
BCG	Bacillus Calmette-Guérin
bp	Base pair
BRED	Bacteriophage Recombineering of Electroporated DNA
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DNA	Deoxyribonucleic Acid
ds	Double stranded
EGFP	Enhanced Green Fluorescent Protein
Fig	Figure
Fw	Forward
GFP	Green Fluorescent Protein
Hyg	Hygromycin
IPP	Institute Pasteur Protection
Kan	Kanamycin
LB	Luria-Bertani broth
LPS	Lipopolysaccharides
mAGP	mycolyl arabinogalactan-peptidoglygan
Met	Methionine
MOI	Multiplicity of infection
nt	nucleotide
OADC	Oleic Albumin Dextrose Catalase
O.D.	Optical density
o.n.	Over-night
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
pfu	Plaque forming units
Pr	primer
RNA	Ribonucleic Acid
rpm	Revolutions per minute
Rv	Reverse
ss	Single stranded
TBE	Tris-Borate-EDTA
tRNA	Transfer RNA
UV	Ultraviolet
wt	Wild type

I. Introduction

1. Bacteriophage's main features

Bacteriophages or simply phages, from the Greek "eaters of bacteria", are viruses that infect bacteria. They were discovered twice, by Frederick William Twort in 1915 who described a glassy transformation of "micrococcus", and by Félix D'Herelle in 1917 who observed the destruction of *Shigella* in bacteria cultures. While Twort abandoned his discovery, D'Herelle recognized the potential of this phenomenon, and dedicated his scientific life to bacteriophages and their use in phage therapy (Ackermann, 2011).

As obligate parasites, they use the bacterial cell resources to replicate (Monk *et al.*, 2010) but they also exist in the extracellular environment as supramolecular structures named virions. Bacteriophages are found everywhere in the biosphere, even in volcanic hot springs; however, their main habitats are oceans and soil (Ackermann, 2011).

Estimated in more than 10^{31} on earth, phages are, until now, the most abundant entity on earth and are major agents of microbial evolution (Willey *et al.*, 2008). Viral ecologists calculate that there are about 10^{23} phage infections per second on a global scale, indicating that their population is not only large but also highly dynamic (Hatfull and Hendrix, 2011).

As wide diverse, bacteriophage families are taxonomically organized on the basis of their morphology and nucleic acid structure. There is no universal criteria for either genus or species. The International Committee on Taxonomy of Viruses (ICTV) currently organizes bacteriophages in one named order and 10 families (Ackermann, 2011).

Their genomes can be single or double-stranded, linear or circular, DNA or RNA molecules. With exception of cystoviruses with three dsRNA molecules, phage genomes are single molecules with sizes ranging approximately from 3,5kb encoding 3 or 4 gene products to, in some cases, genomes of several hundred kilobases encoding hundreds of gene products (Willey *et al.*, 2008; Maniloff, 2012).

Most bacteriophages (96%) have a tailed morphology and constitute the order *Caudovirales*. This includes three families which are characterized according to the

morphological features of the tail: *Myoviridae* (contractile tail), *Siphoviridae* (non-contractile long) and *Podoviridae* (short tail). Tailed phages have dsDNA molecules encapsidated in a protein coat (capsid) in the shape of an icosahedron or have elongated heads. At the tail end, these phages can have a base plate, spikes or tail fibers, with specific sites on their tips, that participate in the recognition and attachment to the cell surface of the host (Ackermann, 2006; 2009; 2011; Maniloff, 2012). Polyhedral, filamentous or pleomorphic bacteriophages comprise the others 4%. All their families are small, and are extremely diversified by their basic properties (Ackermann, 2011).

According to the type of infection, bacteriophages can be classified as temperate or virulent. Temperate phages can produce a lysogenic or lytic infection (Fig. 1), unlike virulent phages that are only able to produce lytic infections. In the lytic life cycle, the bacteriophage takes over the host cell biosynthetic machinery for viral nucleic acid replication and proteins synthesis. After encapsidation and production of a certain number of new phages, they accumulate in the cytoplasm until the right moment for their release, which is when they cause lysis of host cell. Once in the environment, the newly formed phages are able to infect new host cells (Hatfull, 2000; Willey *et al.*, 2008; Maniloff, 2012). Lysogenic life cycle happens in certain physiological situations, when some phages integrate their genome (prophage) into the host's genome, maintaining a stable relationship with the host cell and propagating their genetic material to the host daughter cells during cell division. Bacteria hosting a prophage are said lysogenic. Under some stress conditions, the prophage can be excised from the bacteria genome and "switch" to the lytic cycle. (Willey *et al.*, 2008; Maniloff, 2001).

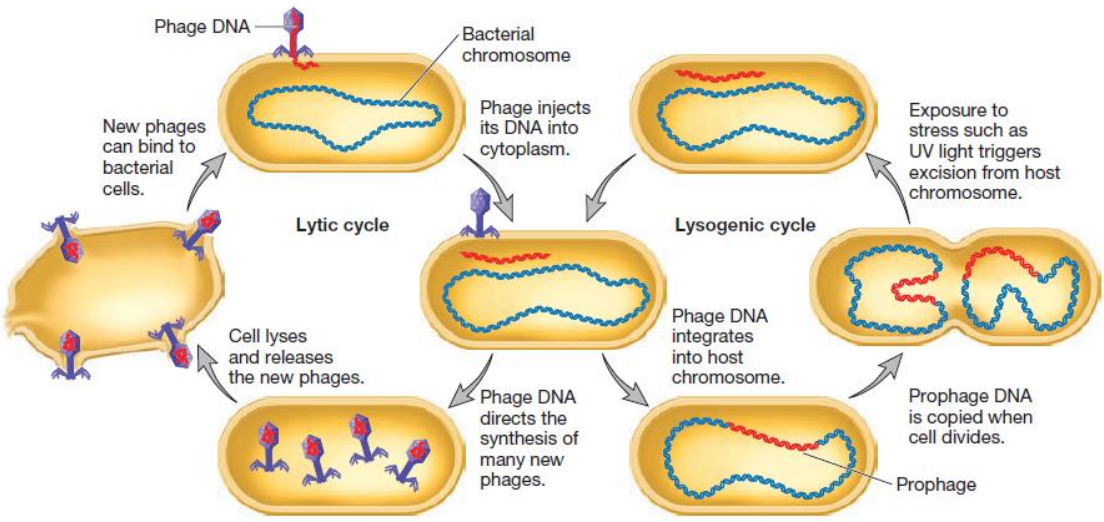

Figure 1 - Lytic and Lysogenic cycles of temperate phages (Adapted from: Willey *et al.*, 2008).

The expected result of bacterial infections is visible lesions or clear areas in a bacteria lawn called plaques. Plaques result from the initial infection of a bacterial cell present in the lawn by a single viable phage, which multiplies within the infected bacteria and lyses releasing the newly formed phages into the medium, leading to subsequent rounds of infection and lysis of nearby host cells (Fig. 2) (Ellis and Delbrück, 1939; Maniloff, 2012). Furthermore, depending on plaque morphologies, it is sometimes possible to distinguish between temperate or lytic phage infection. Temperate phages infection result in turbid plaques due to lysogens growth within those plaques, contrarily to lytic phages in which all of the infected cells lyse, and consequently plaque formation is characterized by clear areas. However, phages sometimes exhibit an intermediate morphology, in which a phage's type of infection is not obvious, and, thus, further examination is necessary (Hatfull, 2000).

Independently of the followed replication cycle, since phages do not have their own metabolism, they are obligate intracellular parasites. Therefore, a virus life cycle will only be successful if all infection steps are productively completed. These include the essential first steps of attachment (or adsorption) and penetration (or injection). Attachment begins with random collisions between the phage and the bacterial cell, leading to specific interactions between the virion and the receptors. Penetration or injection phase occurs when the viral genome enters into the host cell (Maniloff, 2012).

Figure 2 - Bacteriophage plaques formed by Ms6 on a lawn of *M. smegmatis* cells.

2. Bacteriophage receptors in Gram-positive and Gram-negative bacteria

Host recognition by a bacteriophage, followed by attachment to the cell surface, and phage genome penetration into host cell, are the key events for a successful infection. For some tailed phages, the recognition of a bacterial receptor involves a unique central fiber, while to others a cluster of fibers associated with the tail (Kutter *et al.*, 2005).

In general, bacteriophages attach to host cells using a two-step mechanism. In the first step, attachment is specific but reversible, which allows the phage to diffuse two-dimensionally over the bacterial surface until the second receptor is encountered. This second receptor is then bound irreversibly and, at this point, the phage is immobilised and committed to DNA transfer (Garcia-Doval and van Raaij, 2013).

Bacterial receptors are cellular constituents representing different biochemical families, (Barsom *et al*., 2008; Willey *et al*., 2008; Maniloff, 2012) which are eventually different in Gram-positive and Gram-negative bacteria, because these bacteria types differ in their cell structures. Both cell types have a cell wall made up of peptidoglycan and a phospholipid bilayer with membrane proteins (cell membrane). Gram-positive cell wall has a thicker peptidoglycan layer which contain large quantities of teichoic acids; whereas Gram-negative bacteria have a perisplamic space between cell wall and the cell membrane, and also an outer membrane with lipopolysaccharides (LPS), lipoprotein, and porin channels (Willey *et al*., 2008). Therefore, Gram-positive bacteria potential receptors are peptidoglycan elements, teichoic acids, lipoteichoic acids and their associated proteins, whereas in Gram-negative bacteria are porins, lipopolysaccharides and transport proteins (Kutter *et al.*, 2005). However there is little information available about phage receptors' structure in Gram-positive bacteria when compared to Gram-negative. Glucosylated polycerol phosphate is the major and essential teichoic acid in the Gram-positive *Bacillus subtilis* and was reported to be a receptor for several phages including SP01 (Young, 1967). In addition to this, the proteins GamR (*Bacillus anthracis*), Pip (*Lactococcus lactis*) and its orthologue YueB (*Bacillus subtilis*) were also identified as phage receptors (São-José *et al.*, 2004; Kutter *et al.*, 2005). YueB, was the first active virus receptor to be purified, it is a membrane protein from *Bacillus subtilis* and is also a receptor for irreversible adsorption of bacteriophage SPP1 (São-José *et al.*, 2006).

Porins are protein complexes composed of 3 subunits that form a channel in bacteria outer membrane and are used as receptors for phages that infect Gram-negative bacteria. Major proteins of this type in *Escherichia coli* cells are OmpC and OmpF. OmpC is the receptor for phages Hy2, ss4 Tulb and T4 and OmpF is the receptor for phage T2 (Kutter *et al.*, 2005; Rakhuba

et al., 2010). Lipopolysaccharides (LPS) are another constituent of the outer membrane in Gram-negative bacteria, which may also serve as bacteriophage receptors and can be designated as smooth- (S) or rough-type (R). Some bacteriophages might attach to both LPS types, however, phages specific to S-type LPS, display narrow host range specificity determined by large variability of O-antigen structure in bacteria of different taxonomic groups. Bacteriophages recognizing R-type show a broader host range since the structure of LPS core is rather conservative in various species and genera of Gram-negative bacteria. It is well known that receptors for T-phages, specifically T3, T4 and T7 are components of R-type LPS (Rakhuba *et al*., 2010). The receptor for phage lambda in *Escherichia coli* K-12 is LamB an outer membrane protein involved in maltose uptake (Randall-Hazelbauer and Schwartz, 1973). These examples are representative of the variety of phage receptors in Gram-negative bacteria. In addition, there are even some tailed phages that have complex adhesins and can recognize both LPS and membrane proteins (Kutter *et al.*, 2005).

In case of a bacteriophage being able to attach, insert its genome and productively infect a specific bacteria, it means that this bacteria is sensitive to that phage, and thus, it will be part of its host-range (Hyman and Abedon, 2010). Some phages have a narrow host range whereas for others it may be broad. The receptor location on the cell surface, their amount and density, are possible determinants for host range specificity (Rakhuba *et al.*, 2010). In addition, bacterial resistance mechanisms have an important and decisive role in a host-range determination.

3. Bacterial Resistance to Phage Infection

Some bacteria have evolved or may evolve to prevent phage attachment. This resistance may include the presence of physical barriers hiding receptor molecules as well as loss of phage receptor molecules on their hosts, also known as adsorption resistance. In addition there are other mechanisms, which interfere with phage infections, such as prevention of host takeover and abortive infection. In a general point of view, phage-resistance mechanisms serve to limit phage host-range, and those mechanisms are described as “Bacteriophage Resistome”. Despite all, phages in response, have evolved to avoid some of those mechanisms (Hyman and Abedon, 2010).

3.1. Adsorption resistance

Adsorption resistance interferes with the interaction between phage and the host receptor, reducing it. Bacterial receptors for phage adsorptions are molecules with essential roles for the cell maintenance, so, their down regulation or even deletion can be costly to the bacterium. Nevertheless, receptor’s loss or modification are very common. Random mutations leading to receptor’s modification are the most frequent cases, in which a simple alteration can have negative effects on phage binding, without affecting normal functioning of the receptor. In addition, bacteria can also hide phage receptors, including physical barriers such as capsules. Although these barriers consisting of cellular polymers, they might be able to block only certain phages, while other phages display enzymes capable to degrade those polymers (Hyman and Abedon, 2010).

3.2. Prevention takeover

Prevention takeover includes the mechanisms that avoid the takeover of bacteria metabolism like the phage genome uptake block, restriction modification and CRISPR. Uptake blocks prevent the phage DNA from entering the bacterial cytoplasm. This mechanism can be differentiated into mechanisms associated with bacterial chromosome, plasmids, or prophage genes (Hyman and Abedon, 2010). In the latter, known as superinfection exclusion, a preexisting virus prevents a phage from entering the cell by means of a DNA injection blocking mechanism (McGrath *et al.,* 2002). So this mechanism is not developed by the bacterial cell itself, but by the prophage, aiming the elimination of competitors for host resources (Folimonova, 2012).

Bacterial Restriction-Modification systems function as prokaryotic immune systems that attack foreign DNA entering the cell and involves restriction endonucleases and methyltransferases. Restriction endonucleases cleave the non-self DNA such as phage, plasmids

and foreign DNA which is inactivated. All phages might be susceptible to restriction endonucleases, with exception of those that displayed adaptations like RNA genomes and the mutational loss of those recognition sequences (Hyman and Abedon, 2010).

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is another resistance mechanism, and due to its recent discovery is not completely understood. CRISPR loci consist of several non-contiguous direct repeats, separated by stretches of variable sequences known as spacers. Those spacer sequences are segments of "captured" plasmid or viral sequences. Bacteria cells are then resistant to phages harbouring those corresponding sequences (Horvath and Barrangou, 2010; Hyman and Abedon, 2010). This called bacteria "adaptative immune system" uses small guide RNA, also known as CRISPR-RNA (crRNA), to neutralize invading viruses and plasmids (Horvath and Barrangou, 2010).

3.3. Abortive infection

Also referred to as phage exclusion mechanisms, in abortive infection, phage multiplication is blocked leading to the release of few or no progeny particles and to the death of the infected cell. The protein synthesis is blocked when nucleases are activated and cleave the anticodon loop of a specific tRNA. In case of death of both phage and bacterium, the rest of the bacterial population is protected from subsequent infection and survives (Chopin *et al.*, 2005; Hyman and Abedon, 2010).

Since resistance may reduce bacteria's fitness, or in case of the receptor used by the phage be a bacterial virulence determinant, loss of this receptor would dramatically decrease virulence of bacteria (Skurnik and Strauch, 2006). In this case, bacterial evolution to phage resistance would have the benefit of creating mutant bacteria that would no longer be capable of causing disease, and as long as phages are present as selective agents, the resistant, avirulent mutants would replace virulent forms. In addition, phages could even be chosen specifically for this property. Although, there is a possible complication to this hypothesis: subsequent evolution might restore the fitness or virulence of those resistant bacteria, either by selecting susceptibility or by compensatory evolution through second site mutations, as has been observed with antibiotic resistance. For now, it is not yet possible to make a general statement about whether resistance to a single and multiple phages will alter fitness or reduce the virulence of pathogenic bacteria, or even how those alterations will change through subsequent evolution (Levin *et al.,* 2004).

4. Mycobacteria and mycobacteriophages

4.1. Mycobacteria

The genus *Mycobacterium* consists of a wide variety of species occupying many ecological niches and displaying very diverse phenotypes (Pavelka Jr, 2000). It designates a group of aerobic, acid-alcohol fast, rod-shaped from the order *Actinomycetales* (Rastogi *et al.*, 2001).

Mycobacteria is considered as Gram-positive, even though mycobacteria do not retain the Gram stain. Mycobacteria can be divided into slow- and fast- growers. Slow-growing mycobacteria require more than 7 days for the appearance of colonies, and this group includes the pathogenic *Mycobacterium tuberculosis*. The group of fast-growers which includes opportunist mycobacteria such as *Mycobacterium smegmatis,* is characterized by visible growth from dilute inoculation within 7 days (Pavelka Jr., 2000).

There are some mycobacteria species characterized as important animal and human pathogens. *Mycobacterium tuberculosis* and *Mycobacterium leprae*, are two of those pathogenic species within the genus *Mycobacterium*, and the causative agents of the two world's oldest diseases in humans, tuberculosis and leprosy (Hett and Rubin, 2008). One third of the world's population may be infected with tuberculosis bacillus, even not showing symptoms of the disease, since *M. tuberculosis* has the ability to remain dormant for decades within an individual (Hett and Rubin, 2008). It is contagious and airborne, and is the second leading cause of death from a single infectious agent, right after the Human Immunodeficiency Virus (HIV). In 2013, 9 million people fell ill and 1.5 million people died with this disease. Due to an increase in antibiotic resistance, in 2013, 480 000 people developed multidrug-resistant tuberculosis (MDR-TB) and 210 000 people died (WHO, 2014). The mycobacteria cell envelope has unique characteristics, making it impermeable to several compounds, so resistance to several drugs, are partially due to this characteristic (Hett and Rubin, 2008).

