UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

DETECTION AND LOCALIZATION OF LYSA AND GP1 PROTEINS OF THE MYCOBACTERIOPHAGE Ms6

Alessa Ede Valverde da Silva

MESTRADO EM MICROBIOLOGIA APLICADA

2011

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

DETECTION AND LOCALIZATION OF LYSA AND GP1 PROTEINS OF THE MYCOBACTERIOPHAGE Ms6

Dissertação orientada por Prof.ª Dr.ª Madalena Pimentel (FFUL)

e Prof. Dr. Mário Santos (FCUL)

Alessa Ede Valverde da Silva

MESTRADO EM MICROBIOLOGIA APLICADA

2011

DETECTION AND LOCALIZATION OF LYSA AND GP1 PROTEINS OF THE MYCOBACTERIOPHAGE Ms6

Alessa Ede Valverde da Silva

MASTER THESIS

2011

This thesis was fully performed at Microbial Genetics Laboratory, Center for Molecular Pathogenesis, Unit of Retrovirus and Associated Infections (FFUL) under the direct supervision of Prof.ª Dr.ª Madalena Pimentel.

Prof. Dr. Mário Santos was the internal designated supervisor in the scope of the *Master in Applied Microbiology* of the Faculty of Sciences of the University of Lisbon.

AGRADECIMENTOS/ACKNOWLEDGMENTS

Finalizada uma etapa particularmente importante da minha vida, não poderia deixar de expressar o mais profundo agradecimento a todos aqueles que me apoiaram nesta longa caminhada e contribuíram para a realização deste trabalho. As minhas primeiras palavras são dirigidas à minha orientadora Madalena Pimentel por ter me aceitado no seu grupo de trabalho, pela sua disponibilidade e orientação prestada, pelo seu apoio incondicional e compreensão que sempre manifestou.

Agradeço ao meu orientador interno Mário Santos pela disponibilidade e ajuda prestada.

Ao Centro de Patogénese Molecular pela disponibilização de todo o material e afins necessários para a realização deste trabalho.

À Andreia e Filipa pela amizade e ajuda disponibilizada ao longo deste trabalho.

Por último mas não em último, quero agradecer especialmente à minha mãe pelo apoio incondicional e dedicação para que eu conseguisse chegar a esta etapa da minha vida.

Ao meu namorado por todo o amor, apoio e paciência que teve nesta fase final.

ABBREVIATIONS

aa	amino acid
AG	arabinogalactan
Amp	ampicillin
β-gal	beta-galactosidase
BCIP	5-bromo-4-chloro-3-indoxyl phosphate
bp	base pair
BRED	bacteriophage recombineering of electroporated DNA
BS	burst size
CM	cytoplasmic membrane
DNA	deoxyribonucleic acid
ds	double-stranded
EDTA	Ethylenediamine tetraacetic acid
FL	free lipids
GFP	green fluorescent protein
His	Histidine
HRP	horse radish peroxidase
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
kDa	kilodalton
LB	Luria-Bertani
MA	mycolic acids
Met	methionine
m.o.i.	multiplicity of infection
mRNA	messenger ribonucleic acid
nt	nucleotide
OD	optical density
OM	outer membrane
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PG	peptidoglycan
PhoA	alkaline phosphatase
p.m.f.	proton-motive force
PNPP	p-nitro-phenyl-phosphate
RNA	ribonucleic acid
SAR	signal-arrest-release
SDS	sodium dodecyl sulphate
SP	signal peptide
ss	single-stranded
TCA	trichloroacetic acid
TMD	transmembrane domain
Tris	tris(hydroxymethyl)aminomethane
wt	wild type

RESUMO

Bacteriófagos, ou fagos, são vírus que infectam bactérias. Estes organismos são dez vezes mais numerosos do que as bactérias no meio ambiente, sendo estes as entidades mais abundantes na Terra. Os fagos desempenham um grande papel no equilíbrio ecológico da vida microbiana. O estudo dos bacteriófagos, teve início em 1915 por Frederick Twort quando este isolou das fezes de um paciente com disenteria um "agente infeccioso e filtrável". Mais tarde, em 1918, D'Herelle descobriu uma entidade "antagonista" de bactérias, que provocava a sua lise em culturas líquidas e levavam à formação de placas de lise na superfície de meios de cultura com agar semeados com bactérias. D'Herelle criou a designação de bacteriófago para denominar os "vírus que comem bactérias". Apesar de até a data já terem sito descritos inúmeros fagos, apenas alguns têm sido estudados com algum detalhe. Relativamente à natureza dos ácidos nucleicos, os fagos podem apresentar um genoma quer de ADN de cadeia dupla, ADN de cadeia simples, ARN de cadeia dupla ou ARN de cadeia simples e normalmente o genoma é encapsulado em um invólucro proteico denominado cápside. A maioria dos fagos descritos até à data, apresentam cauda e um genoma de ADN de cadeia dupla (aproximadamente 96%). Existe uma variada gama de diferentes tipos de morfologia nos fagos e a sua taxonomia é baseada nos vários tipos de morfologia, presença ou ausência de envelope, pelo seu tamanho e pelo seu tipo de ácido nucleico.

A classificação actual é derivada de um esquema proposto por Bradley em 1967. Presentemente a classificação inclui uma ordem, 14 famílias oficialmente aceites e 37 géneros.

Os bacteriófagos, como todos os parasitas obrigatórios, requerem células hospedeiras para se manter e se reproduzir, aproveitando-se dos seus mecanismos de biossíntese para sobreviver. Com base no tipo de ciclo de vida, os bacteriófagos podem ser de dois tipos: fagos líticos, que apresentam um ciclo de vida lítico; ou fagos temperados, que apresentam ciclo de vida lisogénico e lítico. No ciclo lítico, o fago redirecciona o metabolismo do hospedeiro para a produção de novos vírus, que são libertados durante a lise celular provocada pelos vírus. No ciclo lisogénico, o genoma do fago é inserido no genoma do hospedeiro e sendo então replicado quando o genoma do hospedeiro se replica até que o ciclo lítico seja induzido, por exemplo por agentes mutagénicos. Para iniciar a infecção, os fagos necessitam adsorver à superfície bacteriana, através do reconhecimento de receptores presentes na parede celular e posteriormente ocorre a entrada do genoma viral na célula hospedeira. Após a injecção do genoma viral, este é transcrito pelas polimerases do hospedeiro eventualmente há produção de novas partículas virais pela maquinaria do hospedeiro. Quando o ciclo lítico sucede, na maioria dos casos é necessário ocorrer lise da célula para que os novos viriões possam encontrar novos hospedeiros. Os fagos de ADN de dupla cadeia desenvolveram um sistema de lise em que utilizam uma estratégia denominada "holina-endolisina" para realizar uma lise celular eficiente no hospedeiro.

O fago Ms6 é um micobacteriófago temperado, com genoma constituído por ADN de cadeia dupla, que infecta a bactéria *Mycobacterium smegmatis*. Os micobacteriófagos representam excelentes sistemas modelo para estudar micobactérias hospedeiras. Estes bacteriófagos têm vindo a desempenhar um papel fundamental no desenvolvimento de sistemas genéticos para o estudo da patogénese

molecular das micobactérias patogénicas para o homem, sendo as espécies mais reconhecidas *Mycobacterium tuberculosis* e *Mycobacterium leprae*, os agentes causadores da tuberculose e lepra, respectivamente. *Mycobacterium smegmatis* é considerada uma micobactéria não-patogénica representando assim um bom modelo de estudo.

O micobacteriófago Ms6, sendo de cadeia dupla, utiliza a estratégia “holina-endolisina” para induzir uma lise efectiva do hospedeiro. As endolisinas são enzimas muralíticas com actividade hidrolítica das ligações específicas que se estabelecem na camada de peptidoglicano e as holinas são proteínas membranares de pequenas dimensões, que conduzem a uma alteração do potencial de membrana como consequência da abertura de poros na membrana, actuando assim como controladora da activação da endolisina ou permitindo o seu acesso à mureína.

Na maioria dos casos descritos até à data, o acesso da endolisina ao seu alvo é feito pela passagem através dos poros formados pela holina. Actualmente têm sido descritos fagos que produzem endolisinas cujo transporte para o meio extracitoplasmático é independente da holina, utilizando os sistemas de transporte bacterianos. Estas endolisinas apresentam na sua estrutura uma sequência sinal, que permite a translocação da proteína, através da membrana citoplasmática, utilizando o sistema *sec* (translocase) do hospedeiro, até ao espaço periplásmico. Foram já descritas endolisinas que apresentam na região N-terminal um péptido sinal, como é o caso da Lys44 produzida pelo fago fOg44 de *Oenococcus oeni*. Recentemente foram descobertas outras endolisinas, de fagos que infectam bactérias Gram-negativas, que apresentam na sua extremidade N-terminal uma região de carácter hidrofóbico, designada por *Signal-Arrest-Release* (SAR), que permite igualmente, o transporte da endolisina através da membrana citoplasmática com o auxílio do sistema *sec*,, com a particularidade de não ser clivada. Apesar de nestes casos a holina não participar no transporte da endolisina, pensa-se que esta possa desempenhar um papel na activação de lisinas que se encontram no espaço periplásmico e na determinação do tempo de lise.

A cassete lítica do fago Ms6 é constituída por 5 genes. Para além dos genes da endolisina (*lysA* ou *gp2*) e holina (*hol* ou *gp4*), a cassete lítica do Ms6 inclui mais 3 genes, *gp1*, *gp3* e *gp5*.