The cytoplasmatic membrane, that is similar to other bacterial cytoplasmatic membranes, is also surrounded by a peptidoglycan cell wall which, and in contrast with other bacteria is attached to arabinogalactan that is esterified to mycolic acids forming a mycolyl arabinogalactan-peptidoglygan (mAGP) complex. These covalently linked mycolic acids comprise all or part of the inner leaflet of the outer membrane. The outermost leaflet of the outer membrane is composed by glycolipids, phospholipids, and species-specific lipids. Among these constituents, and similarly to Gram-negative bacteria, there are also porins. Outside the outer membrane, mycobacteria have a layer of proteins, polysaccharides and a small amount of lipids, known as capsule (Hett *et al*., 2008).

4.2. Mycobacteriophages

Mycobacteriophages are bacteriophages that infect mycobacterial hosts (Hatfull *et al.*, 2008). The first mycobacteriophages were isolated using *M. smegmatis* as host in the late 1940s, followed by isolation in *M. tuberculosis* (Hatfull, 2014).

All mycobacteriophages isolated until now are dsDNA with long contractile tails or long non-contractile tails. Although only few mycobacteriophages have been characterized in detail, and the total numbers of mycobacteriophage genomes deposited in GenBank being until this date 422 (The Mycobacteriophage Database, 2014), several hundred of new mycobacteriophages have already been isolated (5604) (The Mycobacteriophage Database, 2014). This interest is due to their important role in the development of genetic and diagnostic tools for pathogenic mycobacteria, including *Mycobacterium tuberculosis,* which is infected by approximately 10% of those mycobacteriophages. Mechanisms of mycobacteriophage infection and host range determination remain poorly studied and the knowledge about mycobacterial phage receptors is minimal (Hatfull *et al*., 2008; Chen *et al*., 2009).

Phage typing studies have revealed that mycobacteriophages can have a wide variety of preferences for bacterial hosts. While some phages like D29 are able to infect different species of mycobacteria, including slow growing mycobacteria such as *M. tuberculosis*, others like Barnyard have a narrow range, being capable of infecting only *M. smegmatis* mc^2155 (Rybniker *et al.*, 2006).

In a recent study, Jacobs-Sera *et al*. (2012) tested the preference of a large collection of mycobacteriophages for *M. tuberculosis* and other strains of *M. smegmatis*, and showed that few phages were able to infect this slow-growing species. In contrast, many of the phages tested, infect other strains of *M. smegmatis*. They also found that there is a correlation between genome type and host range. Only phages from Cluster K and Subclusters A2 and A3 efficiently infected *M. tuberculosis*.

4.3. Mycobacteria receptors for mycobacteriophage adsorption

Although mycobacteriophage host preferences are expected to be strongly related to the availability of specific cellular receptors, and mycobacteria phage-resistant mutants can be isolated in a simple manner, few have been identified and little is known about their chemical nature (Furuchi and Tokunaga, 1972; Hatfull, 2010). As previously mentioned, mycobacteria cell envelope is a complex protective barrier for the cell, however, it must contain defined components for both import and secretion of macromolecules, which similarly to other bacteria, might function as receptors for mycobacteriophage adsorption (Barsom *et al*., 1996).

In 1972, Furuchi *et al.* purified and characterized the receptor of *M. smegmatis* for the adsorption of mycobacteriophage D4. It was characterized as an apolar mycoside C of the mycobacteria, although the subsequent events required for contact with the membrane and translocation of the mycobacteriophage genome across the membrane, have not yet been described (Hatfull, 2010). The author also stated in his article, that previous groups have reported that glycolipids may be receptors for the attachment of mycobacteriophage Phlei to the previously designated *M. phlei* [M. smegmatis ATCC 19249 (Khoo *et al.*, 1996)]; lipopolysaccharides for GS7 and D4 that infect *M. smegmatis* and phospholipids for D29 mycobacteriophage (Furuchi *et al*., 1972). In 1996, Barsom and colleagues reported that *M. smegmatis mpr* gene overexpression confers resistance to mycobacteriophages L5 and D29, maybe because it alters the structure of the host cell wall or membrane, thereby inhibiting productive phage DNA injection. In 2010, Hatfull reported that *mpr* gene normal function, as well as the reason why its overexpression leads to phage resistance, are still completely unknown. An isolated study in 2009 revealed that glycopeptidolypids (GPL) are the receptors for phage I3 in *M. smegmatis*, and defects in its biosynthesis confer resistance to I3 (Chen *et al.*, 2009). These are all the published studies about mycobacteria receptors for phage infections, which leads to the conclusion that little is known about mycobacteria phage receptors and, in general, all these are quite old and somehow inconclusive.

4.4. Recombinant mycobacteriophages as genetic tools

Since the discovery of the first mycobacteriophages, due to the lack of efficient tools for molecular genetic analysis, investigations focused on the biology of those known mycobacteriophages and their potential applications. The absence of molecular genetics methods, represented substantial impediments for the development of systems for genetic manipulation of mycobacteria, especially in *M. tuberculosis*. This was overcome with the

construction of libraries of mycobacterial DNA in *E. coli*, and later, with the development of vectors that behave as phages in mycobacteria and as large plasmids in *E. coli*, known as shuttle phasmids (Pelicic and Reyrat, 1998; Hatfull, 2014). Shuttle phasmids were, until recently, the most powerful method for the construction of recombinant mycobacteriophages. However, the relatively large size of mycobacteriophage genomes for shuttle phasmid construction restricts the number of phages for which this technique is applicable. Besides, shuttle phasmid require complex genetic constructions with several steps of manipulation and screening (Marinelli *et al.*, 2008; van Kessel *et al.*, 2008). In 2008, Marinelli and colleagues proposed a novel approach for the construction of mutant mycobacteriophages named BRED – Bacteriophage Recombineering of Electroporated DNA.

Recombineering refers to genetic engineering using proteins required for homologous recombination. This method, is used in *E. coli*, and takes advantage of λ Red system, or from the Rac profage, in which *rec*E and *rec*T genes enhance significantly the frequencies of homologous recombination (van Kessel *et al*., 2008)

BRED also refers to recombinant DNA engineering by homologous recombination, where mycobacterial strains constructed for recombineering, contain an extra chromosomal plasmid in which *gp60* and *gp61* genes are present and under the control of the inducible acetamidase promoter, thus conferring high levels of homologous recombination. This genes encode for an exonuclease (*gp60)* and its associated recombinase (*gp61*), are both from mycobacteriophage Che9c and similar to those present in *E. coli rec*E and *rec*T respectively. The principle is the same as the latter, ssDNA substrates used for recombineering only require the recombinase, whereas for dsDNA both exonuclease and its associated recombinase are essential (Marinelli *et al.*, 2008; 2012).

BRED technology allows the construction of unmarked deletions, in-frame internal deletions, insertions, base substitutions, precise gene replacements, between others, and it does not apply the complex construction of cloning systems like shuttle phasmid, nor does it require a selectable marker. (Marinelli *et al*., 2008; 2012).

4.5. Mycobacteriophage Ms6

In 1989, the Molecular Microbiology Laboratory of the Pharmacy Faculty of the Universidade de Lisboa isolated the mycobacteriophage Ms6 from a spontaneously induced culture of *M. smegmatis* HB5688. Ms6 is a temperate phage composed of an isometric polyhedral head, hexagonal in shape and with a long non-contractile tail. Its genome is a double stranded DNA molecule, with over 50kb and a CG content of 62% (Portugal *et al*., 1989).

Although the complete sequence of Ms6 genome is not yet completely annotated, some regions are already characterized including the lysis cassette (Fig. 3). The Ms6 lysis cassette is composed of five genes clustered downstream of a strong promoter region (Garcia *et al.*, 2002). Like all dsDNA phages, Ms6 uses the holin-endolysin strategy to achieve lysis of its host, however, the model of lysis is different from those described so far (Pimentel, 2014).

The gene responsible for endolysin is *lysA*, and the holin-like genes are *gp*4 and *gp*5 (Garcia *et al*., 2002; Catalão *et al.*, 2011a; 2011b). *lysA* was shown to translate two proteins which were designated Lysin384 and Lysin241 according to their aminoacids (aa) number (Catalão *et al.*, 2011b). In addition to these, two accessory lysis genes were also identified: *gp*1 and *lysB*. *gp*1 gene, encodes a chaperone-like protein, which is involved in LysA protein delivery to the peptidoglycan, and is necessary to achieve efficient lysis (Catalão *et al.*, 2010). *lysB* gene product, LysB, is an enzyme with lipolytic activity, which targets the *M. smegmatis* outer membrane by cleaving the ester bond between mycolic acids and the arabinogalactan of the mAGP complex (Gil *et al.*, 2008, 2010). Accordingly to the lysis genes organization, Ms6 seems to be closely related to phages of cluster F, subcluster F1 (Catalão *et al.*, 2011b).

In addition to the lysis cassette, the genetic elements involved in Ms6 integration were also identified. The *integrase* gene, which distances around 1.8 kb from *gp5* (Fig. 3), encodes a 373 aa protein which is responsible for Ms6 integration into the 3' end of *M. smegmatis* $tRNA^{Ala}$ gene. Near the 5' end of the *integrase* gene is a 26bp core sequence, where the recombination between the phage DNA and the bacterial genome occurs (Freitas-vieira *et al.*, 1998). Downstream of the *integrase*, and transcribed in the opposite direction, is *pin* gene, a 489bp Ms6 gene which encodes a membrane protein with 122 amino acids (aa) and acts in a super-infection exclusion mechanism (Pimentel, 1999). It was suggested that Pin protein blocks the initial phage infection step, between phage adsorption and injection of its DNA, preventing it from reaching the cytoplasm (Pimentel, 1999).

So far, the mycobacteria receptor for phage Ms6 and the ability to infect other mycobacteria is not yet know.

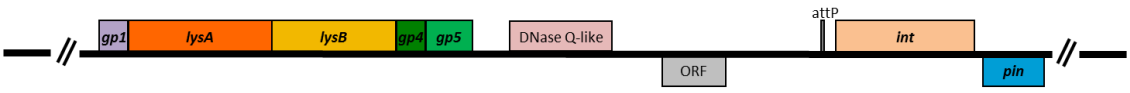

Figure 3 – Genetic organization of the Ms6 lysis cassete and prophage integration region.

II. Objectives

The main goal of this project was to **construct Ms6 derivatives mutants** that could be used to study mycobacteriophage-mycobacteria interactions. Our ultimate purpose was to answer some questions regarding the first steps of infection:

- Is Ms6 able to infect other mycobacteria?
- Is the ability to replicate only dependent on the presence of a phage receptor?

To help answer these questions, in this project it was proposed to:

- ✓ Evaluate Ms6 ability to infect other mycobacteria;
- ✓ Develop molecular tools that will allow the follow up of a Ms6 infection, by constructing Ms6 reporter phages;
- ✓ Evaluate their action inside *M. smegmatis* and other mycobacteria.

The development of mycobacteriophages as genetic tools is unquestionably necessary to help understand the interactions between mycobacteriophages and mycobacteria, and also their possible application as diagnostic and therapeutic agents.

With this project, it was expected to achieve more insights about Ms6 interaction with mycobacteria and provide new genetic tools for future studies.

III. Material and Methods

1. Bacterial Strains, Bacteriophages and growth conditions

Bacterial strains, phages and plasmids used throughout this study are listed in Table 1. *E. coli* was grown at 37°C in Luria-Bertani (LB) broth with shaking or agar, supplemented with 30µg mL^{-1} kanamycin (Sigma-Aldrich) or 250µg mL^{-1} hygromycin (Roche), when appropriate. *M. smegmatis* mc^2155 was grown at 37°C on Middlebrook 7H9 (BD™ Biosciences) with shaking, or Middlebrook 7H10 medium (BD™ Biosciences), both supplemented with 0.5% Glucose. When necessary, 50 µg mL^{-1} hygromycin was also added. For induced conditions 15µg mL^{-1} kanamycin, 0.2% succinate and 0.2% acetamide were added to 7H9.

All other mycobacteria were also grown at 37°C on Middlebrook 7H10 or Middlebrook 7H9 supplemented with 10% OADC (Oleic acid, Albumin, Dextrose, Catalase) (BD™ Biosciences) and 0.05% Tween 80 (Amresco®) with shaking for fast growing mycobacteria. For phage infections Tween 80 was omitted and 1mM $CaCl_2$ included.

When needed, cell concentrations were estimated by measuring the optical density (O.D.) at 600nm ($O.D._{600}$) using a spectrophotometer (Biophotometer-Eppendorf).

Phage stocks were prepared, after overnight (o.n.) elution (4°C) of the phages from several confluent lysis plates, with Phage Buffer (1M $MgSO_4$, 0.068M NaCl, 0.1M Tris-HCl [pH 7.5]). Those samples were then filtered and stored at 4°C. Phage titers were determined by making serial dilutions of the appropriate phage stock, incubated with *M. smegmatis* mc^2155 at 37°C 30min and then plated as top agar lawns on 7H10.

Table 1 - List of bacterial strains, phages and plasmids used throughout this study.

Strains, bacteriophages or plasmids	Description	Reference
Escherichia coli		
DH5α::pMN437	pMN437;Hyg^r	(Song *et al.*, 2008)
JM109::pSMT3	*rec*A1, *end*A1 gyrA96; *thi*; *hsd*R17; *sup*E44; *rel*A1; Δ (*lac-pro*AB), [F' *tra*D36, *pro*AB, *laq*I^{q}ZΔM15]	Stratagene
Mycobacterium smegmatis		
mc^{2}155 (wt)	High-transformation-efficiency mutant of *M. smegmatis* ATCC 607	(Snapper *et al.*, 1990)
mc^{2}155::pJV53	*M. smegmatis* mc^{2}155 carrying a plasmid expressing recombinerring functions; Kan^r	(van Kessel and Hatfull, 2007)
Fast- and Slow-growing Mycobacteria		
Mycobacterium fortuitum		American Type Culture Collection (ATCC) 6841 B
Mycobacterium aurum		A+ (Institute Pasteur Production – IPP)
Mycobacterium vaccae		SN 901 (IPP)
Mycobacterium tuberculosis H37Ra		ATCC 25177
Mycobacterium bovis BCG		ATCC 35734 (IPP)
Bacteriophages		
Ms6 wt	Temperate mycobacteriophage isolated from *M. smegmatis*	(Portugal *et al.*, 1989)
Ms6*lysB*::*egfp*	Replacement of *lysB* gene for *enhanced green fluorescen protein (egfp)* gene	This study
Ms6*pin*:: gfp_m^{2+}	Replacement of phage inhibition (*pin*) gene for gfp_m^{2+} (green fluorescent protein) gene	This study
Plasmids		
pEGFP-N1	Expresses Enhanced green fluorescent protein (EGFP), optimized for brighter fluorescence and higher expression; Kan^r	(Cormack *et al.*, 1996)
pMN437	P_{smyc}-*msp*A; ColE1 origin; PAL5000; P_{smyc}-gfp_m^{2+}; Hyg^r	(Song *et al.*, 2008)
pSMT3	Shuttle vector for mycobacteria; Hyg^r	Addgene plasmid 32362 O'Gaora, 1998

2. Electrocompetent cells and electroporation

Recombineering *M. smegmatis* electrocompetent cells were prepared as described previously (van Kessel and Hatfull, 2007). Briefly, *M. smegmatis* mc^{2}155::pJV53 cells were grown until O.D.$_{600}$ of ~0.4 and expression of the recombineering proteins was induced with 0.2% acetamide for 3h at 37°C with shaking. After 1h30 of incubation on ice, cells were pelleted by centrifugation and washed three times with ice-cold 10% (v/v) glycerol (AppliChem). At the end, the pelleted cells were resuspended in 1/20 volume with ice-cold 10% (v/v) glycerol and 100 µL aliquots were kept at -80°C until use.

All other mycobacteria including *M. smegmatis* mc^2155 were prepared as described in Parish and Brown (2009). Cells were grown until OD_{600} 0.7-0.8 and after 1h30 of incubation on ice, the procedure was the same as described above.

All mycobacteria used were electroporated using 2mm cuvettes (VWR®) in a Gene Pulser™ Electroporation System (Bio-Rad) with pulse settings of 2.5Kv, 1000Ω, 25µF and time constants read for quality transformation analysis.

3. Phage DNA Extraction

Phage DNA extraction was adapted from Sambrook and Russell (2001). To the required volume of phage suspension, 50µg mL^{-1} Proteinase K (AppliChem) and 0.5% SDS were added. After 1h at 56°C, 3M KCl were added, and after 30min on ice, the mixture was centrifuged at 16,900 xg for 5min at 4°C. The aqueous phase containing the DNA was transferred into a new tube and an equal volume of Phenol/chlorophorm/isoamyl alcohol (25:24:1) mixture (AppliChem) was added thrice, mixed and centrifuged (5 min, 4°C, 16,900 xg) to separate phases. This procedure was repeated once with Chlorophorm. DNA was precipitated by addition of an equal volume of isopropanol, 10% sodium acetate 3M, incubated at-20°C for at least 1h and then centrifuged for 45min (4°C, 16,900 xg). Finally, the DNA pellet was washed with 70% (v/v) ethanol, dried at 37°C and resuspended in an appropriate volume of ultra-pure water. DNA samples were stored at -20°C.

4. Plasmid DNA

Plasmid DNA extractions from *E. coli* strains were performed using NZYMiniprep kit (NZYTech) and stored at -20°C.

5. DNA Manipulation and Purification

All oligonucleotides and primers (Thermo Scientific) used in this work are listed in the appendix section (Appendix I).

DNA fragments were amplified by PCR using *Pfu* (Promega) or KOD (Hot Start Master Mix- Novagen®) DNA polymerase for high fidelity synthesis or NZYTaq (2X Green Master Mix – NZYTech) for screenings. All PCR proceedings were made according to manufacturer's instructions in a standard thermocycler (VWR® DOPPIO Thermal Cycler). Genomic DNA, plasmids, and PCR products were analysed by agarose gel electrophoresis, in 0.7 to 1.5% (w/v) (Sambrook and Russell, 2001) agarose gels (NZYTech), containing 3µL GreenSafe Premium (NZYTech) per 50mL of 0.5X TBE. Their approximate molecular mass was determined using DNA

Ladders (NZYTech; Thermo Scientific) and detected under UV light with the use of a transiluminator or BioRad system with an incorporated camera, for pictures capture.

For subsequent reactions, PCR products were purified using MinElute® PCR Purification Kit (Qiagen®) or by extracting the DNA fragment from agarose gel using Invisorb® (Invitek) and eluted in minimal volumes of ultra-pure water and stored at -20°C. DNA concentrations and purity were determined by spectrophotometry using NanoDrop® (ND-1000 - Thermo Scientific).

6. Infection Assay and Spot Test

M. fortuitum, *M. vaccae*, *M. aurum*, *M. bovis* BCG and *M. tuberculosis* H37Ra were grown to an $O.D._{600}$ of 0.5-0.8 or until a high cell density was visible reached (turbid). For an infection assay, cells were pelleted, resuspended in phage suspension at a MOI (Multiplicity Of Infection) of 100, incubated at 37°C for different periods of time and then plated as top agar lawns on 7H10.For the spot assay a turbid cell suspension of each mycobacteria, were plated in top agar lawns in 7H10, and after cooling, 10µL spots were made from different phage dilutions.

7. Construction of Ms6 mutant phages

Mutant phages were constructed using the Bacteriophage Recombineering of Electroporated DNA (BRED) technique in *M. smegmatis* as described previously (Marinelli *et al*., 2008). To replace the Ms6 *lysB* gene with *egfp* gene (Ms6*lysB*::*egfp*), *egfp* was first amplified by PCR from pEGFP-N1 using the pair of primers **Pr *egfp* 5' Fw** / **Pr *egfp* 3' Rv**. The *egfp* gene fragment was then extended by PCR using two 75bp extender primers, **Pr repl *egfp*/*lysB* Fw1** and **Pr repl *egfp*/*lysB* Rv1** which have 25bp of homology with the template and add 50bp of homology upstream and downstream of the sequence flanking the *lysB* gene. A second set of primers, **Pr repl *egfp*/*lysB* Fw2** and **Pr repl *egfp*/*lysB* Rv2**, was used to extend the previous generated substrate which added additional 50bp of homology flanking the *lysB* gene on either end of the substrate.

To construct Ms6*pin*::gfp_m^{2+} and Ms6*pin*::*hyg*, the same strategy was used. gfp_m^2 was amplified from plasmid pMN437 and *hyg* from pSMT3 with the appropriate primers, **Pr *gfpm++* Fw** / **Pr *gfpm++* Rv** and **Pr *hyg* pSMT3 Fw** / **Pr *hyg* pSMT3 Rv** respectively. The first set of extension primers used were **Pr repl *pin-gfpm2+* Fw** / **Pr repl *pin-gfp*m2+ Rv** and **Pr repl *pin-hyg* Fw** / **Pr repl *pin-hyg* Rv** for Ms6*pin*:: gfp_m^2 and Ms6*pin*::*hyg* respectively, and the second set of extender primers used were the same for both substrates **Pr Ext oligo-del *pin* Fw** / **Pr Ext oligo-del *pin* Rv**.

The final products were purified and co-electroporated with Ms6 genomic DNA, into 100µL of electrocompetent *M. smegmatis* mc²155:pJV53. Cells were resuspended in 900µL of 7H9 medium supplemented with glucose and $CaCl_2$, incubated at 37°C with shaking for 1h30 and plated on 7H10 as top agar lawns with *M. smegmatis mc²155.* After o.n. incubation at 37°C, individual plaques were picked into 100µL phage buffer, incubated at 4°C o.n. and screened by PCR for mutant detection. **Pr *lysA* 3' Fw screen** / **Pr screen repl *lysB*/*egfp* Rv** ; **Pr screen repl *pin*-*gfpm2+* Fw** / **Pr screen del *pin* Rv** and **Pr screen repl *pin*-*hyg* Fw** / **Pr screen del *pin* Rv** pair of primers were used for screening of Ms6*lysB*::*egfp*, Ms6*pin*:: gfp_m^2 and Ms6*pin*::*hyg*, respectively. Mixed primary plaques containing both mutant and wild-type DNA were eluted as described above, and serial dilutions were plated with *M. smegmatis* mc²155 as top agar lawns on 7H10. Individual secondary plaques were again picked, eluted and screened. This procedure was repeated until a pure phage mutant was obtained.

8. Fluorescence Detection of *M. smegmatis* infected with mutant Ms6 derivatives

M. smegmatis mc²155 cells were grown to an $O.D._{600}$ of 0.6. Then, 10 or 20mL were pelleted by centrifugation and resuspended in 1 or 2mL of phage suspension (Ms6*lysB*::*egfp* or Ms6*pin*::gfp_{m+}^{2+}) supplemented with 1mM $CaCl_2$, using a MOI of 10 for Ms6*lysB*::*egfp* or MOI 100 for Ms6*pin*::gfp_m^{2+}. As a control, the same assay was performed with Ms6 wt. The mixture was incubated for 50min at 37°C for phage adsorption, and then 100µL 0.4% H_2SO_4 was added for 5min in order to inactivate non-adsorbed phages. The suspension was neutralized with 100µL 0.4% NaOH and then diluted 1:10 fold with 7H9 medium properly supplemented and pre-warmed, and the incubation was continued for around 300min at 37°C, at low shaking, and every 15 to 30min ~400µL aliquots were collected and kept on ice. Ten microliters of each time point were evaluated by bright field and fluorescence microscopy (Axioscope A1, Carl Zeiss, Germany) with an 63X objective with oil immersion, filter sets suitable for UV and images were analysed using ZEN LITE imaging software acquired by AxioCam HRm (Carl Zeiss, Germany). Optical density ($O.D._{600}$) and fluorescence intensity of EGFP or GFP_m^{2+} were evaluated with a microplate reader (Tecan Infinite® M200, Tecan Austria GmbH) with excitation and emission wavelengths set at 488 and 520nm respectively. Two hundred microliters of each time point were applied to plate wells, and fluorescence optics of the microplate reader were set to top and bottom of the well in addition to multiple reads per well for each sample. Wells with only medium (7H9 supplemented) and with *M. smegmatis* mc²155 infected with Ms6 wt suspension, were also measured as blanks or control of tested conditions. Final fluorescence values were obtained by

subtracting the obtained measurements for the control (*M. segmatis* mc^{2}155 infected with Ms6 suspension).

9. Isolation of lysogenic cells and Immunity test - Sensibility to Ms6 assay

To isolate lysogenic cells following an infection with Ms6*pin*::*gfp*$_m^{2+}$, an isolated phage plaque formed on *M. smegmatis* mc^{2}155 lawns, was piqued and *Mycobacterium* cells were isolated on a 7H10 medium plate. The resulting isolated colonies, were grown in solid and liquid media and were analysed by fluorescent microscopy

To confirm lysogeny, an immunity assay was performed. Ten microliters of Ms6*pin*::*gfp*$_m^{2+}$ phage stock with titer of 4x10^{10} pfu mL^{-1}, were drained in line on a 7H10 plate. Perpendicular to this, *M. smegmatis* cells, were crossed on the media and incubated at 37°C.

10. Transfection Assay

From 300 up to 1200ng genomic DNA from Ms6 wt, Ms6*lysB*::*egfp* or Ms6*pin*::*gfp*$_m^{2+}$ were electroporated into electrocompetent cells of *M. smegmatis* mc^2 155, *M. fortuitum*, *M. vaccae*, *M. aurum*, *M. bovis* BCG and *M. tuberculosis* H37Ra, resuspended in 1mL of 7H9 medium supplemented with glucose plus $CaCl_2$ and incubated o.n. at 37°C, or 3h at 37°C for *M. smegmatis* mc^2. Cell suspensions were then centrifuged for 20min, resuspended in 1mL of phage buffer, lysed by sonication (2 cycles of 5s with 30s resting on ice between cycles) and again centrifuged for 20min. The resulting supernatants were filtered (0.2µm, 13mm). Three hundred microliters were incubated then with 200µL of an o.n. culture of *M. smegmatis* mc^{2}155 for 30 min at 37°C and plated as top agar lawns on 7H10. From o.n. incubations, 100µL of each mycobacteria suspensions were collected, kept on ice and analysed by fluorescence microscopy.

IV. Results

1. Infection Assay and Spot test

The ability of a bacteriophage to infect and propagate in susceptible bacteria, is experimentally demonstrated by the incubation of that phage with the bacteria, spread the mixtures as top agar lawns on solid medium, and after appropriate incubation conditions, expect to observe plaques formation. (Ellis and Delbrück, 1939; Maniloff, 2001)

T**o understand if Ms6 is able to infect other mycobacteria besides its host,** the first approach was to incubate Ms6 phage with fast- and slow- growing mycobacteria, in two assays, an infection and a spot test, in expectation of plaque or clear spot formation. After 3 days of incubation for fast-growing, and more than 7 for slow-growing mycobacteria, none of them showed any sign of plaque or spot formation, as observed with *M. smegmatis* mc^2155 (wt), used as control, suggesting that Ms6 is not able to infect those mycobacteria.

2. Construction of Ms6 mutant phages

At the end of the infection cycle, infected cells lyse to release progeny phages, which is visible through plaque formation within the agar (Hatfull, 2000). However, the efficiency of plating can be affected by several factors, including the specific host strain, first steps failure of phage infection and after infection a low number of phage particles produced which would enable visible plaques (Kutter, 2009). Taking these facts into consideration, and since we did not observed the formation of phage plaques, we decided to **construct reporter phages** and use them to follow a phage infection through the production of fluorescent proteins. To achieve this, we took advantage of the BRED technology (Marinelli *et al.,* 2008). With BRED, it is possible to construct deletions, insertions and point mutations in mycobacteriophage genomes (Marinelli *et al.,* 2008). Our approach was to construct Ms6 derivative mutants expressing a fluorescent protein which would allow the monitoring of the infection process. Our strategy was to replace a Ms6 gene for the reporter gene. To make a successful gene replacement, first we needed to be sure that the gene to be replaced is non-essential for phage survival.

Fluorescent proteins have been extensively used as reporters because they do not cause disruption to cell metabolism or toxicity. Green Fluorescent Protein (GFP), from the jellyfish *Aequorea victoria* (Shimomura *et al.*, 1962), emits green fluorescence when exposed to blue light (UV) without the substrate or cofactor addiction, revealing to be a very good instrument for gene monitoring and protein expression in living organisms (Chalfie *et al.*, 1994).

F. Gil *et al.* (2011) has previously shown that the deletion of lysis gene *lysB* is non-essential for phage viability. However, as mentioned before, *lysB* is part of the lysis cassette and thus the LysB protein would only be produced if a lytic cycle occurs in a late stage of infection. So it was important to construct an additional mutant phage in which the gene to be replaced would be expressed during lysogeny.

Although the Ms6 genome sequence is not yet completely annotated, in addition to the lysis cassette, the region encompassing the *integrase* gene was already sequenced and characterized. Gene *pin* encodes a membrane protein predicted to be involved in a superinfection exclusion mechanism, and thus expected to be expressed during lysogeny.

Regarding this, *lysB* and *pin* were the genes chosen to be replaced by a *gfp* gene. So we used two modified *gfp* genes, *egfp* (720bp) and gfp_m^{2+} (735bp) to replace Ms6 non-essential genes, *lysB* (999bp) and *pin* (489bp) respectively.

2.1. Ms6*lysB*::*egfp*

The low fluorescence intensity of GFP wild type, lead to Zhang and his colleagues make changes in the GFP coding sequence. The originated Enhanced Green Fluorescent Protein (EGFP) was 35 times brighter than GFP wt, increasing the sensitivity of the reporter protein *(Zhang et al.*, 1996). On this basis, the first recombineering mutant phage to be constructed would have *egfp* as reporter gene, in place of *lysB* gene and under regulation of Ms6 lytic cassette.

After amplification of *egfp* gene from plasmid pEGFP-N1 (Cormack *et al.*, 1996) the resulting 720bp fragment was then extended with primers **Pr repl *egfp*/*lysB* Fw1** and **Pr repl *egfp*/*lysB* Rv1** to introduce 50bp homologous to the region flanking the sequence to be replaced, on either side of the DNA segment (Fig. 4.I.). In a second extension round with the pair of primers **Pr repl *egfp*/*lysB* Fw2** and **Pr repl *egfp*/*lysB* Rv2**, additional 50bp homologous to the flanking region were added to either side of the template (Fig.4.II.) resulting in the final 920bp recombineering substrate (Fig.4.III.). Two hundred nanograms or 400ng of Ms6 genomic DNA (Table 2) were co-electroporated with the recombineering substrate into *M. smegmatis* mc^2155::pJV53 cells previously induced to express the recombineering functions (Fig.4.IV.) at quantities varying from 50ng to 640ng (Table 2). Six co-electroporations were performed, testing different amounts of both substrate and genomic DNA, and four electroporations using genomic

DNA for control conditions (Table 2). With exception of the control, the plaques obtained were screened for the desired replacement, using primers **Pr *lysA* 3' Fw screen**, which hybridizes at the end of *lysA*, an existing gene upstream of *lysB*, and **Pr screen repl *lysB*/*egfp* Rv**, which hybridizes inside the *egfp* gene (Fig.4.V.). The desired mutation was only detected in the recombination events where 400ng of Ms6 DNA and 320ng of the recombineering substrate were used. From those primary plaques containing both mutant and wild-type phage DNA, serial dilutions were made, and plated with *M. smegmatis* wt. Individual secondary plaques were again picked, eluted and screened. After two screening rounds, the desired mutant phage was efficiently recovered from one original mutant-containing plaque. The absence of PCR amplification with a pair of primers within the *lysB* gene confirmed purify of the obtained mutant phage (Fig.4.V.). This recovered mutant, was tested for the presence of *lysB* to ensure the its purity (negative control). Only the wt phage used as control resulted in the amplification of a 582bp fragment for *lysB* gene presence (Fig.4.VI.). A fragment from the mutant phage DNA amplified and sequenced, using **Pr *lysA* 3´ Fw screen** and **Pr ORF4 3' HindIII Rv**, two primers which hybridize upstream and downstream of the *egfp* gene (Fig.4.VII.). Sequencing results confirmed the successful recombination.

Table 2 - Amounts of Ms6 DNA and substrate used for co-electroporation, the respective number of phage plaques obtained and mutants detected in primary screenings to obtain Ms6*lysB*::e*gfp.*

Amount of Ms6 DNA	Amount of substrate	Number of plaques	Positives from screening PCR
400ng (Control)	-	36	-
400ng (Control)	-	64	-
800ng (Control)	-	230	-
800ng (Control)	-	270	-
200ng	640	145	-
400ng	640	37	-
400ng	320	42	1
400ng	320	89	2
400ng	60	71	-
400ng	50	200	-