Catalão e colegas (2010) mostraram que a endolisina LysA é exportada com a ajuda do produto do gene *gp1*, de uma forma independente da holina. É demonstrado que o gene *gp1* codifica uma proteína “chaperone-like” que se liga à endolisina, auxiliando a sua exportação para o ambiente extra-citoplasmático, o que é necessário para uma eficiente lise do fago Ms6. Actualmente não se sabe como o endolisisna permanece inactiva até a holina determinar o tempo de lise. Assim, o objectivo principal deste trabalho foi investigar o papel do produto do gene *gp1* no transporte de LysA e outras proteínas sem péptido sinal, para o periplasma. Para alcançar este objectivo, construiu-se um fago Ms6 mutante onde o gene *gp1* contém na sua extremidade 3 ' uma sequência que codifica para uma cauda C-Myc permitindo a produção de uma proteína recombinante Gp1-C-Myc. Esta proteína foi detectada com o anticorpo anti C-Myc na fracção solúvel e membranar quando células de *M. smegmatis* foram infectadas com o fago mutante Ms6Gp1-CMyc.

O mesmo resultado foi obtido em *E.coli* confirmando que Gp1 se distribui entre os compartimentos solúveis e membranares. Para apoiar este resultados foram utilizados duas proteínas repórter, a fosfatase alcalina bacteriana (PhoA) como repórter para a localização periplasmática e a proteína verde fluorescente (GFP) como repórter para a localização citoplasmática. A capacidade de Gp1 translocar proteínas foi também apoiada pelo resultado obtido com fusões com a PhoA e GFP. Os resultados da fusão da Gp1-PhoA revelaram actividade da fosfatase alcalina, enquanto a fusão Gp1-GFP resultou em perda de emissão de fluorescência.

Na experiência realizada com a proteinase K obteve-se um perfil alterado de proteínas quando LysA foi sujeita a um tratamento com proteinase K, sugerindo que a LysA pode estar localizada no periplasma na presença de Gp1.

Os resultados apresentados reforçam o papel da Gp1 em auxiliar outras proteínas, sem um péptido sinal, para atravessar a membrana para o periplasma.

Os resultados obtidos neste trabalho experimental abrem novas perspectivas no sentido de oportunidades para novos estudos em vias de secreção de micobactérias. Também uma melhor caracterização do genoma dos bacteriófagos poderá levar a contribuição para a identificação de novos alvos terapêuticos na parede das micobactérias.

Palavras-chave: Micobacteriófagos; *lise,* endolisina; holina; micobactéria;

ABSTRACT

Bacteriophages are viruses that infect bacteria. The majority described to date (≈96%) are endowed with a tail and contains a double-stranded DNA (dsDNA) genome. The phage Ms6 is a temperate phage that infects *Mycobacterium smegmatis*. Like dsDNA phages, Ms6 uses the holin-endolysin mechanism to accomplish lysis of the host. In addition to endolysin (*lysA*) and holin (*hol*) genes, Ms6 genome encodes three accessory lysis proteins, Gp1, LysB and Gp5. Endolysins are proteins that degrade the rigid murein layer, and holins are small membrane proteins that control the activation of the endolysin or its access to the murein. Recently it has been shown that the endolysin *LysA* is exported with the help of the *gp1* gene product, which is involved in its delivery to the peptidoglycan, in a holin-independent manner. Thus, the purpose of this project is to investigate the role of *gp1* gene product in the transportation of *LysA* and other proteins without a peptide signal, to the periplasm. To achieve this we constructed a mutant Ms6 phage where the *gp1* gene carries, at its 3' end, a sequence coding for a C-Myc tag allowing the production of a recombinant protein that was detected with an antibody anti C-Myc in the soluble and membrane fraction of *M. smegmatis* infected cells. The same result was obtained in *E.coli* confirming that Gp1 distributes between the soluble and membrane compartments. The ability of Gp1 to translocate proteins is suggested by the obtained result with alkaline phosphatise (PhoA) and green fluorescence protein (GFP) fusions. A Gp1-PhoA fusion results in alkaline phosphatase activity while a Gp1-GFP fusion results in loss of fluorescence emission. In addition, an altered profile of the proteins when LysA is subject to a proteinase K treatment suggests that LysA may be periplasm localized in the presence of Gp1. The results presented here strengthen the role of Gp1 in helping other proteins, without a peptide signal, to cross the membrane to the periplasm.

Keywords: Mycobacteriophages; lysis, endolysin; holin; recombineering; mycobacteria;

TABLE OF CONTENTS

Agradecimentos/Acknowledgments i

Abbreviations ii

Resumo........ iii

Abstract........ vi

Introduction 1

Classification........ 1

The phage life cycle 2

Bacteriophage-Induced host cell lysis........ 4

The sec-Mediated Lysis 7

Mycobacteriophages........ 9

Mycobacteriophage Ms6 and its Lysis Operon........ 9

Objectives........ 12

Materials and Methods 13

Bacteriophages, bacterial strains, plasmids and growth conditions 13

DNA extraction from Ms6 wt........ 14

One-step growth experiments 14

Gp1 expression in *M. smegmatis*-infected cells 15

Cellular fractionation of *M. smegmatis*........ 15

Protease accessibility experiments........ 16

Plasmid construction........ 16

Detection and quantification of Alkaline Phosphate activity 17

Expression and determination of solubility from Gp1 and Gp1-GFP in E. coli cells...... 17

PCR reactions.. 18

Results 19

Construction of the Ms6Gp1-CMyc mutant 19

Evaluation of the Ms6Gp1-CMyc growth parameters 21

Gp1 expression in *M. smegmatis*-infected cells 22

Subcellular localization of Gp1 *in M. smegmatis*-infected cells.................................... 23

Expression and localization of Ms6 Gp1 in *E. coli*.. 24

Analysis of a Ms6 Gp1 fusion to Alkaline Phosphatase... 25

Analysis of a Ms6 Gp1 fusion to Green-fluorescent protein... 28

Protease accessibility experiments of LysA in *E. coli* 31

Discussion/ Conclusion 33

References.. 35

INTRODUCTION

Bacteriophages (phages) are viruses that have the ability to infect bacteria. They are among the most common biological entities in the biosphere, which play a major role in the ecological balance of microbial life (Pedulla 2003; Deschavanne, 2010). They also play a profound role in the evolution of their host (bacteria) and represent one of the three major mobile genetic elements that contribute significantly to horizontal gene transfer in bacterial genomes by transduction (Boyer *et al.*, 2008). Whole-genome sequencing studies in bacteria have shown that phage elements contribute significantly to sequence diversity and can potentially influence bacterial pathogenicity.

Phages are ten times more abundant than bacteria, with an estimated 10^{32} bacteriophages on earth (Hanlon, 2007). On a global scale, it is estimated that ~ 10^{25} phages initiate an infection every second (Pedulla *et al*., 2003; Brüssow *et al*., 2004; Hendrix, 2005; Hatfull, 2010) and in each of those infections the phage encounters DNA (bacterial or prophage DNA), and can potentially recombine to generate new genomic arrangements (Canchaya *et al.,* 2003; Hatfull *et al*., 2006).

Bacteriophages were first discovered by Frederick Twort in 1915, when he described the glassy transformation of "*Micrococcus*" colonies by a transmissible agent, and by Felix d'Herelle in 1917 (Weinbauer, 2003; Kutter *et al*., 2004) when he observed and described the lysis of *Shigella* (Ackermann, 2003). D'Herelle created the term *bacteriophages* for these infectious agents lysing the bacteria, which literally means "eaters of bacteria". He also developed several techniques that are still in use, postulated the intracellular multiplication of viruses, and was the pioneer to the phage therapy of infectious diseases (Ackerman, 1997).

Classification

There exists a range of different morphological types of phages and taxonomy is based on their capsid shape, the presence or absence of an envelope, their size and also by their nucleic acid type.

The actual phage classification is derived from a scheme proposed by Bradley in 1967. Currently the classification includes one order, 14 families officially accepted and at least five other are awaiting for the classification, and 37 genera (Ackermann, 2009).

Phage genome consists either of double-stranded DNA (dsDNA), single-stranded DNA (ssDNA) or very rarely of single-stranded RNA (ssRNA) and normally the phage's genome is encapsulated in a protein coat (capsid). Bacteriophages can be tailed or tailless. Phages that have a tail in its structure, belong to the order *Caudovirales* and represents around 96% of all the phages described so far. Each phage of the *Caudovirales* order possesses a double-stranded DNA structure and is grouped in one of three families: *Myoviridae* (with long contracted tails), *Siphoviridae* (with long non contracted tails) and *Podoviridae* (with short tails). The remaining phages can be either filamentous, polyhedral or pleiomorphic and are grouped in the families *Microviridae* (ssDNA), *Leviridae* (ssRNA), *Cystoviridae*

(dsRNA with envelope) or *Inoviridae* (ssDNA with long filaments) (Ackermann, 2003; Ackermann, 2009; Hanlon, 2007).

The phage life cycle

Viruses are obligate pathogens that require host cells in order to replicate (Short *et al*., 1999). Each phage multiplies and produces new viral particles inside bacteria using the host biosynthetic machinery to survive (Young, 1992; Young *et al*., 2002).

Upon infection of the bacterial host, bacteriophages can follow several life cycles: lytic, lysogenic, pseudolysogenic and chronic infections (Weinbauer, 2004). In the lytic cycle, the phage redirects the host metabolism towards the production of new phages, which are released during the lysis of the host cell. In the lysogenic cycle, the phage's genome remains in the host in a dormant stage (as prophage or as plasmids) and replicates along with the host, until the lytic cycles is induced (Weinbauer, 2004; Hanlon, 2007; Ackerman 2009). The life cycle of a temperate phage is exemplified in Figure 1.

In chronic infections, the cell is infected and phage progeny is continuously released from the host cell by budding or extrusion without lysing the cell. In persistent infections (pseudolysogeny) phages multiply in a fraction of the population (Weinbauer, 2004).

Infection of cells by viruses requires direct contact between a virus particle and its host (Short *et al.,* 1999). To initiate an infection, a phage virion has to first adsorb to the surface of a susceptible host cell (Shao *et al*., 2008). Bacterial virus do not randomly attach to the surface of a host cell, they fasten to specific receptor sites that may be any one of a wide variety of cell surface components, including proteins, oligosaccharides, teichoic acids, peptidoglycan or lipopolysaccharides. (Kutter *et al.,* 2004; Guttman *et al.,* 2004; Hanlon, 2007;). In some cases, the attachment site might be present on the cell capsule, flagella or even conjugative pili. At first, the attachment is reversible but then it becomes irreversible and is followed by transfer of phage genetic material into the host cell. The entrance of the phage genome into the host cell can occur by a variety of mechanisms, depending on the physiology of the virus (Weinbauer, 2004; Hanlon, 2007). In some phages, the bases present on the phage DNA are chemically modified to confer protection against attacks by cellular restriction and nuclease enzymes. After DNA injection, the viral genome is transcribed by the host cell RNA polymerase, which starts to synthesize early mRNAs to take over the metabolic machinery of the bacterium, redirecting its metabolic processes to the manufacture of new virus particles (Hanlon, 2007). Early mRNA synthesis is followed by the phage DNA replication and later mRNA synthesis. Late mRNAs direct the synthesis of three kinds of proteins: phage structural proteins, proteins that help with the phage assembly without becoming part of the virion structure, and proteins involved in the cell lysis and phage release (Guttman *et al*., 2004). Afterwards, the construction and assembly of new phage particles within the host cell takes place, and in the majority of the phages, the release of the new virions implies host cell lysis (Hanlon, 2007). In some phages, the cell lysis is mediated by the production of late proteins, which attack the bacterial peptidoglycan (Young *et al.,* 2006; Hanlon, 2007).

In the lysogenic cycle, temperate bacteriophages do not automatically enter a lytic cycle, but alternately, integrate their DNA into the host cell DNA after infection (Figure 1). The bacterial cells are then designated lysogenic cells. When the bacterial DNA replicates, the phage DNA replicates simultaneously and so each daughter cell will withhold the viral DNA (prophage). Each prophage produces a repressor protein that blocks the transcription of its own genes and also those of similarly related bacteriophages (Weinbauer, 2004). The existence of a prophage can therefore grant, upon a bacterial cell, some sort of immunity to infection by the same or close related bacterial virus. Lysogenic cells may undergo several rounds of division, but sporadically, one cell can lyse spontaneously and release the phage's progeny. Nerveless, a population of lysogenic cells may be induced to enter the lytic cycle through environmental stress, such as treatment with mutagenic agents or exposure to ultraviolet light (Hanlon, 2007).

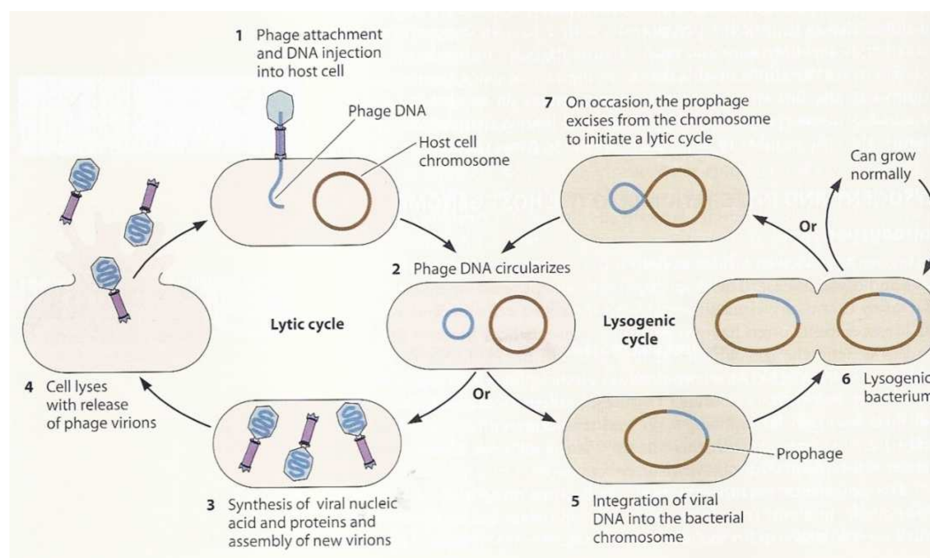


Figure 1. Two lifestyles of a temperate bacteriophage: lysogenic and lytic cycle (Shaechter *et al.*, 2006).

When a prophage escapes regulation by the repressor, its DNA is cut free allowing it to follow a lytic cycle. The excision of prophage DNA is frequently imprecise and bacterial genes adjacent to the prophage DNA may be incorporated into the infectious phage DNA and then transferred to subsequent host cells (Hanlon, 2007). This mechanism is known as restricted transduction, and is responsible for the horizontal transfer of genes, from one bacterial cell to another (Canchaya *et al.*, 2003; Hanlon, 2007). The acquisition of prophages would be an irrelevant process for the evolution of pathogenic bacteria if phages wouldn't transfer useful genes to the lysogen cell, which are known to increase the survival fitness of lysogens (Brüssow *et al.,* 2004).

Bacteriophage-Induced host cell lysis

Most bacteriophages must lyse their host cell to release the newly assembled progeny virions to the extracellular environment (Young *et al.*, 2006) Lysis is a programmed event of major importance regarding the phage survival and ecological fitness (Wang, 2006). A sharply time-defined and efficient release of phage progeny is crucial to maximize both burst-size (number of phage progeny released) and the opportunity to infect new hosts (Wang, 2006; São-José *et al.*, 2007; Shao *et al.*, 2008).

If lysis occurs too early, prior to the assembly of a significant number of progeny, the infectious cycle would not be profitable. On the other hand, if lysis happens too late it might preclude the phage's progeny from taking advantage of other potential host cells in the environment of the infected cell. It is likely that these opposing evolutionary pressures are balanced, giving a characteristic optimal lysis time for each bacteriophage (Bernhardt *et al.*, 2002).

The major barrier to lysis is the continuous meshwork of peptidoglycan, which is a stable structure that allows the bacterial envelope to resist internal osmotic pressure (Young *et al.*, 2000; Wang, 2006). Filamentous phages have a unique morphology and morphogenesis and can extrude from the cytoplasm without lethal consequences for the host. The rest of the phages must either degrade or compromise the peptidoglycan to provoke lysis (Young *et al.*, 2000; São-José *et al.*, 2007).

There are at least two distinct strategies for bacteriophages lysis: one employed by phages with small single-stranded nucleic acids and the other by phages with double-stranded genome. Single-stranded DNA and ssRNA phages accomplish lysis with a single lysis protein that does not encode a muralytic enzyme activity. This protein causes lysis through acting as a specific inhibitor of an enzyme of the multi-step pathway of murein biosynthesis, thus inhibiting the cell wall synthesis (Young *et al.*, 2000; Bernhardt *et al.*, 2002).

In double-stranded DNA (dsDNA) phages, two complementary functions have independently evolved to accomplish both rapid progeny release and optimal time lysis (São-José *et al.*, 2007). Double-stranded DNA phages evolved a lytic system that uses a "holin-endolysin" strategy to accomplish an efficient host cell lysis. These phages produce, during the lytic cycle's late phase of gene expression, a soluble muralytic enzyme called endolysin or lysin. Endolysin is a rather generic term used to describe a range of bacteriophage-encoded peptidoglycan hydrolases. These enzymes are also known as phage lysozymes, lysins, or muralytic enzymes. They are characterized by their ability to target bonds in the peptidoglycan (PG) layer of the host cell wall. Therefore, endolysins degrade the rigid murein layer and release the newly assembled virions by lysis (Loessner, 2005). Generally, endolysins have muralytic activity against one of the three different types of covalent bonds (glycosidic, amide, and peptide) of the peptidoglycan polymer of the cell wall. However, a small number of endolysins with more than one type of muralytic activity have been described, such as, endolysins encoded by *Streptococcus agalactiae* bacteriophage B30 (muramidase and peptidase) (Pritchard et al., 20004), *Staphylococcus aureus* phage

Φ11 (endopeptidase and amidase) (Navarre et al., 1999) and *S.agalactiae* phage NCTC 11261 (endopeptidase and muramidase) (Cheng et al., 2005).