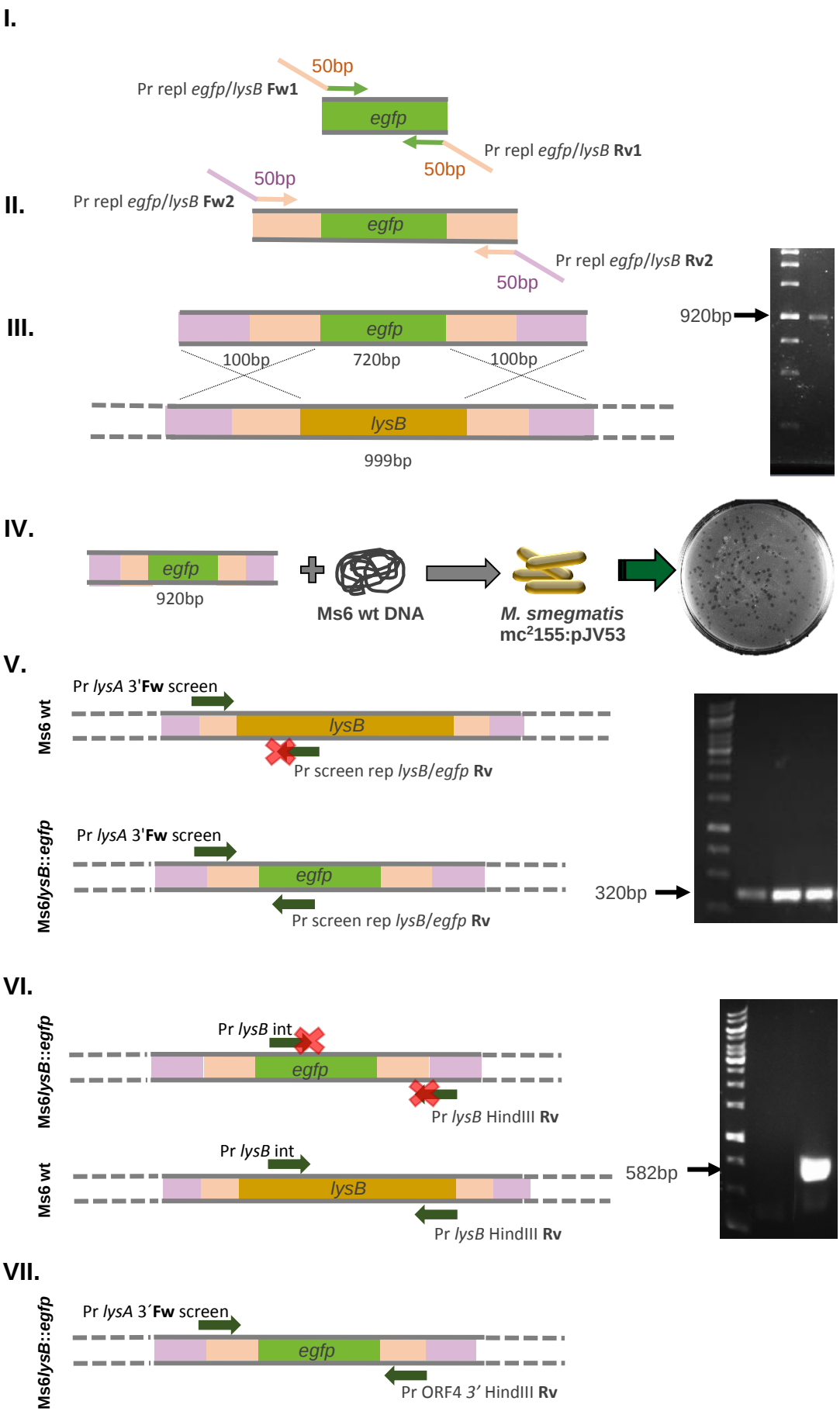
I.
50bp
Pr repl egfp/lysB Fw1
egfp
Pr repl egfp/lysB Rv1
50bp
II.
50bp
Pr repl egfp/lysB Fw2
egfp
Pr repl egfp/lysB Rv2
50bp
III.
egfp
100bp
720bp
100bp
lysB
999bp
920bp
IV.
egfp
920bp
Ms6 wt DNA
M. smegmatis
mc²155:pJV53
V.
Ms6 wt
Pr lysA 3'Fw screen
lysB
Pr screen rep lysB/egfp Rv
Ms6lysB::egfp
Pr lysA 3'Fw screen
egfp
Pr screen rep lysB/egfp Rv
320bp
VI.
Ms6lysB::egfp
Pr lysB int
egfp
Pr lysB HindIII Rv
Ms6 wt
Pr lysB int
lysB
Pr lysB HindIII Rv
582bp
VII.
Ms6lysB::egfp
Pr lysA 3´Fw screen
egfp
Pr ORF4 3´ HindIII Rv

Figure 4 - Schematic representation of the strategy used for the replacement of *lysB* gene for *egfp* gene and subsequent screenings, with the respective primers. I.- First extension PCR, in which primers added 50bp to each end of the *egfp* gene. II.- The resulting fragment was then used as template to add other 50bp to each end of the fragment. III.- The fragment obtained from the extension PCR had 100bp of homology to either side of the *lysB*, the target gene. IV.- This 920bp recombineering substrate was co-electroporated with Ms6 DNA into *M. semgmatis* mc^2155::pJV53 cells. After incubation on a lawn of bacteria, the obtained phage plaques were picked and screened. V.- In case of Ms6*lysB*::*egfp* presence, a 320bp fragment was amplified. On the contrary, if only Ms6 wt was present, there was absence in amplification. VI.- The purified Ms6*lysB*::*egfp* was screened to confirm *lysB* inexistence, only the control (Ms6 wt) amplified a 582bp fragment. VII.- A fragment from the purified Ms6*lysB*::*egfp* was amplified for sequentiation.

2.2. Ms6Δ*pin*

In opposition to *lysB* non-essentiality, there was no previous studies proving that *pin* was non-essential for Ms6 function. Regarding this fact, before the construction of phage mutants replacing *pin* gene, it was necessary to confirm that *pin* was a non-essential gene. For that purpose the first approach was to create a mutant phage lacking this gene (Ms6Δ*pin*), to confirm it. Again, we used the BRED technology to construct Ms6Δ*pin*. The approach was similar to that of gene replacements; however, the initial substrate was not a gene amplified from a plasmid, but a synthetized 100nt oligo with 50nt of homology upstream and downstream of the target region to delete. This oligo would be amplified with 75nt extender primers which would add 50bp of homology to either end of the 100nt oligo, forming a 200bp dsDNA substrate.

After several attempts we could not obtain the expected 200bp fragment, but only a PCR product of approximately 150bp (Fig. 5). Regarding this, we decided to proceed with the construction of Ms6 mutant phages with *pin* gene replaced by a reporter gene.

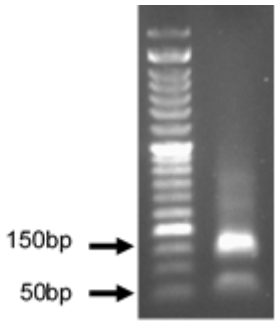

Figure 5 - Agarose gel of the originated fragment after 100nt oligo extension PCR for *pin* delection recombineering substrate. Comparing with molecular DNA ladder (NZYDNA Ladder VI - NZYTech).

2.3. Ms6*pin*::*gfp*$_m^{2+}$

In 2010, Steinhauer *et al.* constructed a *gfp* gene with a shorter distance between Shine-Dalgarno sequence and start codon, and two additional mutations to enhance GFP thermostability. These alterations revealed a 2-fold fluorescence increase, when compared with *egfp* in *M. bovis* BCG. This gene was denominated *gfp*$_m^{2+}$, and revealed to be a superior *gfp* variant in fast- and slow-growing mycobacteria. For this reason *gfp*$_m^{2+}$ gene was chosen to replace *pin* gene in Ms6 genome.

The strategy to construct Ms6*pin*::*gfp*$_m^{2+}$ was similar to that of Ms6*lysB*::*egfp*. *gfp*$_m^{2+}$ gene and its optimized Shine-Dalgarno sequence were amplified from plasmid pMN437 (Song *et al.*, 2008). Then, the amplified gene was extended through two additional PCR rounds (Fig.6.I-II.) in order to obtain a final recombineering substrate of 935bp with 100bp homology to either ends of *pin*, the target gene (Fig.6.III.). However, it was not possible to obtain a pure 935 recombineering substrate, and an additional fragment >800bp was always amplified (Fig.6.III.). By reasons that we cannot explain, we could never separate the two fragments, so we decided to proceed with the recombineering event using the mixed fragments. Again, we tested different amounts of the recombineering substrates and genomic DNA for co-electroporation (table 3 for tested conditions; Fig. 7.I. for schematic representation of co-electroporation). After o.n. incubation on a lawn with *M. smegmatis* wt, the resulting plaques predicted to contain a mixture of mutant and wt phages (Fig. 7.I.). These were piqued and screened with **Pr screen repl *pin-gfpm2+* Fw** and **Pr screen del *pin* Rv**, designed to amplify a DNA segment of 584bp containing the *gfp*$_m^{2+}$ gene. In case of *gfp*$_m^{2+}$ absence, the Fw primer could not hybridize, and the amplification was not verified (Fig.7.II.). From all plaques screened (with exception of controls), the desired PCR product was only obtained when 400ng of Ms6 DNA were co-electroporated with 200ng of the recombineering substrate (Table 3). This positive plaque contained the mutant and the wt phage, so Ms6*pin*::*gfp*$_m^{2+}$ was purified by re-plating serial dilutions. We obtained the desired mutant phage after two rounds of purification. This was confirmed by PCR using a pair of primers that flank the replaced region. As observed in figure 7.III., we only obtained the expected 1013bp fragment corresponding to the presence *gfp*$_m^{2+}$. A fragment of 767bp corresponding to the presence of gene *pin* was only detected in the Ms6 wt DNA control. (Fig. 7.III.). To confirm the construction of Ms6*pin*::*gfp*$_m^{2+}$ we amplified a DNA fragment using primers **Pr Primu1** and **Pr Primu2**, which flank the replaced region (Fig.7.IV.).

Surprisingly we detected 5 nucleotide substitutions and one nucleotide insertion on the DNA sequenced (appendix 2). However, only one occurred within the *gfp*$_m^{2+}$ gene, an A/G substitution 12bp upstream of the stop codon. This substitution resulted in a change of a

Glutamic Acid, an acidic polar aa, to Glycine, a neutral non-polar aa. The other four substitutions occurred out the coding sequences, one downstream and three upstream of the *pin* gene. The nucleotide insertion was also in a non- coding region, between the *pin* gene and the *integrase* gene, thus not affecting the expression of gfp_m^{2+}. Table 4 lists all mutations analysed from DNA sequencing results.

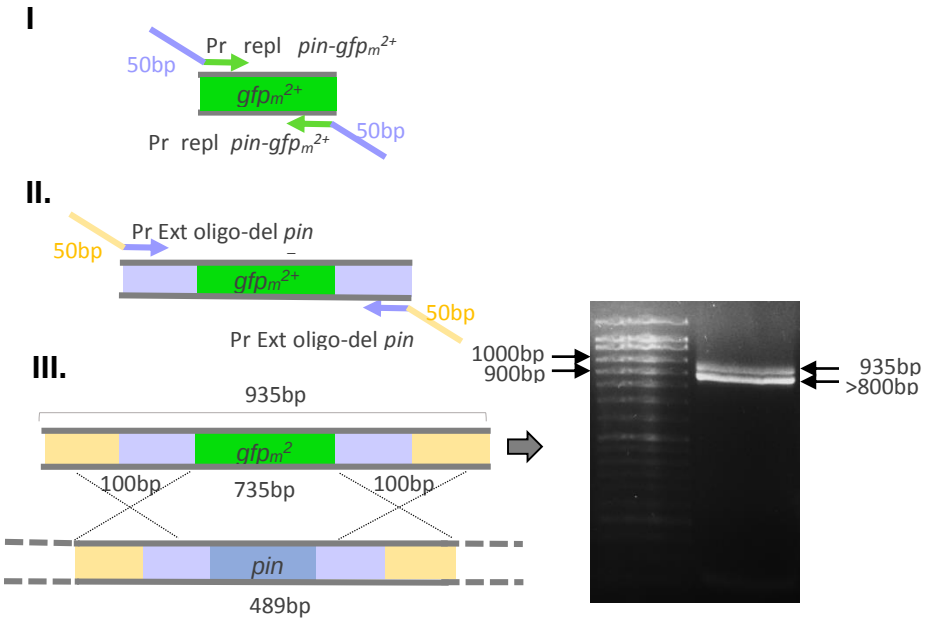

Figure 6 - Schematic representation of the strategy used for the replacement of *pin* gene for gfp_m^{2+} gene. I. First extension PCR, in which a pair of primers added 50bp to each end of the gfp_m^{2+} gene. II. The resulting fragment was then used as template to add other 50bp to each end of the fragment. III. The 935bp fragment obtained from the previous extension PCR had 100bp of homology to either side of the *pin*, the target gene. Additionally a fragment >800bp was also amplified, however it was not possible to separate them.

Table 3 - Amounts of Ms6 DNA and substrate used for co-electroporation, the respective number of phage plaques obtained and mutants detected in primary screenings to obtain Ms6*pin*::gfp_m^{2+}.

Amount of Ms6 DNA	Amount of Substrate	Number of plaques	Positives from screening PCR
200ng (Control)	-	14	-
400ng (Control)	-	46	-
100ng	200	2	-
100ng	300	4	-
100ng	400	2	-
200ng	200	5	-
200ng	400	9	-
200ng	600	13	-
400ng	200	28	1

Table 4 – Detected mutations in the DNA sequence of the replaced segment in Ms6*pin*::gfp_m^{2+}.

Ms6 genome position		mutations		aa change
28 719	30bp downstream gfp_m^{2+} Stop	Insertion	C	-
28 737	13bp downstream gfp_m^{2+} Stop	Substitution	C→G	-
28 763	12bp upstream gfp_m^{2+} Stop	Substitution	A→G	Glutamic Acid→ Glycine
29 482	16bp upstream gfp_m^{2+} Start	Substitution	A→G	-
29 503	37bp upstream gfp_m^{2+}Start	Substitution	A→T	-
29 545	79bp upstream gfp_m^{2+} Start	Substitution	A→T	-

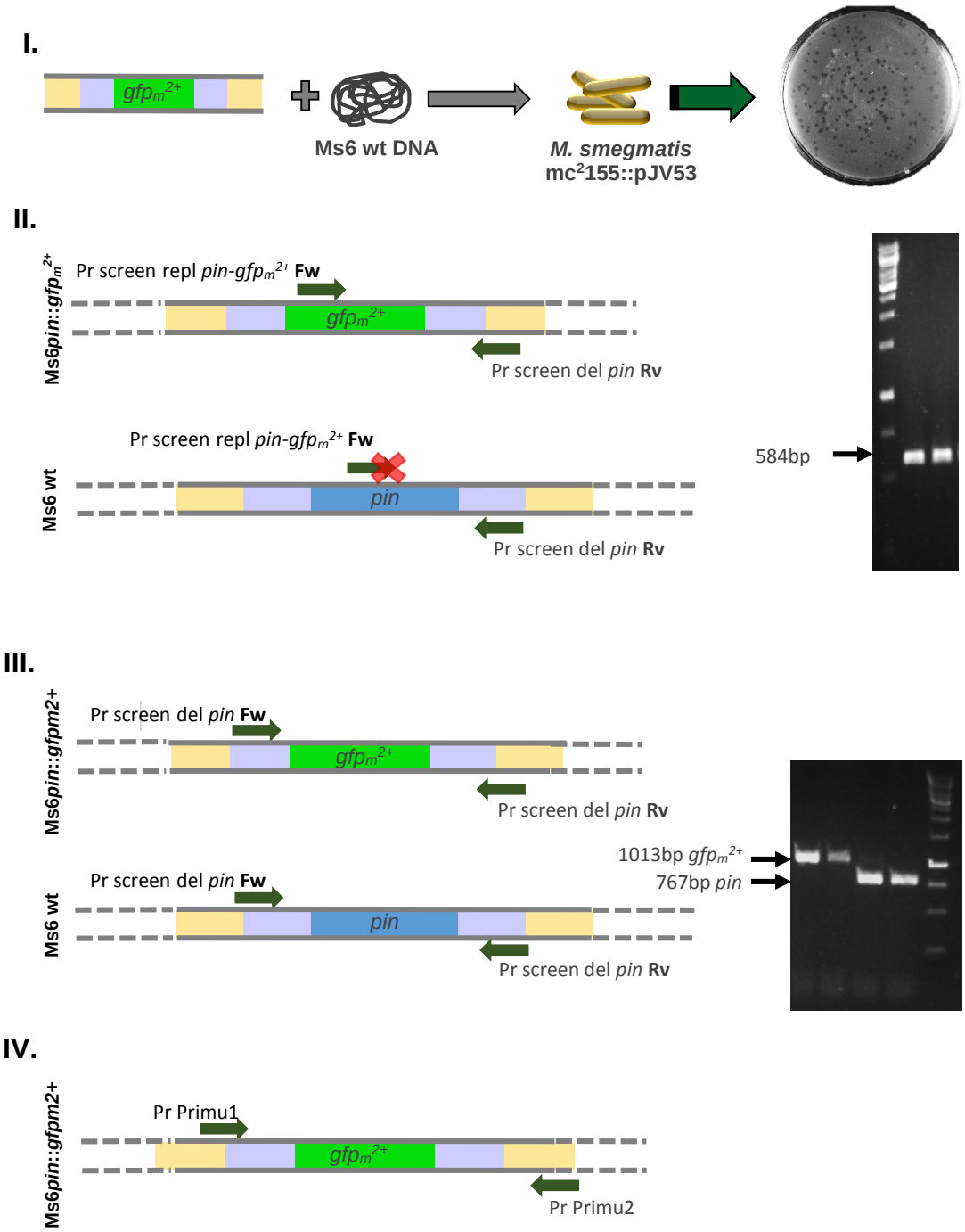

Figure 7 – Schematic representation used to construct and purify Ms6*pin*::gfp_m^{2+}. I.-The gfp_m^{2+} recombineering substrate was co-electroporated with Ms6 DNA into *M. semgmatis* mc²155::pJV53 cells. After incubation on a lawn of bacteria, the obtained phage plaques were picked and screened. II.- Screening for gfp_m^{2+} presence. In case of Ms6*pin*::gf_{pm}^{2+} presence, a 484bp fragment was amplified. On the contrary, if only Ms6 wt was present, there was absence in amplification. III.- When purified, Ms6*pin*::gfp_m^{2+} was screened to confirm *pin* inexistence and gfp_m^{2+} presence. In this screening, the control (Ms6 wt) amplified a 767bp fragment, whereas the purified mutant phage (Ms6*pin*::gfp_m^{2+}) a 1013bp fragment. IV.- The replaced region was amplified and the resulting fragment was then sequentiated.

3. Fluorescence detection of *M. smegmatis* infected with the mutant Ms6 derivatives

After successfully obtaining recombineering mutant phages, Ms6*lysB*::*egfp* and Ms6*pin*::$gfp_{m^{2+}}$, it was crucial to confirm GFP expression.

This confirmation was obtained by infecting, Ms6's host, *M. smegmatis*, with the recombineering mutant phages, in a time course assay. We followed GFP expression by both fluorescence microscopy and also spectrofluorometry. For both detection and measurements assays, a control assay was made, in which *M. smegmatis* wt was infected with Ms6 wt.

3.1. Ms6*lysB*::*egfp*

M. smegmatis was infected with Ms6*lysB*::*egfp* phage at a MOI 10 and every 30min an aliquot was collected to be observed in a fluorescence microscope. Emission of fluorescence started to be detected after 90-120min post phage adsorption, although with weak intensity, and only few cells per field (Fig.8.I.) which is in agreement with the time that LysB starts to be produced (Gil, 2011; Gigante, unpublished). At around 270-300min, an increase of bacilli's number was observed, as well as an increase in fluorescence intensity. By this time period, most bacilli observed were fluorescent (Fig.8.II.). At 330 min post infection, fluorescent cells were still observed, however bacilli were not well defined as in previous observations, which suggests that many cells have already lysed (Fig.8.III. and IV.). In addition, there was a decrease in fluorescent cells in comparison to previous time observations.