Depending on the enzymatic specificity, endolysins can be divided into five major functional types: (i) N-acetylmuramidases (lysozymes), (ii) endo-b-N-acetylglucosaminidases, and (iii) lytic transglycosylases, which all cleave the sugar moiety of peptidoglycan; (iv) endopeptidases, which cleave the peptide moiety; and (v) N-acetylmuramoyl- L-alanine amidases, which cut the amide bond between both moieties (Figure 2) (Borysowski *et al.,* 2006)

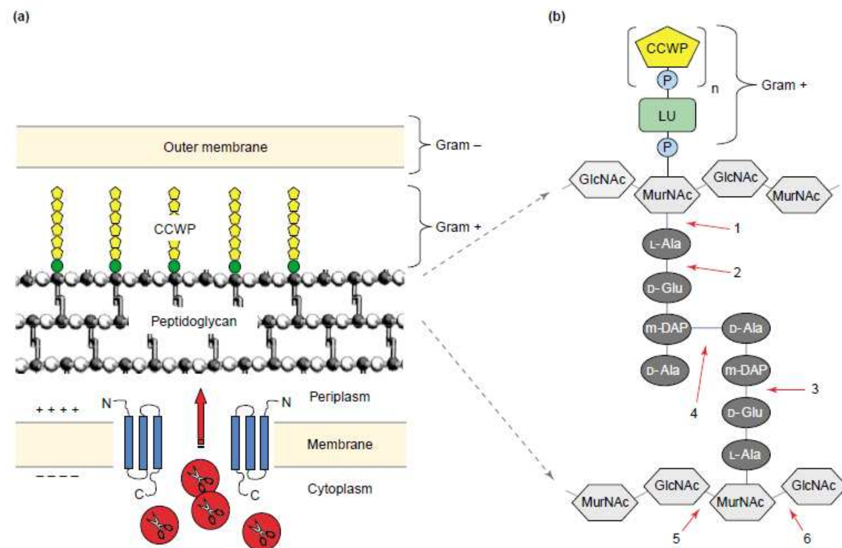


Figure 2. ***Bacterial cell wall structure and endolysin targets***. **(a)** Schematic representation of the bacterial cell wall, and of how phage endolysins, like λR, gain access to their substrate. Holin proteins (blue) insert themselves into the cytoplasmic membrane and can oligomerize, thereby forming membrane lesions. The endolysins (red) pass through these pores to access the peptidoglycan. **(b)** The bonds potentially attacked by endolysins of different enzymatic specificities are indicated by numbers: 1) N-acetylmuramoyl-L-alanine amidase; 2) L-alanoyl-D-glutamate endopeptidase; 3) D-glutamyl-m-DAP endopeptidase (this activity has not yet been identified in a phage endolysin); 4) interpeptide bridge-specific endopeptidases; 5) N-acetyl-b-D-glucosaminidase; and 6) N-acetyl-b-D-muramidase (also known as muramoylhydrolase and 'lysozyme') and lytic transglycosylase (Figure from Loessner, 2005).
Abbreviations: CCWP, carbohydrate cell wall polymer; GlcNAc, N-acetyl glucosamine; LU, linkage unit; m-DAP, meso-diaminopimelic acid; MurNAc, N-acetyl muramic acid; P, phosphate group

The endolysins structure, from phages infecting Gram positive host, is composed by two domains, positioned at the N and C-terminus connected by a short linker (L) (Fig. 3). Generally the N-terminal domain contains the catalytic activity of the enzyme, and the C-terminal domain binds to a specific substrate (usually carbohydrates) found in the cell wall of the host bacterium (Fischetti, 2005; Lossner, 2005). Cleavage requires interaction of the binding domain with its cell wall substrate. This offers some degree of specificity to the lysins because these substrates appear to be found only in

enzyme-sensitive bacteria. Sequence comparisons of enzymes with the same activity indicate that the catalytic region is highly conserved whereas the C-terminal region is variable (Fischetti, 2005).

Figure 3. Basic structure of phage lytic enzymes (Fischetti, 2005).

Most endolysins described so far lack a secretory signal sequence and are thus unable to access the murein layer. During a phage infection these endolysins accumulate in the cytoplasm, until another lysis protein, the holin, forms pores in the cytoplasmic membrane that allow the endolysin to reach its target (Figure 2) (Young *et al.*, 2000). Holins are small hydrophobic proteins that form a hole in the cell membrane and serves to release or activate the endolysin at a programmed time, (Young, 1992; São-José *et al.*, 2003; Loessener, 2005) regulating the timing of the lysis. Holins are thus, subject to intense evolutionary pressure to achieve lysis at an optimal time (Young, 2005; Wang *et al.*, 2006). Unlike endolysins, holins are more clustered and frequently unique with respect to their primary structure. Holins are small proteins, with a hydrophilic and highly charged C terminal sequence (São-José *et al.*, 2003). They are currently grouped into three classes, based on their membrane topology. Class I, with three transmembrane domains (TMDs) (N side out, C side in), and class II, which have two TMDs (N side in, C side in). Both classes I and II have multiple, unrelated gene families, but only one gene family, the T4 holin gp*t* and its relatives in T-even phages, defines a class III, which have one TMD (N side in and Cside out (Young, 2005).

During late gene expression, holins progressively accumulate in the membrane and when the proton motive force (PMF) is sufficiently depleted, the holins undergo some massive rearrangement, resulting in oligomers and form holes in the cytoplasmic membrane through which some endolysins can then pass through (Loessener, 2005; Young, 2005; São-José *et al.,* 2007). The structure alterations introduced in the holin may either prevent lysis or lead to premature or delayed lysis (São-José *et al.*, 2007). Holins can be prematurely triggered by membrane depolarization with energy poisons such as cyanide and dinitrophenol (Wang *et al.*, 2000; Young, 2005).

In addition to holins, phages can also synthesize other proteins, such as antiholin, an inhibitor of the holin activity. If present, the inhibitor is usually transcribed from the same reading frame as the holin (dual start motif), but contains a functionally defective transmembrane domain (Bernhardt, 2002; Loessener, 2005). In the case of the λ phage, the *S107* (antiholin) differs from the holin gene product, *S105*, by two residues, Methionine and Lysine (Young, 2005). Antiholins can be either soluble or membrane-bound (São-José *et al.*, 2007). In others cases, the antiholin, can be coded by an independent gene. In 20 out of 46 cases of phages infecting Gram-positive hosts in which a typical holin-lysin cassette was reported, a second holin-like gene was located immediately upstream of the hol-lys pair (São-José *et al.*, 2003). In most cases, the endolysin and holin genes are adjacent and frequently clustered with other

genes providing supplementary lysis functions in a ‘‘lysis cassette’’, encoding up to five proteins (Young, 2002).

The best-characterized example of holin-endolysin systems is the λ phage, which is composed by 5 genes (São José *et al.*, 2007). The first genes of the late operon are the phage’s lysis genes: S, R, Rz and Rz1. The *S* gene encodes the holin (S_{105}) and the antiholin (S_{107}) and *R* encodes the endolysin with a transglycosylase activity (Wang *et al.*, 2003; Young, 2005).The DNA sequence in the lysis cassette contains two additional genes *Rz* and *Rz1,* that encode an inner membrane and a lipoprotein, respectively. Together, they interact and form a complex allowing the fusion of inner and outer membrane (Young *et al.,* 2006).

The lysis cassette of the P2 phage contains the genes *Y*, *K*, and *lysABC*, with *Y* encoding a class I holin, *K* encoding the endolysin, and the *lysBC* genes resemble the *Rz* and *Rz1* genes of λ phage and lysA seems to be an antiholin (Young, 2005).

Genes encoding similar gene products were identified in several other phages infecting Gram-negative hosts (Summer *et a*l., 2007). Therefore, the mechanism of bacteriophage lysis suggests that the complexity of the phage’s lytic cassettes depends on their hosts. Hosts with a more simpler envelope require the phage to possess a simple lytic cassette, while hosts with a complex envelope require the phage to have a more complex lytic cassette (Gil *et al.*, 2010).

The sec-Mediated Lysis

One of the characteristics of lysis regulation is that the phage endolysins would not reach the cell wall before the formation of holin lesions (see Figure 4).

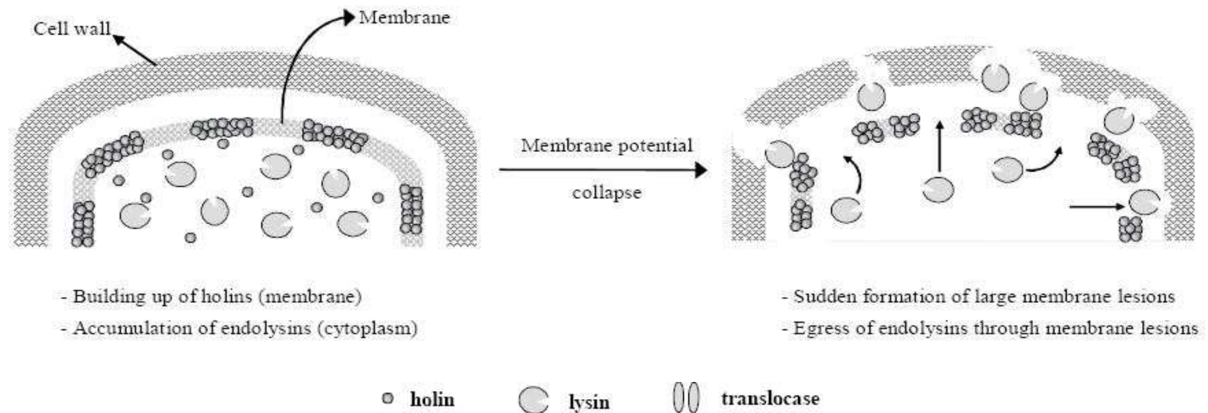


Figure 4. Model representation of host-lysis strategies of phages producing non-secreted endolysins (São-José *et al.*, 2003).

This was long thought to be universal, but recently it was shown that some phages produce endolysins whose transport across the cytoplasmic membrane is not dependent on holin function. These endolysins have a signal sequence in its structure that allows for translocation of the protein through the host cytoplasmic membrane to the periplasmic space involving the host *sec* system (Xu *et al.,* 2004). This

paradigm began with studies by Santos and colleagues on the phage fOg44, which grows on the Gram-positive bacterium *Oenococcus oeni*. This phage encodes an endolysin (Lys44), which possesses a cleavable N-terminal signal sequence that also functions in *Eschericha coli*. Lys44 is continuously exported during assembly to the extracytoplasmic environment by *sec* machinery and its signal sequence is proteolytically removed by the leader peptidase, generating an active enzyme (Xu *et al*., 2004; Young, 2005; São-José *et al*., 2007; Catalão *et al.,* 2010). In spite of the presence of a secretory enzyme, fOg44 has a holin gene. Moreover, fOg44 infections, showed that the mature endolysin was already detectable long before lysis was achieved, demonstrating that the *sec*-mediated export of the endolysin is not sufficient to provoke lysis.

These remarks give raise to a number of questions about the role and mode of action of holins and how lysis timing is regulated in phages where the endolysin is *sec*-exported (Xu *et al*., 2004; São-José *et al*., 2007). The authors proposed that if endolysin can use host endogenous pathways to reach their substrate, the holin function would seem expendable. Their model suggests that, in phages producing secreted endolysins, the activity of the target endolysins would be inhibited in the cell wall until dissipation of the membrane potential by the holins (Figure 5), therefore, holins would activate the exported endolysins by collapsing the membrane potential, instead of releasing them (São-José *et al*., 2007).

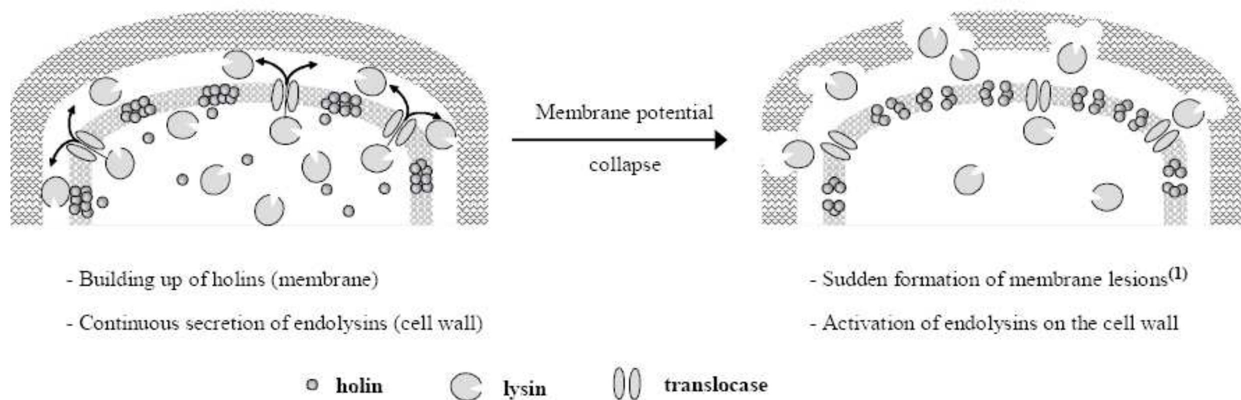


Figure 5. Model representation of host-lysis strategies of phages producing peptide-bearing endolysins (São-José *et al*., 2003).

More recently, it was demonstrated that a number of endolysins from phages that infect Gram-hosts, have intrinsic export signals, named signal-arrest-release (SAR) domains (Xu *et al.,* 2004; Xu *et al,*. 2005). A particular interesting case is that of the endolysin of enterobacteriophage P1 (*Lyz*) and the lambdoid coliphage 21 (R^{21}): they do not require the holin for export, but differ from the fOg44 secretory endolysin in that export is mediated by a N-terminal transmembrane domain (TMD), which is not proteolytically cleaved.. The N-terminal domain of P1Lyz phage is both necessary and sufficient, not only for the export of this endolysin through the membrane, but also for its release into the periplasm. This class of endolysins also requires the host *sec* translocon to cross the cytoplasmic membrane. The

unusual N-terminal domain, rich in residues that are weakly hydrophobic, functions as a SAR sequence, which acts as a normal signal-arrest domain to direct the endolysin to the periplasm in a membrane tethered form where it remains enzymatically inactive. This functional regulation is essential to avoid premature lysis of the infected host (Sun *et al.,* 2009). After that, SAR allows P1's Lyz to be released from the membrane, as a soluble active enzyme in the periplasm (Xu *et al.*, 2004; Xu *et al.*, 2005; Young, 2005).

A model for triggering of lysis with SAR endolysins was proposed: firstly the SAR endolysin is tethered in an inactive form to the energized membrane and the holin protein is accumulated without affecting the proton-motive force (pmf). At the programmed lysis time, the holin triggers, disrupting the membrane sufficiently to abolish the pmf, and possibly assists the release of the endolysin from the membrane, which results the activation of the endolysin (Xu *et al.*, 2004).

Interestingly the class represented by the lambdoid bacteriophage 21, utilizes endolysins having N-terminal SAR domains and pinholins, as opposed to the large-hole-forming holins (small hole size) (Park *et a*l., 2006; Park *et al.*, 2007). The pinholins are not large enough to allow endolysin to cross through them. The term ''pinholin'' has been proposed to differentiate the small-hole (pinhole) characteristic of the phage's 21 holin from the canonical holins that form large, non-specific holes (Pang *et al., 2009*).

Mycobacteriophages

Mycobacteriophages are virus that infect mycobacteria hosts. The interest in these phages derives, in part, from the medical significance of their hosts (Hatfull, 2000; Hatfull, 2006). Mycobacteria are acid-fast staining bacteria with characteristic waxy cell walls, containing a peptidoglycan-arabinogalactan polymer linked to a long chain of mycolic acids (MA). The cell wall provides an extraordinarily efficient permeability barrier which contributes to the high intrinsic resistance to many drugs (Hoffmann, 2008). Mycobacteria are divided into two groups based on their growth rate: slow-growers such *Mycobacterium tuberculosis* and fast-growers such *Mycobacterium smegmatis*. Many mycobacterial species are human and animal pathogens, the most recognized being *M. tuberculosis* and *M. leprae*, the causative agents of tuberculosis and leprosy, respectively (Hatfull *et al.*, 1994; Hatfull, 2006; Zuber, 2008). Even though *M.smegmatis is* considered non-pathogenic, it provides a popular model for studying the virulence mechanisms of the pathogenic mycobacteria (*Arora et al.,* 2008).

Mycobacteriophage Ms6 and its Lysis Operon

Mycobacteriophage Ms6 is a temperate phage that infects *Mycobacterium smegmatis*.
Electronic microscopy studies revealed that these phage particles are composed by an isometric polyhedral head with 80 nm in diameter, hexagonal form and an extensive non-contractile tail 210 nm long. The Ms6 genome is constituted by a molecule of dsDNA with a length over 50 kbp (unpublished

data) and GC content of 62%. The morphological characteristics of the mycobacteriophage allowed for its classification in the *Siphoviridae* group (Portugal *et al.,* 1989).

Freitas-Vieira and colleagues (1998) have characterized the genetic elements involved in the site-specific integration events between phage Ms6's DNA and the mycobacterial genome recombination mechanism required for chromosomal integration. This recombination is catalysed by a phage-encoded *recombinase* and involves a common core sequence present in both phage *(attP)* and bacterial *(attB)* genomes (Freitas-Vieira *et al.,* 1998).

In 2002, the genetic organization and some transcriptional control elements of the mycobacteriophage Ms6 lysis functions were described. Garcia and colleagues identified and isolated, from Ms6, a strong promoter region (P_{lys}) by using transcriptional fusions with *lacZ* reporter gene. Two tandem σ^{70}-like promoter sequences (*P1* and *P2*), which are recognized by the host's *RNA polymerase*, were found in this region by genetic analysis. Transcription of the lysis genes is dependent on the P_{lys} promoter located about 6kb away from the integration locus (Figure 6A), which is positioned in the middle of the Ms6 genome. Furthermore, an intrinsic transcription termination signal was detect in the leader sequence upstream of the first open reading frame (ORF). These data suggest that an anti-termination mechanism may be involved in the regulation of Ms6 lysis genes transcription (Garcia *et al*., 2002).

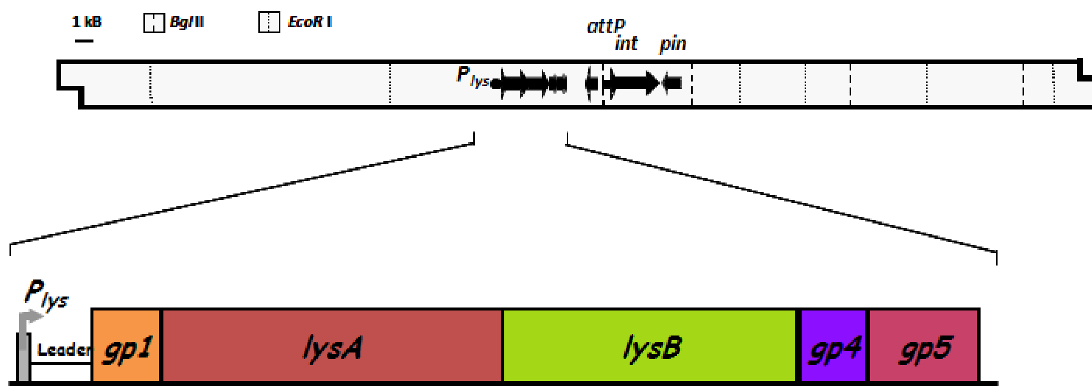


Figure 6. (A) Schematic representation of *Plys* on the 57-kb Ms6 DNA. **(B)** Genetic organization of the Ms6 lysis locus. Figure Adapted from Garcia *et al*., 2002.