These results were confirmed by fluorescence quantifications using a spectrofluorometer. Until 210min post phage adsorption, fluorescence was relatively low. At 240min, fluorescence emission increased exponentially until the end of the assay, in agreement with the above observations. Although emission of fluorescence could be already detected by fluorescence microscopy at 90-120min post adsorption, the levels do not seem to be detected by spectrofluorometer.

It is important to mention, that plate readers identify fluorescence emission in a wavelength set, and do not distinguish fluorescent cells from surrounding fluorescence.

Top and bottom measurements, were quite similar, revealing the same pattern over time (Fig. 9). O.D.600 was relatively constant through time between ≈0.14 and 0.17, and doubled at 210min, which is in agreement with *M. smegmatis* time division

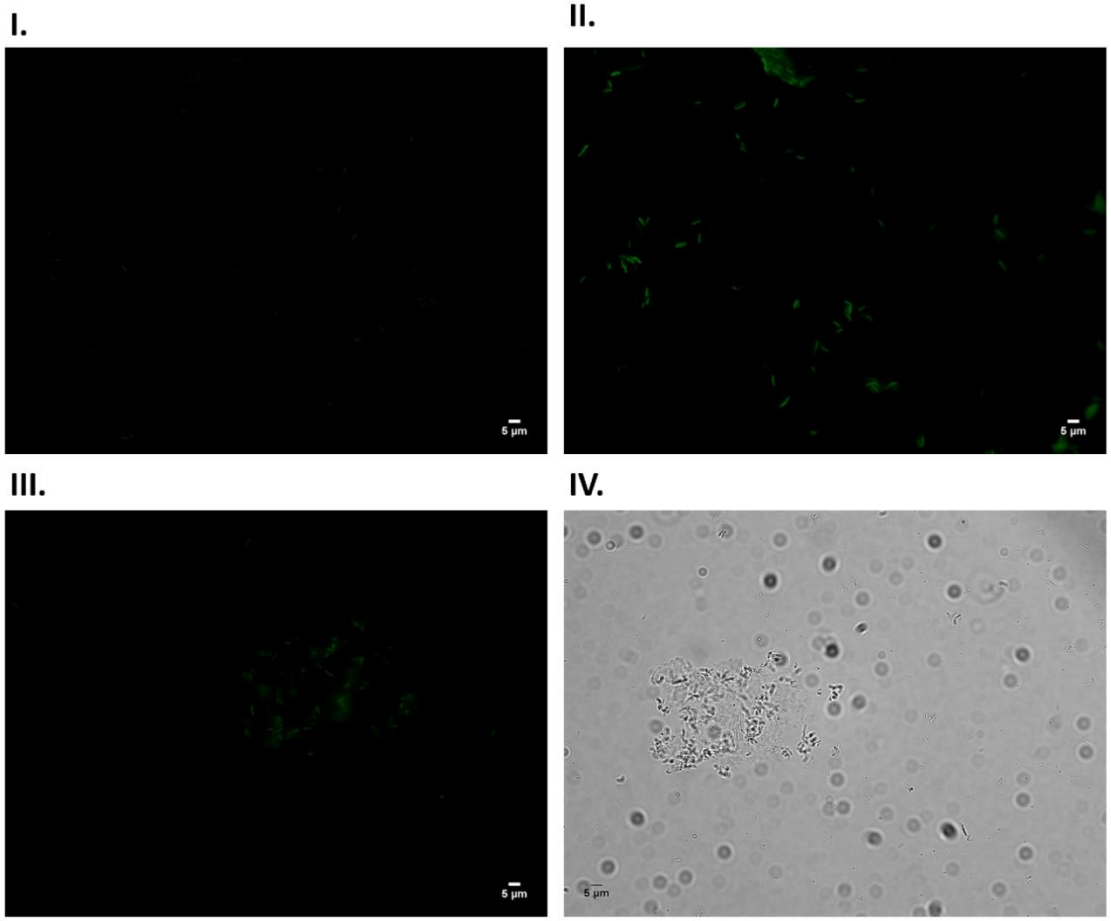

Figure 8 - Fluorescence Microscopy (63X) – *M. smegmatis* mc²155 infected with mycobacteriophage Ms6*lysB*::*egfp* after 240 min (I.), 270 min (II.) and after 330 min (III.-IV.) post adsorption at 37°C 100 rpm

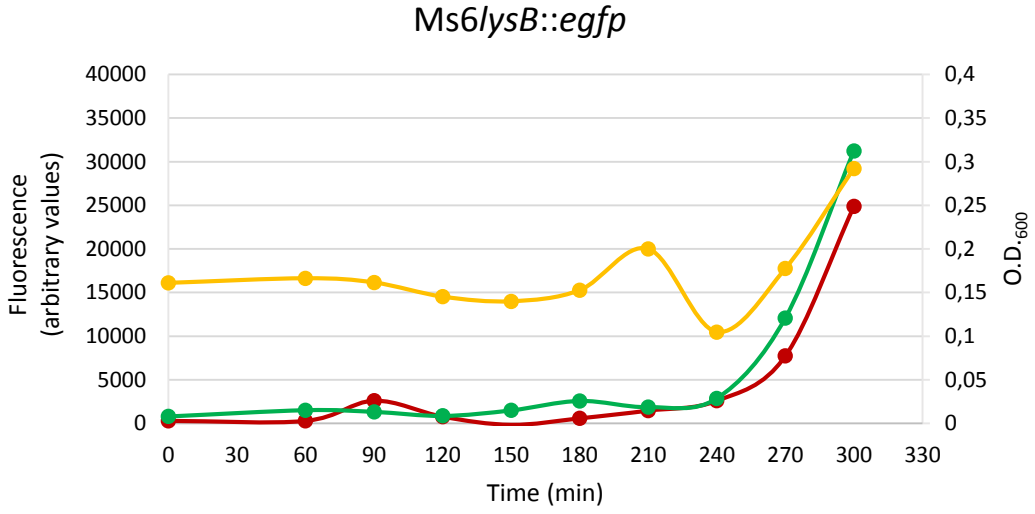

Figure 9 - Fluorescence and O.D.$_{600}$ measurements in a time course assay of *M. smegmatis* wt cells infected with Ms6*lysB*::*egfp* (MOI 10). Bottom (red) and top (green) measurements do not show any significant difference. O.D.$_{600}$ measurements are represented in yellow.

3.2. Ms6*pin*::gfp_m^{2+}

When *M. smegmatis* wt was infected with Ms6*pin*::gfp_m^{2+} at a MOI of 10 the emission of fluorescence was very low, almost undetectable by fluorescence microscopy. When MOI was increased to 100, after 30-60min post phage adsorption, fluorescence cells started to be observed. At 150min, for example, and comparing with the bright field, it is possible to state that most observed cells were fluorescent, however, some more intense than others (Fig.10. I.-IV.). At around 210-240min, most observed cells were fluorescent. Until the end of the assay, at 300min approximately, fluorescence was still observed, but the cell number in the field has decreased (Fig.10. V.-VI.). These observations were confirmed by fluorescence measurements and $O.D._{600}$ monitoring (Fig. 11). At around 180-210min it was possible to detect a more pronounced increase in fluorescence, earlier than the fluorescence observed with *M. smegmatis* wt infected with Ms6*lysB*::*egfp (*240min post phage infection). By this time the $O.D._{600}$. values began to decrease, and this reduction continued until the end of the assay. In all assays with *M. smegmatis* wt infected with Ms6*pin*::gfp_m^{2+}, even with lower $O.D._{600}$ values similar to those observed in Ms6*lysB*::*egfp* infection, this same pattern was still verified.

This results are opposite to the observed $O.D._{600}$ in *M. smegmatis* wt infected with Ms6*lysB*::*egfp,* in which, after a small decrease in $O.D._{600}$ values, an increment began to be observed.

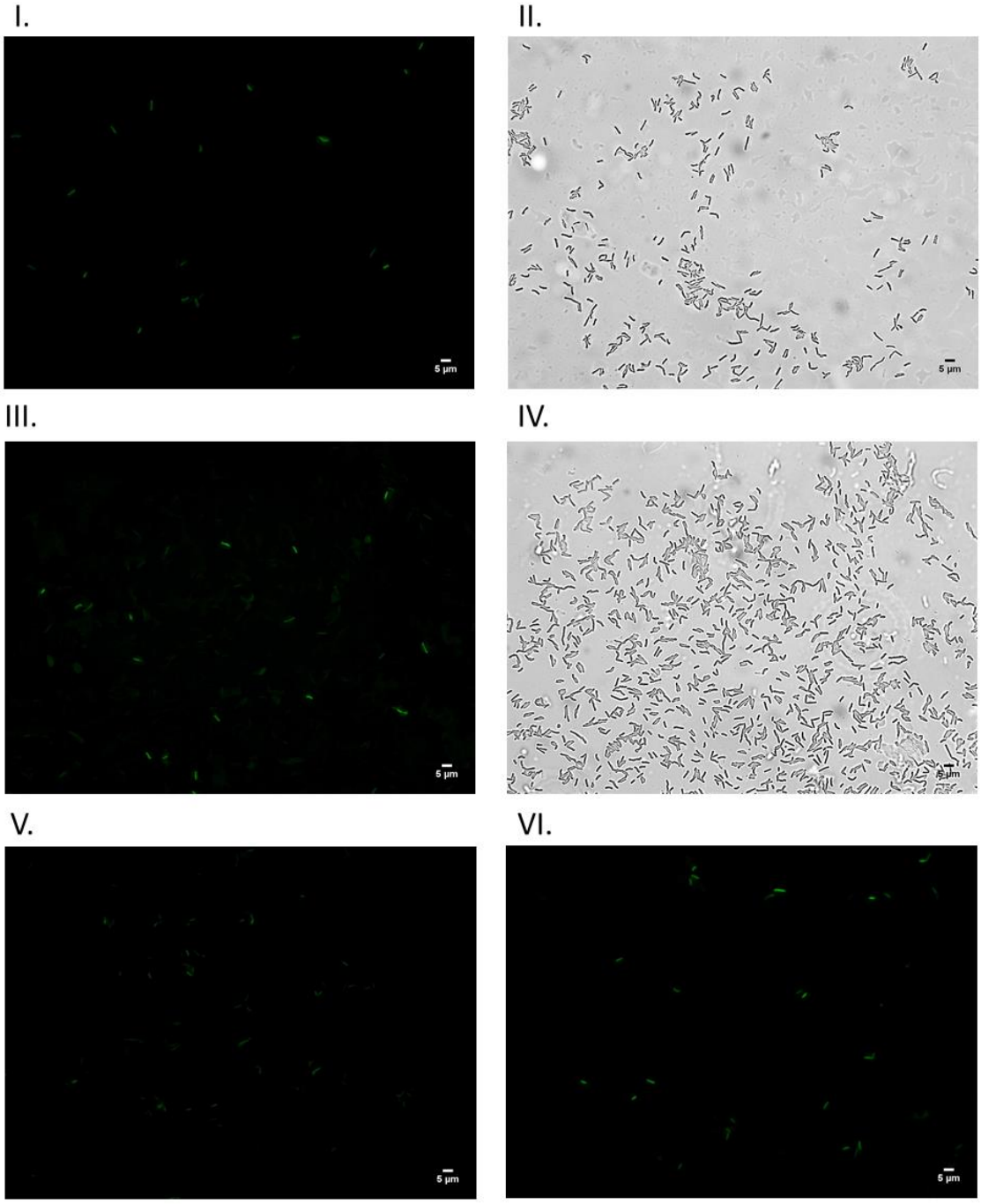

Figure 10 - Fluorescence Microscopy (63X) – *M. smegmatis* mc^2155 infected with mycobacteriophage Ms6*pin*::gfp_m^{2+} after 120min (I.-II.), 150min (III.-IV.), 270min (V.) and 300min (VI.) post adsorption at 37^{o}C 100 rpm.

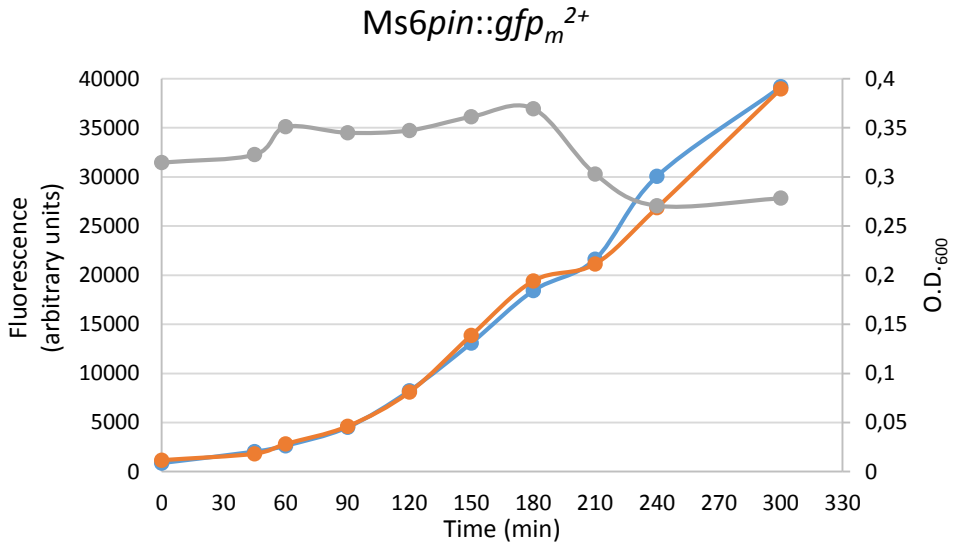

Figure 11 - Fluorescence and O.D.$_{600}$ measurements through time (post adsorption) when Ms6*pin*::*gfp*$_m^{2+}$ infects *M. smegmatis* wt cells. Bottom (blue) and top (orange) measurements do not show any significant difference. O.D.$_{600}$ measurements are represented in grey.

To confirm that expression of *gfp*$_m^{2+}$ occurs during lysogeny, we selected bacterial cells from within a phage plaque and replicated them in a solid medium. Observations under UV light revealed the presence of fluorescent colonies. Isolated colonies were then checked for lysogeny by performing an immunity test. For this, a drop of Ms6*pin*::*gfp*$_m^{2+}$ was streaked on 7H10 medium and cells were crossed perpendicularly to the phage streak (Fig. 12). Cell growth beyond the phage streak confirmed lysogeny. Under fluorescence microscopy, these cells grown in liquid medium, emitted an intense fluorescence (Fig.13). These observations confirm that Ms6 pin gene is expressed during lysogeny.

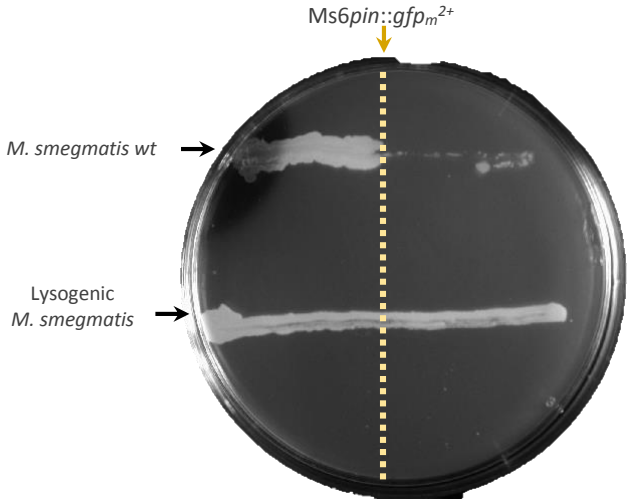

Figure 12 - Determination of *M. smegmatis* mc²155 cells sensibility to Ms6*pin*::*gfp*$_m^{2+}$ phage.

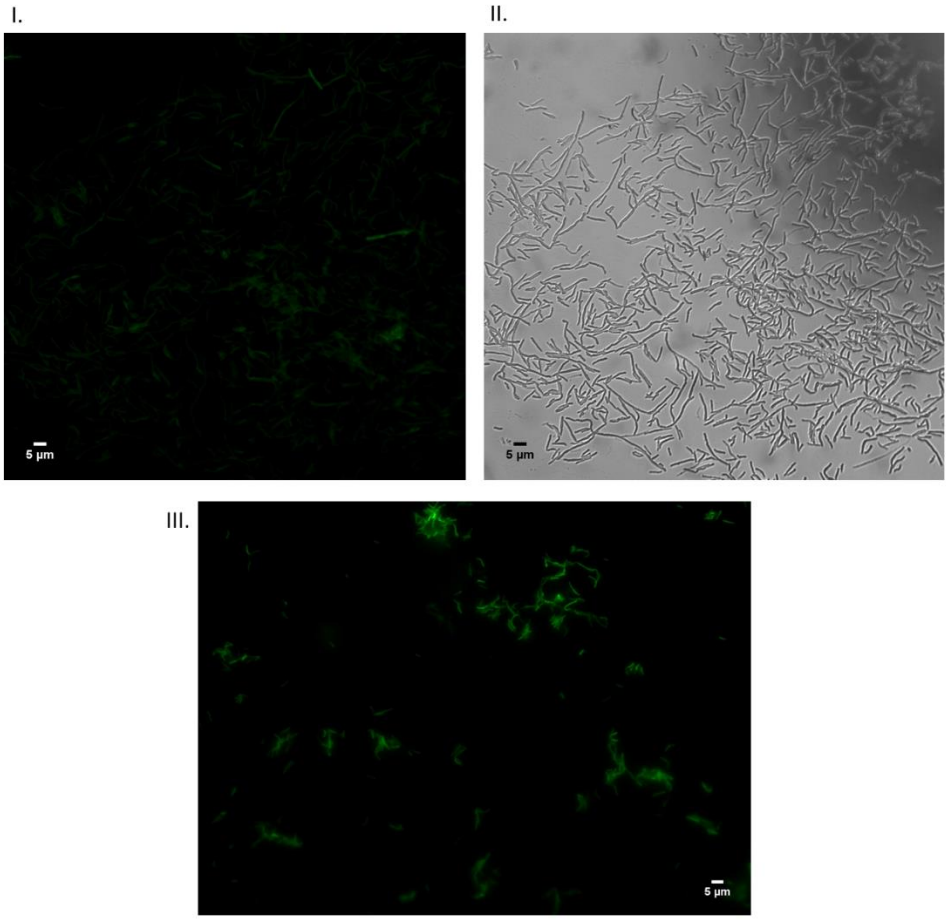

Figure 13 - Fluorescence Microscopy (63X) – *M. smegmatis* mc²155 infected with mycobacteriophage Ms6*pin*::gfp_m^{2+} in lysogenic life state. I.; II. –Mycobacteria grown in solid medium after 3 days (colonies were dissolved in liquid medium for microscopy observation); III. – Mycobacteria grown in liquid medium.