The lytic *cassette* of Ms6 consists of five genes. Apart from the endolysin (*lysA*) and holin (*hol*) genes, the Ms6 lytic cassette includes three extra genes (*gp1, gp5 and gp3*) (Fig.6B). The ***gp1*** gene, 231 bp long, is separated from the +1 position by the leader sequence and has the potential to encode a 77-amino acid (aa) protein that was recently identified has a chaperon-like protein that specifically interacts with the N-terminal region of LysA and is involved in its delivery to the peptidoglycan in a holin-independent manner (Catalão, *et al*., 2010).

gp2 or lysA, with 1152 bp, starts at a GTG codon that overlaps the gp1 TGA stop codon, which is in a different reading frame. It encodes the Ms6 endolysin with 384 aa (Garcia *et al.,* 2002). Analysis of the deduced amino acid sequence have identified an amidase domain and and its hydrolase activity was already demonstrated; the protein was shown to cleave the bond between L-Ala and D-muramic acid (Piechota *et al.*, unpublished). Recently it was shown that during an Ms6 infection, *lysA* synthesizes two proteins: Lysin384 translated from the start codon at position 1 and Lysin241 translated from a second start codon at position 430 of the nucleotide sequence (Catalão *et al.,* 2011a). It was demonstrated that both proteins are necessary for an efficient lysis, although it's not known the reasons why Ms6 produces two endolysins.

Downstream of *lysA* is ***gp3*** **(*lysB*)**, with 996 bp and starts at an ATG codon that overlaps, in a different reading frame, the TGA stop codon of *gp2* and encodes a 332 aa protein (Garcia *et al.,* 2002). The LysB enzyme was shown to have lipolytic activity that targets the mycobacteria's outer membrane (Gil *et al.*, 2008; Gil *et al.*, 2010).

Mycobacteria, apart of having a mycobacterial outer membrane composed of mycolic acids and free lipids, have a second barrier, an arabinogalactan-peptidoglycan layer (Hoffmann *et al.,* 2008). Recently, Gil *et al.,* 2010 have shown that Ms6's LysB cleaves the ester bond between mycolic acids and arabinogalactan. In addition, LysB also hydrolyzes other lipids containing mycolic acids present in the envelope of mycobacteria, such as the trealose dymicolate, a lipid involved in pathogenesis of slow growing mycobacteria (Gil *et al.,* 2010). Payne *et al.,* (2009) also reported that mycobacteriophage Giles Lysin B is a novel mycolarabinogalactan esterase, which cleaves the mycolylarabinogalactan bond to release free mycolic acids. The authors proposed that LysB acts at a late stage in the lysis, severing the connection of the mycobacterial outer membrane to the cell wall, providing a faster and more complete lysis of the host cell (Figure 7).

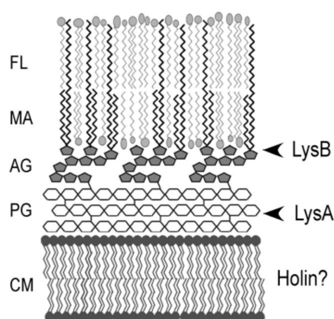


Figure 7. A model for mycobacteriophage lysis of mycobacteria. Mycobacterial cell walls are unusual in that the cytoplasmic membrane (CM) is surrounded by a peptidoglycan layer (PG) to which a network of arabinogalactan (AG) is covalently attached. A mycobacterial outer membrane consisting of mycolic acids (MA) and free lipids (FL) is covalently attached *via* an ester linkage of mycolic acids to arabinogalactan. LysA — assisted by holins encoded by at least some of mycobacteriophages — perform an essential step in lysis involving degradation of the peptidoglycan layer, and the lysis is completed through LysB-mediated cleavage of the outer membrane from arabinogalactan (Payne *et al.,* 2009).

The ***gp4*** **(*hol*)** gene begins with ATG codon located 10 nucleotides downstream of the *lysB* stop codon and encodes a 77 aa protein with holin-like activity. Ms6 *hol* shares some structural characteristics with class II holins (Garcia *et al.,* 2002). The Ms6 Gp4 possesses two TMDs, TMD2 and TMD1, which TMD1 has characteristics of a SAR domain. Unlike some holins, such as λ S, Gp4 lacks a dual-start motif.

Catalão and colleagues suggests that Gp4 is a pinholin because the present of a SAR domain is followed by a typical TMD, which is analogous to other pinholins already characterized, such as the holin of phage 21 (Catalão *et al.*,2011b).

At last, ***gp5*** with 372 bp starts at an ATG codon that overlaps the *gp4* TGA stop codon in a different reading frame. It encodes a 124 aa protein (Garcia *et al.*, 2002) with a predicted single TMD at the N-terminal region. Catalão and their colleges hypothesized that Gp5 might function as a holin-like protein because they demonstrated that the overexpression of Gp5 in *E. coli* results in a drastic inhibition of the cell growth (Catalão *et al.*, 2011b).

As mentioned above, Catalão and colleagues showed that endolysin *LysA* is exported with the help of the *gp1* gene product, in a holin-independent manner. *gp1* is shown to encode a chaperon-like protein that binds the endolysin, assisting its export to the extra-cytoplasmic environment and which is required for an efficient lysis in Ms6. Currently is not known how the endolysins remain inactive until holin determines the time of lysis. Thus, the purpose of this project is to investigate the role of the *gp1* gene product in the export of *LysA* and other proteins, without signal peptide, from the cytoplasm to the periplasm.

Objectives

The main goal of this study is to investigate the *gp1* gene product localization in *M. semgmatis* and evaluate its role in the transportation of *LysA* and other proteins without a peptide signal, from the cytoplasm to the periplasm.

To achieve this, the specific objectives are:

- Construct a Ms6 mutant phage where the *gp1* gene carries, at its 3' end, a sequence coding for a C-Myc tag allowing the production of a recombinant protein that can be detected with an antibody anti C-Myc in *M. smegmatis* infected cells.
- Evaluate the capacity of Gp1 to translocate proteins across the cytoplasmic membrane, by fusing Gp1 with reporter proteins.
- Determine LysA localization in the presence of Gp1.

METHODS/EXPERIMENTAL PROCEDURES

Bacteriophages, bacterial strains, plasmids and growth conditions

Phages, bacterial strains, plasmids and oligonucleotides used in this study are listed in Table 1. Ms6 phage was isolated from *Mycobacterium smegmatis* strain HB5688 (Portugal *et al.*, 1989; Snapper *et al.*, 1990). *M. smegmatis* was grown in Middlebrook 7H9 (Difco™) supplemented with 0.5% glucose and 1 mM $CaCl_2$ at 37 ℃ overnight under vigorous shaking, or on M iddlebrook 7H10 also supplemented. The phage stocks were obtained from confluent lysis in supplemented 7H10 after infection of *M. smegmatis* with the appropriate phage dilution, and then eluted for at least 4h at 4ºC with SM buffer (50 mM Tris.HCl pH 7.5, 10 mM MgSO4.7H2O, 100 mM NaCl) supplemented with 1 mM $CaCl_2$. *E. coli* strains were grown at 37ºC with shaking, in Luria-Bertani (LB) broth or agar supplemented with 100 µg/mL ampicilin or 30 µg/mL kanamycin for plasmid selection.

Table 1. Strains, mycobacteriophages, plasmids and oligonucleotides used in this study.

Strain, mycobacteriopphage, plasmid or oligonucleotide	Description	Reference/source
Mycobacterium smegmatis		
mc^2155	High-transformation-efficiency mutant of *M. smegmatis* ATCC 607	Snapper *et al.,* 1990
mc^2155 (pJV53)	Mycobacteria plasmid expressing recombineering functions	van Kessel and Hatfull, 2007
Bacteria		
Escherichia coli		
JM109	*recA endA1 gyr96 thi hsdR17 supE44 relA1 Δ(lac-proAB) [F' traD36 proAB lacI*^{q} *ZΔM15*	Stratagene
BL21 (DE3)	F^- *ompT hsdS*$_B$ ($r_B^-m_B^-$) *gal dmc* (DE3)	Novagem
Mycobacteriophages		
$Ms6_{wt}$	Temperate bacteriophage from *M. smegmatis*	Portugal *et al.,* 1989
$Ms6_{Gp1\ C\text{-}Myc}$	C-Myctag insertion at the 3' end of Ms6 *gp1* gene	
$Ms6_{LysAHis6}$	His_6tag insertion at the 3' end of Ms6 *lysA* gene cIts857 Sam100	Catalão *et al.,* 2010
Plasmids		
pQE30	Expression vector, T5 promotor; Amp^r	QIAGEN
pET29a(+)	Expression vector, T7 promotor; Kan^r	Novagem
pCSJ8	The first 30 aa of Lys44 fused to PhoA in pAH312	São-José, 2002
pMJC49	PhoA in pET 29ª	Catalão, unpublished
pMJC1	gp1 cloned in pET29b (*Bam*HI/*Hind*III)	Catalão *et al.*, 2010
pMJC50	gp1- PhoA cloned in pet 29a	Catalão, unpublished
$pAS3_{GP1\text{-}PhoA}$	gp1- PhoA cloned in pQE30 (*Bam*HI/*Hind*III)	This study