4. Transfection assay

As already mentioned, the first step of a phage infection is adorption to a bacterial receptor. Since we did not verify plaque or spot formation when Ms6 was incubated with fast- and slow-growing mycobacteria, it was hypothesized that these mycobacteria might not have Ms6 receptors or the number of phages produced were too low to allow plaque visualization. So we were interested in understanding if Ms6 is able to replicate in those mycobacteria when overcoming the first steps of infection. To achieve this we inserted the Ms6 derivative mutant DNA by transformation. If Ms6 was able to replicate in those mycobacteria by using reporter phages, we would expect to observe fluorescence emission.

For transfection of Ms6*lysB*::*egfp* or Ms6*pin*::*gfpm2+*, different DNA amounts ranging from 400 to 1600ng were tested. After electroporation, different incubation times for both fast- and slow-growing mycobacteria, were also tested. Transfected cells were analysed by

fluorescence microscopy, but fluorescence was not detected in any of those tested mycobacteria cells, including *M. smegmatis* wt.

After incubation, transfected cells were plated in *M. smegmatis* lawns in order to evaluate the possibility of phage release and consequently plaque formation. However we did not observe plaque formation, even with an additional sonication step to overcome possible lysis difficulties in transfected mycobacteria.

5. Attempts to construct Ms6*pin*::*hyg*

Again, we used BRED to construct the mutant Ms6*pin*::*hyg* where the *pin* gene was substituted by a DNA fragment containing both the hygromycin resistance (*hyg*) gene and its promoter region. A fragment of 1119bp was amplified from plasmid pSMT3 and then extended in two PCR rounds to add 100bp homology in either end of the regions flanking the *pin* gene. The first round of PCR after *hyg* amplification was the 100bp addition (50bp homology in either end flanking the target gene) (Fig.14.I). After a second 100bp extension PCR (Fig.14.II), the resultant 1320bp fragment (Fig.14.III) was purified and co-electroporated with Ms6 DNA into *M. smegmatis* mc^2155::pJV53 (Fig.14.IV). To test the best conditions to obtain a higher number of plaques, we co-electroporated different amounts of the Ms6 DNA and of the recombineering substrate. The resulted plaques were screened for *hyg* presence (Fig. 14.V). From all 43 plaques screened resultant from co-electroporations, 29 were positive for *hyg* presence together with wt phage (table 5). However, after several secondary plates made, and subsequent screenings performed, in an attempt to recover the mutant phage, we were not able to isolate it, which suggests that this construction may lead to an unstable mutant phage.

Table 5 - Amounts of Ms6 DNA and substrate used for co-electroporation, the respective number of phage plaques obtained and mutants detected in primarily screenings to obtain Ms6*pin*::*hyg*.

Amount of DNA	Amount of substrate	Number of plaques	Positive PCR screening
100ng (Control)	-	3	-
400ng (Control)	-	12	-
50ng	450	0	-
50ng	600	1	1
100ng	300	1	1
100ng	450	4	3
200 ng	450	8	4
400ng	300	29	20

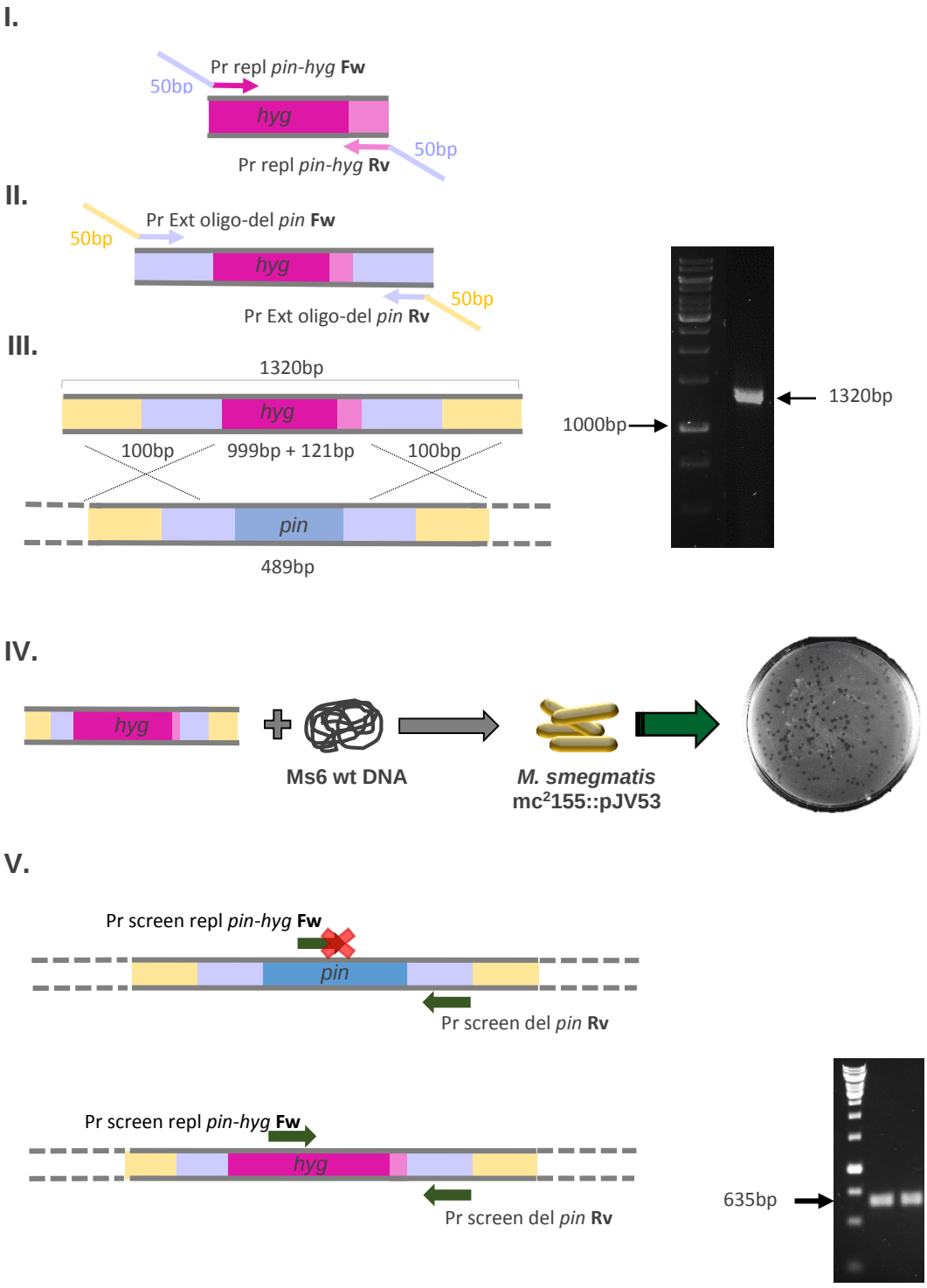

Figure 14 - Schematic representation of the strategy used for the replacement of *pin* gene for *hyg* gene and subsequent screenings. I.- First extension PCR, in which primers added 50bp to each end of the *hyg* gene and its promoter region (represented in light pink). II.- The resulted fragment was then used as template to add other 50bp to each end of the fragment. III.- The fragment obtained from the previous extension PCR had 100bp of homology to either side of the *lysB*, the target gene. IV.- This 1320bp recombineering substrate was co-electroporated with Ms6 DNA into M. semgmatis mc²155::pJV53 cells. After incubation on a lawn of bacteria, the obtained phage plaques were picked and screened. V.- In case of Ms6*pin*::*hyg* presence, a 635bp fragment was amplified. On the contrary, if only Ms6 wt was present, there was absence in amplification. We could not obtain purified Ms6*pin*::*hyg*.

V. Discussion

Ms6 was isolated in 1989 from a strain of *M. smegmatis*, and so far, the only known host is *M. smegmatis* mc^{2}155. In the current work, we were interested in analyse the ability of Ms6 to infect and replicate in other fast- and slow-growing mycobacteria. For this purpose, in a first approach, the fast-growing *M. aurum*, *M. vaccae* and *M. fortuitum*, and the slow-growing *M. bovis* BCG and *M. tuberculosis* H37Ra were incubated with Ms6, in an infection and spot test assay. After incubation at 37°C, no phage plaques or clear halos were obtained for any mycobacteria except for Ms6 host, i.e. *M. smegmatis*. However, the absence of plaque formation does not necessarily mean that the phage is unable to infect or replicate in other hosts. A phage plaque is a clearing zone in a bacterial lawn resulting from the burst of infected cells (Abedon and Yin, 2009). Plaque formation is dependent on an initial phage adsorption into the bacteria. In addition, for a plaque to be visible, there must be produced a substantial number of progeny phages in each infected cell, so certain phage plaques can be difficult to observe because of their small size due to incomplete lysis (Kropinski *et al.,* 2009).

In a host range study of 14 mycobacteriophages, Ribnyker *et al*. (2006) reported that phages D29, L5 and Bxz2 were able to form plaques on all of the slow growing species tested except for *M. marinum* and one strain of *M. scrofulaceum*. The authors reported that none of the phages tested formed plaques on two strains of *M. marinum.* In the case of TM4 and D29 phages, they suggested that this absence of plaque formation might not be due to a receptor-receptor binding protein mismatch but rather to an intracellular inhibition of phage replication, since the shuttle phasmids derived from TM4 and D29 were highly efficient in delivering the mycobacterial transposon Tn5367 to *M. marinum* (Rybniker *et al.*, 2003).

In other study, a collection of 220 sequenced mycobacteriophages were used to determine which were able to infect *M. tuberculosis* and various strains of *M. smegmatis*. The authors stated that when plating those phages with some *M. smegmatis* strains, the efficiency of plating was reduced when compared with *M. smegmatis* mc^{2}155 (Jacobs-Sera *et al.*, 2012).

Taking this information into consideration, our main goal was to construct Ms6 derivative mutants using BRED technology (Marinelli *et al.*, 2008), in which phage replication could be easily followed in mycobacteria.

GFP, is a common reporter protein to monitor bacterial cells. It is very stable, can be detected in real-time, and does not require any exogenous substrate, complex medium or expensive equipment (Errampalli *et al.*, 1999; Parish & Brown, 2009). These characteristics and the fact that GFP expression can be evaluated in live bacteria without cells lysis, made GFP an ideal reporter. So we wanted to construct a Ms6 mutant derivative phage, in which a non-essential gene was replaced by and enhanced *gfp* gene *(egfp)* which would produce a protein 35-times brighter than GFP (Zhang e*t al.*, 1996). In 2011, F.Gil constructed a Ms6 derivative phage defective for *lysB* gene (Ms6Δ*lysB*) using BRED technology, and proved that *lysB* is a non-essential gene for Ms6 viability. Considering this we decided to construct Ms6*lysB*::*egfp,* in which *lysB* gene was replaced by *egfp* gene. Since *lysB* would only be expressed in case of a lytic infection occurs, it was also necessary to construct an Ms6 derivative mutant, in which the reporter gene would be expressed in the lysogenic life cycle. M. Pimentel (1999) proved that Pin protein is a transmembrane protein with a role in a superinfection exclusion. The results of that previous study (Pimentel, 1999), lead us to consider that *pin* was a non-essential gene, and therefore could be an ideal target for gene replacement with a role in the lysogenic life cycle. Regarding this we planned the construction of Ms6*pin*::gfp_m^{2+} in which *pin* gene (489bp) was replaced by gfp_m^{2+} gene (735bp), an improved *gfp* variant in fast- and slow-growing mycobacteria (Steinhauer *et al.*, 2010).

There are few published works in which BRED technology was applied (Marinelli *et al.*, 2008; Catalão *et al.*, 2011; da Silva *et al.*, 2013). This technology has mainly been applied to construct deletions, point mutations and small insertions in mycobacteriophages with efficiencies ranging in general from 1-20% (Marinelli *et al*., 2008;da Silva *et al.*, 2013). Only two mutants constructed by DNA replacement have yet been reported (Marinelli *et al*., 2008; da Silva *et al*., 2013). Marinelli *et al.* (2008) replaced gene 54 of mycobacteriophage BP genome for a *gfp* gene, in which the final replacement substrate had 942bp. da Silva *et al.* (2013) constructed D29::P_{hsp60}-*egfp,* a D29 derivative mutant phage in which a 472bp non-essential region was replaced by 1143bp P_{hsp60}-*egfp* cassette.

Our strategy was based on gene replacements to construct Ms6*lysB*::*egfp* and Ms6*pin*::gfp_m^{2+}. Although different proportions of Ms6 wt DNA with recombineering substrates were tested, our best conditions to obtain Ms6*lysB*::*egfp* were the co-electroporation of 400ng of genomic DNA and 320ng of recombineering substrate, which resulted in a frequency of 1.12% for pure mutant phage. In the case of *pin* replacement for gfp_m^{2+}, 400ng of genomic DNA and

200ng of recombineering substrate were sufficient for a successful recombination, resulting in a frequency of 3.57% for pure mutant phage. With lower concentration of genomic DNA and higher recombineering substrate, we did not detect mutant phages on any recovered phage plaque. The BRED protocol (available at The Mycobacteriophage Database*)*, suggests for general construction of recombineering phages 50-100ng of genomic phage DNA and 100-400ng of recombineering substrate as the optimal concentrations (The Mycobacteriophage Database, 2014). da Silva *et al.* (2013), used 100ng of D29 phage DNA and 600ng-700ng of recombineering substrate, to obtain a DNA replacement in phage D29 with a percentage of mutant recovery of 1%. Our optimal conditions, were clearly different from the previous studies. This may be a consequence of the different conditions applied: phage and substrate DNA quality, competency of the electrocompetent cells, and even *M. smegmatis* mc^2155 cells used for bacterial lawns. However, these results reveal that even with different conditions, BRED is still very efficient.

Both replacements were confirmed by DNA sequencing, and mutations were not verified in Ms6*lysB*::*egfp*. On the contrary, Ms6*pin*::gfp_m^{2+} revealed one nucleotide substitution within *gfpm2+* gene, 12bp upstream of the stop codon, in addition to 4 substitutions and one insertion in the vicinity of the gfp_m^{2+} gene. These mutations, may be the result of an anomalous recombination, in which two fragments, instead of only one, were used as recombineering substrates. It should be reminded that it was not possible to isolate the single 935bp recombineering substrate, so co-electroporation was performed using two recombineering substrates, the supposed one with 935bp, and a shorter, >800bp. The reasons for the amplification of these two fragments may be due to an annealing problem of primer Fw, Pr Ext oligo-del *pin* Fw. This primer function was to add an additional 50bp homology to the 5' end of the gfp_m^{2+} recombineering substrate (see Fig.7.II. in Results). *pin* gene is localized downstream of *int* (integrase) gene, after *int* terminator (Pimentel, 1999) (Fig.16). Terminator sequences (rho-independent) contain nucleotides that form hydrogen bonds within the strand, which creates a hairpin-shaped structure (Wiley *et al.*, 2008). This 75bp extension primer, contains the terminator sequence, therefore this region might have been unstable during PCR amplification, and somehow originated a shorter fragment.

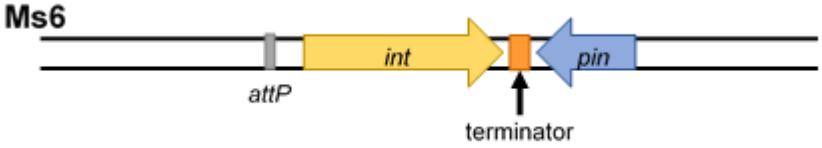

Figure 15 - Representation of *attP* site, *int* gene, terminator for *int* and *pin* gene in Ms6 genome (adapted from Pimentel, 1999).

Before the construction of phage mutants replacing *pin* gene, it was necessary to confirm that *pin* was a non-essential gene. For that purpose we planned the construction of a mutant Ms6Δ*pin*. However the construction of the 200bp recombineering substrate for *pin* deletion was not possible. The 75bp extension primer that would add 50bp homology to the 5' end of the 100nt deletion oligo, was Pr Ext oligo-del *pin* Fw, the same primer that was later used in gfp_m^{2+} recombineering substrate. Most likely, we were not able to obtain a 200bp recombineering substrate for recombination, due the same reasons explained above. Despite the verified mutations and the non-success in constructing Ms6Δ*pin*, we were successful in constructing a viable Ms6*pin*::gfp_m^{2+} phage. This also confirms that *pin* is a non-essential gene.

GFP has been used to study mechanisms such as signal transduction, host-pathogen responses, marker for bacteria detection and secreted protein localization and quantification (Awais *et al.*, 2006; Parish and Brown, 2009; Piuri *et al.*, 2009; Rondón *et al.*, 2011). The GFP fluorescence activity can be detected with minimal handling, without preparation of lysate, using a fluorescence microscope, a fluorometer, a fluorescence-activated cell sorting machine, or an imaging microplate reader (Furtado and Henry, 2002). To confirm GFP expression in Ms6*lysB*::*egfp* and Ms6*pin*::gfp_m^{2+}, *M. smegmatis* wt was infected, in separate assays, with these mutant phages. From all those available options for GFP activity detection, we chose to use fluorescence microscopy. This is the most direct way to observe fluorescence, does not require complex sample preparation, and is very reliable for visualization (Marjanovič *et al.*, 2014). After visual confirmation of GFP production, we concluded that we had two major tools to analyse Ms6 behaviour inside its host. So, before going any further, we decided to analyse GFP expression through time when Ms6*lysB*::*egfp* and Ms6*pin*::gfp_m^{2+} mutant phages infect *M. smegmatis*. *M. smegmatis* wt was infected with Ms6*lysB*::*egfp* at MOI 10. After 50min for phage adsorption, cells' suspension infected by phage was incubated for 5h30 (330min) approximately. Although fluorescent cells were detected after 90 to 120min, their number kept being very low, until 240min when it started to increase. At 270-300min the maximum fluorescence peak was observed. Fluorescence intensity started to diminish after 300min.

After constructing D29::P_{hsp60}-*egfp* phage, da Silva *et al.* (2013), used it to infect *M. smegmatis* wt at a MOI 10 during 6h. The authors observed fluorescence after 90min and stopped fluorescence observation at 180min (da Silva *et al.*, 2013). *lysB* gene is part of Ms6 lytic cassette, meaning that *lysB* will only be expressed at the end of Ms6 lytic infection, when lytic proteins are expressed to disrupt the cell wall and allow progeny phages release (Gil *et al.*, 2008). da Silva *et al.* (2013), replaced a 472bp non-essential region in D29 genome, that will be transcribed earlier than the lytic cassette. This means that the expression profile of these phages is different, preventing further comparisons.

Additionally to visualization on fluorescence microscopy we aimed to quantify emission of fluorescence using spectrofluorometry. Samples collected during our time course assay were analysed in a plate reader for fluorescence and $O.D._{600}$ measurements. Similarly to fluorescence microscopy, fluorescence started to increase at 90-120min, however with this method it continued until the end of the assay. Of note is the fact that spectrofluorometry measures total GFP fluorescence (Furtado and Henry, 2002). At the end of the assay, lysis had already occurred (after 240min) and most fluorescence detected is a consequence of EGFP released to the medium. Therefore, while with microscopy fluorescence seemed to diminish, which actually did inside *M. smegmatis* cells, the fluorescence released to the medium continued to be detected by spectrofluorometry. With this results we conclude that both methods can be used to follow lytic cycle and therefore Ms6*lysB*::*egfp* is a useful tool to use in future studies, regarding cell lysis.

Pin protein was described has having a function in a superinfection exclusion mechanism. Once a cell is infected by a phage, this mechanism prevents newly phage DNA translocations to the cytoplasm from same phage type. This mechanism functions as consequence of the existence of a prophage (Pimentel 1999). A higher MOI value leads to a higher number of infected cells. This selective pressure favours lysogeny, because the host density is too low in proportion to the phage's number, so lysogeny is less likely to drive the host to extinction (Paul and Jiang, 2001). A MOI value of 100 was needed to observe fluorescent *M. smegmatis*' cells infected with Ms6*pin*::gfp_m^{2+}. This value, 10x higher than for Ms6*lysB*::*egfp* infection, favoured lysogenic cycle in *M. smegmatis* cells.

Due to *pin'*s role, it was expected that this gene would have an early expression, together with genes necessary to follow the lysogenic cycle. This was confirmed by the observation of fluorescence 30 to 60 min post phage adsorption. By 180-210min the majority of the cells were fluorescent with an increased intensity. This was due to a higher number of cells producing GFP_m^{2+} and the accumulation of this protein in the cytoplasm. Although the number of fluorescent cells has diminished after 270min to 300min post phage adsorption, the perceived fluorescence in each cell did not seem to decrease, which seems to be due to cells that lysed while lysogenic cells keep fluorescent. These observations were again confirmed by spectrofluorometry.

Gene *pin* encodes a membrane protein, and thus it is expected that the amount of protein produced is low, in order to not be toxic to the cells. However we did not observe a reduced fluorescence in cells infected with Ms6*pin*::gfp_m^{2+} when compared with the fluorescence resulting from an infection with Ms6*lysB*::*egfp*. This might be due to the fact that GFP_m^{2+} fluorescence is higher than EGFP (Steinhauer *et al.*, 2010). In addition, the start codon of

gfp$_m^{2+}$ is an AUG while the start codon of *pin* is UUG, resulting in an increase in the protein production, since AUG is a stronger pairing with fMet-tRNA and consequently translation is more efficient (Kozak, 2005). Although previous studies on *pin* did not measured its expression in lysogeny, in this work we observed that selected lysogenic cells are fluorescent and thus *pin* is expressed during lysogeny. This means that the constructed Ms6*pin*::*gfp*$_m^{2+}$ is a good tool to follow integration of Ms6 as a prophage in mycobacteria.

In order to achieve one of the main objectives in this study, which was to understand if Ms6 is able to replicate inside tested mycobacteria, Ms6*lysB*::*gfp* or Ms6*pin*::*gfp*$_m^{2+}$ genomes were transfected into fast- and slow-growing mycobacteria. Unfortunately we did not obtained results, neither with fluorescence microscopy nor with plaques formation. Together with technical problems, this may be also due to the low electroporation efficiency of those mycobacteria. An attempt of fluorescence detection with spectrofluorometry could have been made, however, since this method proved to be sensitive to cell concentration, a cell density this high would probably lead to misleading results. Since all results from this assay were negative, Ms6 inability to follow one of it cycles inside those cells, and its consequent degradation remains a hypothesis. However Ms6's host, *M. smegmatis,* used as control, also did not presented positive results, in either approaches. These results lead us to conclude that before making any statement, further studies must be performed, in particular, an optimization of this assay.

Although we succeeded in the construction of the fluorescent reporter phages, our attempts to obtain Ms6 derivative containing the hygromycin resistance cassette were unsuccessful. It was not possible to purify Ms6*pin*::*hyg* even with 29 of 43 screened plaques being positive for mutant phage presence, and after performing several subsequent screenings. The large size (1320bp) of the substrate to replace a gene with approximately half its length (489bp), might have been one reason for Ms6*pin*::*hyg* poor viability. However, da Silva *et al*. (2013) successfully replaced a 472bp non-essential region for 1143bp P$_{hsp60}$-*egfp* cassette, in which was added 200bp of homology to either side of the cassette, resulting in a recombineering substrate of 1343bp. So, large replacements or insertions should be possible, provided that downstream genes expression is not impeded (Marinelli *et al.*, 2008).

Ms6*pin*::*hyg* was constructed as an alternative for mutant phages with *gfp* reporter genes, as an easier way to determine the ability of Ms6 to infect and/or replicate in other mycobacteria. If this scenario happened, colonies resistant to hygromycin would grow in 7H10 supplemented with this antibiotic. If Ms6 was not viable inside those mycobacteria, colonies formation would be absent except for the Ms6 host, *M. smegmatis* wt.

VI. Concluding Remarks and Future Perspectives

For the past decade, bacteriophage research has been going through a renaissance, mainly due to the prospect of their use in phage therapy, food and biotechnology industries. It is essential to go beyond the initial discovery of phage resistance system to try to understand the molecular mechanisms behind all these antimicrobial activities. In this regard, progress in phage biology is undoubtedly needed to fully comprehend bacteriophages, and in this particular case, mycobacteriophage Ms6 biological systems.

The aim of this project was to **construct Ms6 derivative mutants** and use them as tools in a way that would contribute to a better understanding of how Ms6 interacts with mycobacteria.

BRED technology revealed, once again, to be a very good alternative to the existing cloning systems to generate mutations in mycobacteriophages. Even with substantial differences, in particular to co-electroporation amounts and resulting phage plaques, **we successfully constructed and purified two recombineering mutant phages, Ms6*lysB*::*egfp* and Ms6*pin*::*gfp*$_m^{2+}$** with frequencies of 1.12% and 3.57% respectively.

Ms6*lysB*::*egfp* and Ms6*pin*::*gfp*$_m^{2+}$ are two important tools because they have reporter genes replacing non-essential genes, being expressed in Ms6 lytic and lysogenic cycle respectively. These mutant phages, in particular Ms6*pin*::*gfp*$_m^{2+}$, helped us to understand the expression profiles of the replaced genes in *M. smegmatis.*

We proved that when *M. smegmatis* is infected by Ms6*pin*::*gfp*$_m^{2+}$ and it follows lysogenic cycle, GFP_m^{2+} is expressed, and therefore we also proved that *pin* gene is expressed in lysogenic cells. In addition fluorescence remains active for at least a few days.

The use of this two phages in future studies will certainly help unveil some questions regarding Ms6 behaviour inside its host.

However, we were not able to evaluate Ms6*lysB*::*egfp* and Ms6*pin*::gfp_m^{2+} inside other mycobacteria. We did not find the answer to the question regarding the presence of Ms6 attachment receptor in the tested mycobacteria.

Given the answers we obtained in this study and all the questions that still remain to be answered, our future perspectives for the continuation of this project should focus on:

1. Understand if Ms6 is able to follow the lytic or the lysogenic cycle inside fast- and slow-growing mycobacteria;
2. Optimize the transfection protocol with Ms6*lysB*::*egfp* and Ms6*pin*::gfp_m^{2+}*;*
3. Use new strategies to construct a Ms6 mutant expressing an antibiotic resistance.

Our findings would contribute to the understanding of mycobacteriophage's biology, in particular Ms6, and also may be a rich potential for further contributions to develop new tools to study mycobacteria genetics.

References

Ackermann, H.-W., 2006. Classification of bacteriophages, pp. 8–16. In: Calendar, R. (ed), The Bacteriophages, Oxford University Press, New York.

Ackermann, H-W., 2009. Phage classification and characterization, pp.127-140. In: Clokie M.R.J., Kropinski A.M. (ed), Bacteriophages. Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions, Methods in Molecular Biology, 127–140, Humana Press, New York.

Ackermann, H., 2011. Bacteriophage taxonomy. Microbiology Australia. 90–94.

Awais, R., Fukudomi, H., Miyanaga, K., Unno, H., Tanji, Y., 2006. A recombinant bacteriophage-based assay for the discriminative detection of culturable and viable but nonculturable *Escherichia coli* O157:H7. Biotechnology Progress. **22:** 853–859.

Barsom, E.K., Hatfull, G.F., 1996. Characterization of *Mycobacterium smegmatis* gene that confers resistance to phages L5 and D29 when overexpressed. Molecular Microbiology. **21:** 159-170.

Catalão, M.J., Gil, F., Moniz-Pereira, J., Pimentel, M., 2010. The mycobacteriophage Ms6 encodes a chaperone-like protein involved in the endolysin delivery to the peptidoglycan. Molecular Microbiology. **77**: 672–686.

Catalão, M.J., Milho, C., Gil, F., Moniz-Pereira, J., Pimentel, M., 2011a. Functional analysis of the Holin-Like Proteins of Mycobacteriophage Ms6. Journal of Bacteriology. **193:** 2793-803.

Catalão, M.J., Milho, C., Gil, F., Moniz-Pereira, J., Pimentel, M., 2011b. A second endolysin gene is fully embedded in-frame with the *lysA* gene of mycobacteriophage Ms6. PLoS One 6, e20515.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., Prasherf, D.C., 1994. Green Fluorescent Protein as a Marker for Gene Expression. Science. **263:** 802–804.

Chen, J., Kriakov, J., Singh, A., Jacobs, W.R. Jr, Besra, G.S., Bhatt, A., 2009. Defects in glycopeptidolipid biosynthesis confer I3 resistance in *Mycobacterium smegmatis*. Microbiology. **155**: 4050-4057.

Chopin, M.-C., Chopin, A., Bidnenko, E., 2005. Phage abortive infection in lactococci: variations on a theme. Current Opinion in Microbiology. **8:** 473–9.

Cormack, B.P., Valdivia, R.H., Falkow, S., 1996. FACS-optimized mutants of the green fluorescent protein (GFP). Gene.**173:** 33–38.

da Silva, J.L., Piuri, M., Broussard, G., Marinelli, L.J., Bastos, G.M., Hirata, R.D.C., Hatfull, G.F., Hirata, M.H., 2013. Application of BRED technology to construct recombinant D29 reporter phage expressing EGFP. FEMS Microbiology Letters. **344**: 166–72.

Ellis, E.L., Delbrück, M., 1939. The growth of Bacteriophage. Journal of General Physiology. 365–384.

Errampalli, D., Leung, K., Cassidy, M.B. Kostrynska, M., Blears, M., Lee, H., Trevors, J.T., 1999. Applications of the green fluorescent protein as a molecular marker in environmental microorganisms. Journal of Microbiological Methods. **35:** 187-199.

Folimonova, S.Y., 2012. Superinfection exclusion is an active virus-controlled function that requires a specific viral protein. Journal of Virology. **86:** 5554–55561.

Freitas-vieira, A., Anes, E., Moniz-Pereira, J., 1998. The site-specific recombination locus of mycobacteriophage Ms6 determines DNA integration at the $tRNA^{Aia}$ gene of *Mycobacterium spp.* Microbiology. **144**: 3397–3406.

Furtado, A., Henry, R., 2002. Measurement of green fluorescent protein concentration in single cells by image analysis. Analytical Biochemistry. **310:** 84–92.

Furuchi, A., Tokunaga, T., 1972. Nature of the receptor substrance of *Mycobacterium smegmatis* for D4 bacteriophage adsorption. Journal of bacteriology. **111:** 404-411.

Garcia, M., Pimentel, M., Moniz-Pereira, J., 2002. Expression of Mycobacteriophage Ms6 Lysis Genes is driven by two σ^{70}-Like Promoters and is dependent on a Transcription Termination Signal present in the Leader RNA. Journal of Bacteriology. **184:** 3034-3043.

Garcia-Doval, C., van Raaij, M.J., 2013. Bacteriophage Receptor Recognition and Nucleic Acid Transfer, pp. 489-518. In: Mateu, G. M. (ed), Structure and Physics of Viruses, An integrated Textbook. Springer, Netherlands.

Gil, F., Catalão, M.J., Moniz-Pereira, J., Leandro, P., McNeil, M., Pimentel, M., 2008. The lytic cassette of mycobacteriophage Ms6 encodes an enzyme with lipolytic activity. Microbiology. **154:** 1364–71.

Gil, A. Functional Analysis of *lysB* gene from the lysis module of mycobacteriophage Ms6. 2011. PhD thesis, Universidade de Lisboa.

Hatfull, G.F., 2000. Molecular Genetics of Mycobacteriophages, pp. 37-54. In: Hatfull, G.F., Jacobs Jr., W.R., Molecular Genetics of Mycobacteria. American Society for Microbiology, Washington DC.

Hatfull, G.F., Cresawn, S.G.., Hendrix, R.W., 2008. Comparative genomics of the mycobacteriophages: Insights into bacteriophage evolution. Research in Microbiology. **159:** 332–339.

Hatfull, G.F., 2010. Mycobacteriophages: Genes and genomes. Annual Review of Microbiology. **64**: 331-356.

Hatfull, G.F., Hendrix, R.W., 2011. Bacteriophages and their Genomes. Current Opinion in Virology. **1**: 298–303.

Hatfull, G.F., 2014. Molecular Genetics of Mycobacteriophages, pp.81-133. In: Hatfull, G.F., Jacobs Jr., W.R., Molecular Genetics of Mycobacteria (2^{nd} Edition). American Society for Microbiology, Washington DC.

Hett, E.C., Rubin, E.J., 2008. Bacterial growth and cell division: a mycobacterial perspective. Microbiol. Molecular Biology Reviews. **72**: 126–56.

Horvath, P., Barrangou, R., 2010. CRISPR/Cas, the Immune System of Bacteria and Archaea. Science. **327:** 167–70.

Hyman, P., Abedon. S.T., 2010. Bacteriophage host range and bacterial resistance. Advances in Applied Microbiology. **70**: 217-248.

Jacobs-Sera, D., Marinelli, L., Bowman, C., Broussard, G., Guerrero, C., Boyle, M., Petrova, Z., Dedrick, R., Pope, W., SEA-PHAGES, Modlin, R.L., Hendrix, R.W., Hatfull, G.F., 2012. On the nature of mycobacteriophage diversity and host preference. Virology. **434**: 187–201.

Khoo, K., Suzuki, R., Dell, A., Morris, H.R., Mcneil, M.R., Brennan, P.J., Besra, G.S., Biochemie, J., 1996. Chemistry of the Lyxose-Containing Mycobacteriophage Receptors of *Mycobacterium phlei/Mycobacterium smegmatis*. Biochemistry. **35**: 11812–11819.

Kozak, M., 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene**. 361**: 13–37.

Kropinski, A.M., Mazzocco, A., Waddell, T.E., Lingohr, E., Johnson, R.P., 2009. Enumeration of Bacteriophages by Double Agar Overlay Plaque Assay, pp.69-76. In: Clokie, M.R.J., Kropinski, A.M. (ed), Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions, Humana Press, New York.

Kutter, E., Raya, R., Carlon, K., 2005. Molecular Mechanisms of Phage Infection, pp.165-222. In: Kutter, E., Sulakvelidze, A. (ed), Bacteriophages – Biology and Applications. CRC Press, Boca Raton, Florida.