$pAS1_{SP\text{-}PhoA}$	SPLys44-PhoA cloned in pQE30 (*Bam*HI/*Hind*III)	This study
$pAS2_{SP\text{-}PhoA}$	SPLys44-PhoA cloned in pET29a(+) (*Bam*HI/*Hind*III)	This study
pGLO	Plasmid containg the GFPuv gene	Biorad
$pAS4_{GFP}$	*gfp* cloned in pQE30	This study
$pAS5_{Gp1\text{-}GFP}$	*gp1-gfp* in pQE30	This study
pMJC3	*gp1* and *lysA* cloned in pET29b	Catalão *et al.*, 2010
pMJC4	*lysA* cloned in pET29b	Catalão *et al.*, 2010
Oligonucleotides		
Pr Gp1 C-Myc	CTCCATCCCCGTCCTCGGCGGAATCCTCGGGAGCAAACGGGAACAG AAACTGATCAGCGAAGAGGATCTGTGACGGGAGCAAACGGTGA CCACGAAAGATCAAGTCGCCC	This study
Pr Gp1 C-Myc Extended fwd	CTGACCAACCTTCCAGCGCAAGTCATGGACATCATCGACAGCGCGCT GCGCTCCAGACAGCGCGCTGCGCTCCATCCCCGTCCTCGGCGGAATC	This study
Pr Gp1 C-Myc Extended 3' rv	GCATTCGCTGCGGGTGTAGCCGCGCGCCTTGGCTTCGGCGATGGTGA TTTGGGCGACTTGATCTTTCGTGGTCAC	This study
Pr Screening Myc fwd	CTCGGGAGCAAACGGGAACAGAAACTG	This study
Pr Gp1A fwd	CG**GGATCC**ATGGACCGCTTAGGCATCGTCC	Catalão *et al.*, 2010
Pr LysA60aa rv	GTCGAAGCGGTGTGGGTAGGAGCCG	
Pr SPLys44 fwd	GGCGGATCCACACGTAAAAAGTTAAA	
Pr phoA3' rv	CCAAGCTTTTTCAGCCCCAGAGC	
PrGp1 rv, EcoRI	TCGTGGG**GAATTC**CGTTTGCTCCC	This study
PrGFP fwd, EcoRI	GAAAAAAT**GAATTC**CAAAGGAGAAGAAC	This study
PrGFP rv, PstI	GGAATTCA**CTGCAG**TTATTTGTAGAGCTC	This study
PrGFP fwd, BamHI	CCAG**GATCC**GCTAGCAAAGGAG	This study

DNA extraction from Ms6 wt

The extraction of Ms6 genomic DNA was adapted by a protocol from Sambrook and Russel (2003). 600 µl of phage stock were incubated with proteinase K (50 µg/mL) and 0,5% SDS for 1 hour at 56ºC. The mixture was treated three times with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) and one time with chloroform. Finally, the DNA was precipitated with 10% 3M sodium acetate and an equal volume of cold isopropanol. After a maximum centrifugation step of 45 min at 4º, the DNA was washed with 200 µL of 70% ethanol, centrifuged for 15 min, and then the pellet was dried at 37ºC and resuspended in water. The samples were stored at -20ºC and were quantified in Spectrophotometer ND-1000 (Nanodrop).

One-step growth experiments

One-step growth experiments were performed as described in Catalão *et al.,* 2010. *Mycobacterium smegmatis* cells were grown to an OD_{600} 0,5-0,6 and then diluted to an OD_{600} between

0,2-0,3. Then 10 mL were centrifuged and resuspended in 1 mL of a phage dilution, using a multiplicity of infection (m.o.i.) of 1. The mixture was incubated for 50 min at 37ºC to allow adsorption of the phages. 100 µL of 0,4% H_2SO_4 was added and incubated 5 min at 37ºC, to inactivate the non-adsorbed phages. The suspension was neutralized with 100 µL of 0,4% NaOH and diluted to 1:100. 1 mL of the suspension were withdrawn every 30 min until reaching 210 min and 100 µL of serial dilutions of each sample were plated with 200 µL of *M. smegmatis* cells as top agar lawns on 7H10 medium. Phage titer for each sample was determined after 24 h at 37°C.

Gp1 expression in M. smegmatis-infected cells

Examination of Gp1 synthesis in *M. smegmatis* was performed as previously described (São-José *et al.*, 2000). An exponentially-growing culture of *M. smegmatis* was infected with Ms6LysAHis6 or Ms6Gp1 C-Myc at a multiplicity of infection (m.o.i.) of 10 and incubated at 37ºC for 30 minutes. Ten milliliter samples were withdrawn at 30-min intervals, centrifuged (8000rpm 13' at 4ºC) and the pellet was immediately frozen at -20ºC. After thawing, cells were resuspended in 100 µL of 10mM Tris-HCl buffer supplemented with 20 mg/mL of lysozyme and 1 µL of a protease inhibitor cocktail (Calbiochen) (Sambrook *et al.*, 2001). After an incubation period at 37°C for 3 hours, 25 µL of 5x SDS-PAGE sample buffer were added followed by incubation at 100°C f or 5 min to complete cell lysis. The proteins were analyzed by 15% SDS–PAGE followed by Western-blotting, using a horse-radish-peroxidase (HRP) conjugated anti-C Myc antibody (Roche).

Cellular fractionation of M. smegmatis

A 200 mL culture of *M. smegmatis* infect with Ms6Gp1C-Myc was centrifuged and washed twice in a large volume of PBS and resuspended in 5 mL of ice cold Phosphate Buffered Saline (PBS) with 50 µL of a cocktail of proteases inhibitor (Calbiochem).The cell suspension was sonicated with a microtip probe at 50% output for 3 min, incubated in ice for 10 min and spun in centrifuge 30 min maximum speed at 4ºC. The supernatant, which has the cytosol-enriched fraction, was filtered through a 0.22 mm filter to remove membrane fragments that did not pellet, the volume was measured, brought up to 0.8 mL in ice cold PBS and 0.2 mL of 10% Triton X-100 was added (final concentration of 2.0% Triton X-100). The pellet, having the membrane and cell wall-enriched fraction, was resuspended in 0.8 mL of ice cold PBS and 0.2 mL of 10% Triton X-100 (final concentration 2.0% Triton X-100). The fractions were incubated on ice for 1h with intermittent vortexing (approximately every 15 min) and spun in an Eppendorf centrifuge 30 min maximum speed at 4ºC. The supernatant from the cytosol-enriched fraction was transferred to a new tube (thereby having the membrane enriched fraction) and the pellet (cell wall-enriched fraction) was resuspended in 1 mL of ice cold PBS. All fractions were then precipitated with 9 mL of cold acetone at –70ºC for 2h and were afterwards analysed in western-blot (Hatfull, personal communication).

Protease accessibility experiments

The procedure was performed using a protocol from Chou and Kendal (1990). Expression of Gp1 and LysA in *E.coli* was induced from plasmids pMJC3 and pMJC4. *E. coli* cells were grown to an OD_{600} of 0,5 and a 1 mL aliquot was collected, meant for a negative control. The *E.coli* cultures were induced for 1h and then collected and immediately chilled on ice. An equal volume of 10% trichloroacetic acid (TCA) was added to the uninduced sample and one of the induced samples, and then the precipitate was collected by centrifugation (15 000 x g, 30 min at 4ºC), precipitated with ice cold acetone, air dried and resuspended in 100 µL of 2x SDS-PAGE sample buffer. The two remaining aliquots were washed in 30 mM Tris-HCl pH 8 and then resuspended in 0.5 M sucrose, 30 mM Tris-HCl pH 8, containing 20 µg/mL lysozyme and 1 mM EDTA to digest the cell wall and to expose the plasma membrane. Following 20-min incubation on ice, proteinase K (25 g/mL) was added in one of the aliquots while the same volume of enzyme storage buffer was added to the other aliquot. After 20 min of incubation on ice, followed by 15 min room temperature and subsequently precipitation with TCA the samples were washed with ice cold acetone and resuspended in 2xSDS-PAGE sample buffer. All samples were boiled at 100ºC for 8 min and analysed by Western-bolt.

Plasmid construction

To construct plasmids $pAS_{SP\text{-}PhoA}$ and $pAS2_{SP\text{-}PhoA}$, a DNA fragment containing the 5' end, corresponding the first 30 amino acids of the fOg44 endolysin gene (*lys44*) fused with the mature form of phosphate alkaline gene was amplified by Polymerase chain reaction (PCR) from plasmid pCSJ8 (São-José C., 2002). The reaction was performed with Pr SPLys44 fwd and Pr phoA3' rv designed to contain the recognition sequence of *Bam*HI and *Hind*III (Fermentas), respectively. The PCR product was restricted with *BamHI* and *HindIII* and inserted into vector pQE30 (QIAGEN) and pEt29a (Novagen) restricted with the same enzymes. The ligation reaction was performed at 25ºC for 2h with *T4 DNA ligase* (Biolabs) according to the manufacter instructions. Transformations were performed in competent in *E. coli* cells previously treated with 50 $mMCaCl_2$. Recombined cells were selected in LB agar supplemented with the w appropriate antibiotic. The recombinant plasmids were then extracted using FastPlasmid Mini Kit (Eppendorf) and insertion of the DNA fragment of interest was checked by PCR with the same pair of primers used for cloning. To construct plasmid $pAS3_{GP1\text{-}PhoA}$, a DNA fragment containing the Ms6 *gp1* fused with the mature form of phosphate alkaline gene was amplified by PCR from pMJC50 (Catalão, unpublished). The reaction was performed with Pr Gp1A fwd and Pr phoA3' rv designed to contain the recognition sequence of *Bam*HI and *Hind*III (Fermentas) respectively and inserted into pQE30 previously restricted with the same enzymes. After purification, the product was cloned into pQE30, restricted with the same restriction enzymes, and the ligated as above. Plasmid $pAS4_{GFP}$ was constructed by insertion of a DNA fragment coding for the GFPuv. Into pQE30. GFPuv was amplified from pGLO with primers PrGFP

fwd, *BamHI* and PrGFP rv, *PstI*. After purification the product was cloned into pQE30 restricted with *BamHI* and *PstI* and then ligated as previously described. Transformations were performed in competent *E. coli* JM109 cells (for pQE30 derivatives) or BL21 (DE3) (for pET29 derivatives), previously treated with 50 mM$CaCl_2$ (Samborook *et al.*, 2003). Recombinant cells were selected in LB agar supplemented with the appropriate antibiotic. The recombinant plasmids were then extracted using FastPlasmid Mini Kit (Eppendorf) and insertion of the DNA fragment of interest was checked by PCR with the same pair of primers used for cloning.