Kutter, E., 2009. Phage host-range and efficiency of plating, pp.141-149. In: Clokie M.R.J., Kropinski A.M. (ed), Bacteriophages. Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions, Methods in Molecular Biology, 127–140, Humana Press, New York.

Levin, B.R., Bull, J.J., 2004. Population and evolutionary dynamics of phage therapy. Nature in Reviews Microbiology **2:** 166-173.

McGrath, S., Fitzgerald, G. F., van Sinderen, D., 2002. Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. Molecular Microbiology. **43:** 509–520.

Maniloff, J., 2012. Bacteriophages, in: eLs - Encycplopedia of Life Sciences. John Wiley & Sons, Ltd, Chichester, UK.

Marinelli, L.J., Piuri, M., Swigoňová, Z., Balachandran, A., Oldfield, L.M., van Kessel, J.C., Hatfull, G.F., 2008. BRED: A simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. PLoS One 3, e3957.

Marinelli, L.J., Hatfull, G.F., Piuri, M., 2012. Recombineering A powerful tool for modification of bacteriophage genomes. Bacteriophage. **2:** 5–14.

Marjanovič, I., Kandušer, M., Miklavčič, D., Keber, M.M., Pavlin, M., 2014. Comparison of Flow Cytometry, Fluorescence Microscopy and Spectrofluorometry for Analysis of Gene Electrotransfer Efficiency. Journal of Membrane Biology. **247:** 1259-1267.

Monk, A.B., Rees, C.D., Barrow, P., Hagens, S., Harper, D.R., 2010. Bacteriophage applications: where are we now? Letters in Applied Microbiology. **51:** 363-369.

O'Gaora, P. 1998 Expression of genes in mycobacteria, pp. 261–273. In: Parish T, Stoker NG (ed), Mycobacteria Protocols. Humana Press, New Jersey.

Parish, T., Brown, A.C., 2009. Mycobacteria Protocols, Methods in Molecular Biology. Humana Press, New Jersey.

Paul, J.H., Jiang, S.C., 2001. Lysogeny and Transduction, pp.105–125. In: Paul, J. (ed) , Methods in Microbiology, (vol. 30). Academic Press, San Diego.

Pavelka Jr., M.S., 2000. Genetics of Mycobacterial Metabolism, pp. 221-234. In: Hatfull, G.F., Jacobs Jr., W.R., Molecular Genetics of Mycobacteria. American Society for Microbiology, Washington DC.

Pelicic, V., Reyrat, J., 1998. Genetic advances for studying *Mycobacterium tuberculosis* pathogenicity. Molecular Microbiology. **28:** 413–420.

Pimentel, M., 1999. Biologia Molecular da Resistência à Superinfecção por Bacteriófagos: Estudo do Gene *pin* do Micobacteriófago Ms6. PhD Thesis. Universidade de Lisboa

Pimentel, M., 2014. Genetics of phage lysis. Microbiology Spectrum. **2:** 1–13.

Piuri, M., Jacobs, W.R., Hatfull, G.F., 2009. Fluoromycobacteriophages for rapid, specific, and sensitive antibiotic susceptibility testing of *Mycobacterium tuberculosis*. PLoS One 4, e4870.

Portugal, I., Anes, E., Moniz-Pereira, J., 1989. Temperate Mycobacteriophage *from M. smegmatis*. Acta Leprogica **1:** 243–244.

Rakhuba, D.V., Kolomiets, E.I., Szwajcer Dey, E., Novik, G.I., 2010. Bacteriophage receptors, Mechanisms of Phage Adsorption and Penetration into Host Cell. Polish Journal of Microbiology. **59:** 145–155.

Randall-Hazelbauer, L., Schwartz, M., 1973. Isolation of the bacteriophage lambda receptor from *Escherichia coli*. Journal of Bacteriology. **116:** 1436–14

Rastogi, N., Legrand, E., Sola, C., 2001. The Mycobacteria: an introduction to nomenclature and pathogenesis. Review in Science Techniques. **20:** 21–54.

Rondón, L., Piuri, M., Jacobs, W.R., De Waard, J., Hatfull, G.F., Takiff, H.E., 2011. Evaluation of fluoromycobacteriophages for detecting drug resistance in *Mycobacterium tuberculosis*. Journal of Clinical Microbiology. **49:** 1838–1842.

Rybniker, J., Wolke, M., Haefs, C., Plum, G., 2003. Transposition of Tn 5367 in *Mycobacterium marinum*, Using a Conditionally Recombinant Mycobacteriophage. Journal of Bacteriology. **185:** 1745–1748.

Rybniker, J., Kramme, S., Small, P.L., 2006. Host range of 14 mycobacteriophages in *Mycobacterium ulcerans* and seven other mycobacteria including *Mycobacterium tuberculosis*--

application for identification and susceptibility testing. Journal of Medical Microbiology. **55:** 37–42.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

São-José, C., Baptista, C., Santos, M. A, 2004. *Bacillus subtilis* operon encoding a membrane receptor for bacteriophage SPP1. Journal of Bacteriology. **186:** 8337–46.

São-José, C., Lhuillier, S., Lurz, R., Melki, R., Lepault, J., Santos, M.A., Tavares, P., 2006. The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of bacteriophage SPP1 DNA. Journal of Biological Chemistry. **281:** 11464–70.

Shimomura, O., Johson, F.H., Saiga, Y., 1962. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. Journal of Cellular and Comparative Physiology. **59:** 223–239.

Skurnik, M., Strauch, E., 2006. Phage therapy: facts and fiction. International Journal of Medical Microbiology. **296:** 5–14.

Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., Jr, W.R.J., 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Molecular Microbiology. **4:** 1911-1919.

Song, H., Sandie, R., Wang, Y., Andrade-Navarro, M. A, Niederweis, M., 2008. Identification of outer membrane proteins of *Mycobacterium tuberculosis*. Tuberculosis. **88:** 526–44.

Steinhauer, K., Eschenbacher, I., Radischat, N., Detsch, C., Niederweis, M., Goroncy-Bermes, P., 2010. Rapid evaluation of the mycobactericidal efficacy of disinfectants in the quantitative carrier test EN 14563 by using fluorescent *Mycobacterium terrae*. Applied Environmental Microbiology. **76:** 546–54.

van Kessel, J.C., Hatfull, G.F., 2007. Recombineering in *Mycobacterium tuberculosis*. Nature Methods. **4:** 147–52.

van Kessel, J.C., Marinelli, L.J., Hatfull, G.F., 2008. Recombineering mycobacteria and their phages. Nature Reviews in Microbiology. **6:** 851–7.

Willey, J.M., Sherwood, L.M., Woolverton, C.J., 2008. Prescott, Harley and Klein's Microbiology. McGraw Hill, New York.

Young, F.E., 1967. Requirement of Glucosylated Teichoic Acid for adsorption of phage in *Bacillus subtillis* 168. PNAS. **58:** 2377–2384.

Zhang, G., Gurtu, V., Kain, S.R., 1996. An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. Biochemical Biophysical Research Communications. **227:** 707–11.

World Heath Organization (WHO), 2014

- http://www.who.int/campaigns/tb-day/2013/event/en/;
- http://www.who.int/tb/publications/factsheet_global.pdf.

The Mycobacteriophage Database, 2014:http://phagedb.org/

Appendix I

List of oligonucleotides and primers used throughout this study.

Description	Name	Sequence 5´- 3’
Amplification of *egfp* gene from pEGFP plasmid Fw	Pr *egfp* 5' Fw	GTGAGCAAGGGCGAGGAGCTGTTCA
Amplification of *egfp* gene from pEGFP plasmid Rv	Pr *egfp* 3' RV	CTTGTACAGCTCGTCCATGCCGAGA
***egfp* gene+50bp of homology with Ms6 Fw – first extension**	Pr repl *egfp*/*lysB* Fw1	GGCGGAAAAACCCTCGTGGACGCGGTA GCAGAACTGTTGGGCCACTGATGGTGA GCAAGGGCGAGGAGCTGTTCA
***egfp* gene+50bp of homology with Ms6 Rv – first extension**	Pr repl *egfp*/*lysB* Rv1	CGCTCGGCGGCGTCGATCCAGAATGAA CGTGTCAGCATGGGTGCCTCCTACTTGT ACAGCTCGTCCATGCCGAGA
***egfp* gene+50bp of homology with Ms6 Fw – second extension**	Pr repl *egfp*/*lysB* Fw2	GATCCTGCGGCAACTGCGCGGATACAA CCTCACTGGCTGGCCGCAGCTCGGCGG AAAAACCCTCGTGGACGCGG
***egfp* gene+50bp of homology with Ms6 Rv – second extension**	Pr repl *egfp*/*lysB* Rv2	CGGCGCCGAGGGTGGCGATCGCGGTTT GGGCGAATGTGCGTATGGCACGCTCGG CGGCGTCGATCCAGAATG
Hybridize inside *lysA* gene, before *lysB* or *egfp* beginning Fw	Pr *lysA* 3' Fw sreen	GGACGACTGGAGTGACCGCGAAATC
Hybridize inside *egfp* gene Rv	Pr screen repl lysB/egfp Rv	GTCGGGGTAGCGGCTGAAGCACTG
Hybridize inside *lysB* gene Fw	Pr *lysB* int	GCAGGAATCACCGAGGCGCACAAC
Hybridize at the end of *lysB* gene	*lysB* HindIII 3'	CTCCTATGAAGCTTGCGTAGGTAGTC
Hybridize after the end of *egfp* Rv	Pr ORF4 3'HindIII	CGCAAGCTTTGCGGTGGTGACCGC
oligo 100nt for deletion of *pin* gene	Oligo 100nt del *pin*	TACCAGGGGATCTAGATCCCTGGTAGG GCGCCTTTTTGTGTTTGCGGACCACGTG ACGCACTGTCGGTTATCTAATCGTAATA TTCCCATTTGTGGGCTT
Extension of deletion oligo Fw or of replacement oligo Fw – first extension	Pr Ext oligo-del *pin* Fw	CGGAGTTTATGGACGGATTGTTGGGGG ACTTTTAAGACCCAGATGCGCCCTACCA GGGGATCTAGATCCCTGGTA
Extension of deletion oligo Rv Or of replacement oligo Rv – first extension	Pr Ext oligo-del *pin* Rv	CGCAACCGGATGCGTAGACGCTAACGG ATCGTTGCCAGGATCGACACACGAAGC CCACAAATGGGAATATTACGA
screening for *pin* deletion or *pin* replacement Fw	Pr screen del *pin* Fw	CACGGATGTGGATCGGACGAGTTCG
screening for *pin* deletion or *pin* replacement Rv	Pr screen del *pin* Rv	CATGGGCTGCACCATCCGAAATGAAC
Amplification of *gfpm2+* gene from pMN437 plasmid Fw	Pr *gfpm++* Fw	GCAGAAAGGAGGTTAATAATGTCGAAG
Amplification of *gfpm2+* gene from pMN437 plasmid Rv	Pr *gfpm++* Rv	CTACTTGTACAGCTCGTCCATGCCG
***gfpm2+* gene+50bp of homology with Ms6 Fw – First extension**	Pr repl *pin-gfpm2+* Fw	TACCAGGGGATCTAGATCCCTGGTAGG GCGCCTTTTTGTGTTTGCGGACCCTACT TGTACAGCTCGTCCATGCCG

***gfpm2+* gene+50bp of homology with Ms6 Rv – First extension**	Pr repl *pin-gfpm2+* Rv	AAGCCCACAAATGGGAATATTACGATT AGATAACCGACAGTGCGTCACGTGCAG AAAGGAGGTTAATAatgtcga
Screening for *gfpm2+*replacement Fw. Hybridize inside *gfpm2+* gene	Pr screen repl *pin-gfpm2+* Fw	CCAGCTTGTGGCCCAGGATGTTACCG
Hybridize before *gfpm2+* STOP codon	Pr Primu1	CACGTACACGGATGTGGATCGG
Hybridize after *gfpm2+* Shine-Dalgarno added sequence	Pr Primu2	AAGACCAGGAGGAAGCCTTAGG
Amplification of *gfpm2+* gene from pSMT3 plasmid and addition of Shine-Dalgarno sequence Fw	Pr *hyg* pSMT3 Fw	CTCATCAGGCGCCGGGGGCGGTG
Amplification of *gfpm2+* gene and its promoter region from pSMT3 plasmid Rv	Pr *hyg* pSMT3 Rv	CGTCCCCGACGTGGCCGACCAGC
***hyg* gene+50bp of homology with Ms6 Fw – First extension**	Pr repl *pin-hyg* Fw	CTACCAGGGGATCTAGATCCCTGGTAG GGCGCCTTTTTGTGTTTGCGGACCTCAT CAGGCGCCGGGGGCGGTGT
***hyg* gene+50bp of homology with Ms6 Rv – First extension**	Pr repl *pin-hyg* Rv	AGAGTGTCGAGATCGTCATCAAACATG GGCTGCACCTACCGAAATGAACACGTC CCCGACGTGGCCGACCAGCCC
***hyg* gene+50bp of homology with Ms6 Rv – second extension**	Pr ext oligo repl *pin-hyg* Rv	ACCACATTGTTGGCGGGCACTCGTCCAT CGCGTTCAAAATCCGCGCCAGCAGAGT GTCGAGATCGTCATCAAACA
Screening for *hyg* replacement. Hybridize inside *hyg* gene	Pr screen repl *pin-hyg* Fw	TGTTCCCGGTCAGCGGCACC

Appendix II

Sequenced region encompassing the site of *gfpm2+* replacement for *pin* gene:

>GGACGGATTGTTGGGGGACTTTTAAGACCCAGATGCGCCCTACCAGGGGATCTAGATCCCCC
TGGTAGGGCGCCTTTTTCTGTTTGCGGACCCTACTTGTACAGCCCGTCCATGCCGTGGGTGATGCCGGC
GGCGGTGACGAACTCCAGCAGGACCATGTGGTCGCGCTTCTCGTTCGGGTCCTTGGACAGGGCGGACT
GGGTGGACAGGTAGTGGTTGTCGGCAGCAGGACCGGGCCGTCGCCGATCGGGGTGTTCTGCTGGTAG
TGGTCGGCTAGCTGGACGCCACCGTCCTCGATGTTGTGGCGGGTCTTGAAGTTGGCCTTGATGCCGTTC
TTCTGCTTGTCCGCGGTGATGTAGACGTTGTGGGAGTTGTAGTTGTACTCCAGCTTGTGGCCCAGGATG
TTACCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTGACCAGGGTGTCGCCCTCGAACTTG
ACCTCGGCACGCGTCTTGTAGTTACCGTCGTCCTTGAAGGAGATGGTGCGCTCCTGGACGTAACCCTCC
GGCATGGCGGACTTGAAGAAGTCGTGGCGCTTCATGTGGTCCGGGTAGCGGGAGAAGCACTGGACGC
CGTAGGTCAGGGTGGTGACCAGGGTCGGCCACGGGACCGGCAGCTTACCGGTGGTGCAGATGAACTT
CAGGGTCAGCTTGCCGTAAGTGGCGTCGCCCTCACCCTCGCCGGAGACGGAGAACTTGTGGCCGTTGA
CGTCACCGTCCAGCTCGACCAGGATCGGGACGACGCCGGTGAACAGCTCCTCGCCCTTCGACATTATTA
ACCTCCTTTCCGCACGTGACGCACTGTCGGTAATCTAATCGTAATATTCCCATTTGTGGGCTTCGTGTGT
CGAACCTGGCAACGATCCGTTAGCGTCTACGCATCCGGTTGCGA

Original region encompassing the site of *gfpm2+* replacement for *pin* gene:

>GGACGGATTGTTGGGGGACTTTTAAGACCCAGATGCGCCCTACCAGGGGATCTAGATCCCT
GGTAGGGCGCCTTTTTGTGTTTGCGGACCCTACTTGTACAGCTCGTCCATGCCGTGGGTGATGCCGGCG
GCGGTGACGAACTCCAGCAGGACCATGTGGTCGCGCTTCTCGTTCGGGTCCTTGGACAGGGCGGACTG
GGTGGACAGGTAGTGGTTGTCCGGCAGCAGGACCGGGCCGTCGCCGATCGGGGTGTTCTGCTGGTAG
TGGTCGGCTAGCTGGACGCCACCGTCCTCGATGTTGTGGCGGGTCTTGAAGTTGGCCTTGATGCCGTTC
TTCTGCTTGTCCGCGGTGATGTAGACGTTGTGGGAGTTGTAGTTGTACTCCAGCTTGTGGCCCAGGATG
TTACCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTGACCAGGGTGTCGCCCTCGAACTTG
ACCTCGGCACGCGTCTTGTAGTTACCGTCGTCCTTGAAGGAGATGGTGCGCTCCTGGACGTAACCCTCC
GGCATGGCGGACTTGAAGAAGTCGTGGCGCTTCATGTGGTCCGGGTAGCGGGAGAAGCACTGGACGC
CGTAGGTCAGGGTGGTGACCAGGGTCGGCCACGGGACCGGCAGCTTACCGGTGGTGCAGATGAACTT
CAGGGTCAGCTTGCCGTAGGTGGCGTCGCCCTCACCCTCGCCGGAGACGGAGAACTTGTGGCCGTTGA
CGTCACCGTCCAGCTCGACCAGGATCGGGACGACGCCGGTGAACAGCTCCTCGCCCTTCGACATTATTA
ACCTCCTTTCTGCACGTGACGCACTGTCGGTTATCTAATCGTAATATTCCCATTTGTGGGCTTCGTGTGTC
GATCCTGGCAACGATCCGTTAGCGTCTACGCATCCGGTTGCGA