Plasmid pAS5Gp1-GFP was constructed by ligation of three DNA fragments generated as follows: 1) *gp1* gene was amplified from Ms6 DNA with primers PrGp1 fwd, *BamHI* and PrGp1 rv, *EcoRI* and restricted with *BamHI* and *EcoRI*; 2) GFPuv gene was amplified from plasmid pGLO with primers PrGFP fwd, *EcoRI* and PrGFP rv, *PstI* and restricted with *EcoRI* and *PstI*; 3) plasmid pQE30 was restricted with *BamHI* and *PstI*. The ligation of the three fragments was performed at 25ºC for 2h followed by incubation at 16ºC overnight. The reaction mixture was inserted into electrocompetent *E.coli* JM109 cells by electroporation (Hanahan, 1983; Sambrook *et al.*, 2001). Recombinant cells were selected in LB agar with ampicillin and plasmid extraction was performed as above. The presence of the inserted fragment was checked by enzymatic restriction.

Detection and quantification of Alkaline Phosphate activity

For the PhoA screening: *E. coli* recombinant strains were grown in LB agar supplemented with 1mM IPTG and 40 μg/mL BCIP (5-Bromo-4-chloro-3-indolyl phosphate) and the appropriate antibiotic. *For the quantification assay: E coli* cells were grown in LB broth to an OD_{600} of 0.4 - 0.5 at 37ºC with shaking (~ 250rpm) and expression was induced with 1 mM IPTG. 1 mL of culture was collected at times 0', 30' and 90' following induction, and centrifuged at 12 000x g for 1 min at 4ºC. Supernatant was discarded and the cells were resuspended in 1 mL of MOPS-buffer. Cell density was read at 600_{nm} either directly or after dilution with MOPS-buffer. 100µL of 40 mg/mL p-nitro-phenyl-phosphate (PNPP) was added to the assay tubes and the tubes were incubated at 37ºC until a significant yellow color was observed. To stop the reaction, 100 µL of K_2HPO_4 was added. Then the OD was read at 420_{nm} and 550_{nm}. Units of phosphatase activity were calculated by the following formula (Brickman *et al.,* 1975; Belin 2010):

$$1000 \times \frac{OD420 - 1.75 \times OD500}{time\ (\mathrm{min}) \times \mathrm{OD600(of\ culture)}} \times dilution\ factor$$

Expression and determination of solubility from Gp1 and Gp1-GFP in E. coli cells

E. coli recombinant cells carrying plasmids PQE30, pAS4GFP or pAS5Gp1-GFP were grown in LB medium supplemented with ampicillin, at 37℃ wit h shaking until OD_{600} equals 0.6. At this time an

aliquot of 1 mL was collected (noninduced control) and the expression was induced with 1 mM IPTG for 1h and 3h. 1 mL samples were collected, centrifuged at 4000 x g for 10 min at 4℃ and the pellet was resuspended in 100 µL of phosphate buffer (50 mM $NaHPO_4$ and 30 mM $CaCl_2$) and incubated with lysozyme (1mg/mL) and 1 µL of a protease inhibitors cocktail at 4℃ for 30 m in. 4 cycles of freezing (with liquid nitrogen) and thawing (at 42℃) were then ca rried out and the lysates were centrifuged at 15 000 x g for 30 min at 4℃. The supernatant corresponds th e soluble fraction (cytoplasmic proteins), while the pellet corresponds to the insoluble fraction (membrane proteins). The pellet was resuspended in 100 µL SDS-PAGE buffer and the supernatant was added by the same volume of 2 X SDS-PAGE buffers. The samples were boiled for 10 min and analysed by SDS-PAGE followed by Coomassie-Blue staining.

PCR reactions

The reaction mixture for PCR reactions, whose composition is summarized in Table 2, was subject to the following conditions: 1 min at 95 ° C, 30 cycles of 1 min at 95 °C, 1min at 56 °C and 1min at 72 ° C, and 10min at 72 ℃ and then stored at 4° C. The reactions were performed in the BioRad thermocycler MyCycler.

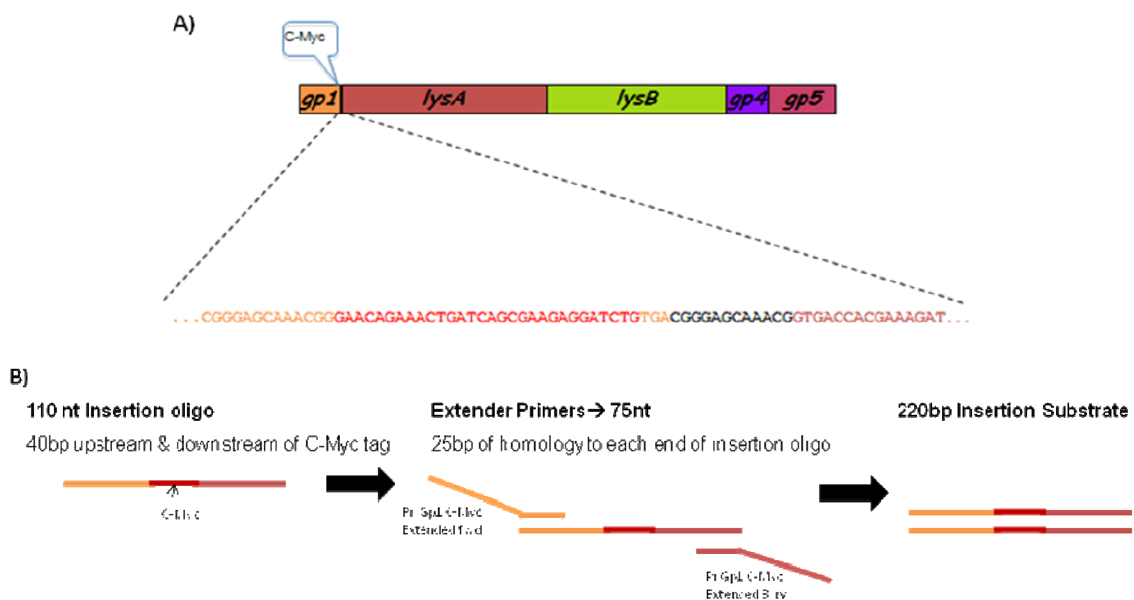
Table 2.

Reaction mixture	Final concentration
polymerase buffer	1x
$MgCl_2$	2,5mM
DNA	2 ng/µL
dNTP	0,25mM
Primer fwd	1pmol/µL
Primer rv	1pmol/µL
DNA plymerase Pfu or taq	0,025U/µL

RESULTS AND DISCUSSION

Construction of the Ms6Gp1-CMyc mutant

To determine the localization of Ms6 Gp1 during infection of *Mycobacterium smegmatis, a Ms6 derivative mutant was constructed,* where the *gp1* gene carries a sequence at its 3’ end that encodes a C-Myc tag. Construction of Ms6 mutant phages was performed using Bacteriophage Recombineering of Electroporated DNA (BRED) as described previously by Marinelli and colleagues (2008). 10-20 ng of a 110 bp oligonucleotide with 40 bp of upstream and downstream homology to the insertion region, was extended by PCR using two 75-bp extender primers, which have 25 bp of homology at each end of the 100-mer (see Table 1 and Figure 8A). The PCR reaction was performed as follows: 1 cycle of 7 min at 95 °C, 30 cycles of 1 min at 95 °C, 1 min at 56 °C and 1 min at 72 °C, and 10 min at 72 ℃ and then stored at 4℃. The final 220-bp dsDNA product was p urified using MinElute PCR Purification Kit (Qiagen) and then co-electroporated with 70 ng of Ms6 wt genomic DNA into electrocompetent *M. smegmatis* (pJV53). The cells were resuspended in supplemented 7H9 medium, incubated at 37ºC for 2 hours with shaking and plated as a lawn in 7H9 molten agar with 200 µL of wild-type *M. smegmatis*. After 24h of incubation at 37ºC, 40 phage plaques were picked and eluted into an eppendorf with 100 µL phage buffer (SM) with 1 mM $CaCl_2$, for 2 hours at room temperature. To detected plaques containing the phage mutant, 21 of them were screened by Polymerase chain reaction (PCR) with primers Pr Screening Myc fwd, which hybridizes with the C-Myc coding sequence, and Pr LysA60aa rv, which hybridizes with the LysA DNA at position 48 and 73 to detect the C-Myc tag insertion (Table 1).
The strategy used to construct the Ms6 mutant phage is represented in Figure 8.



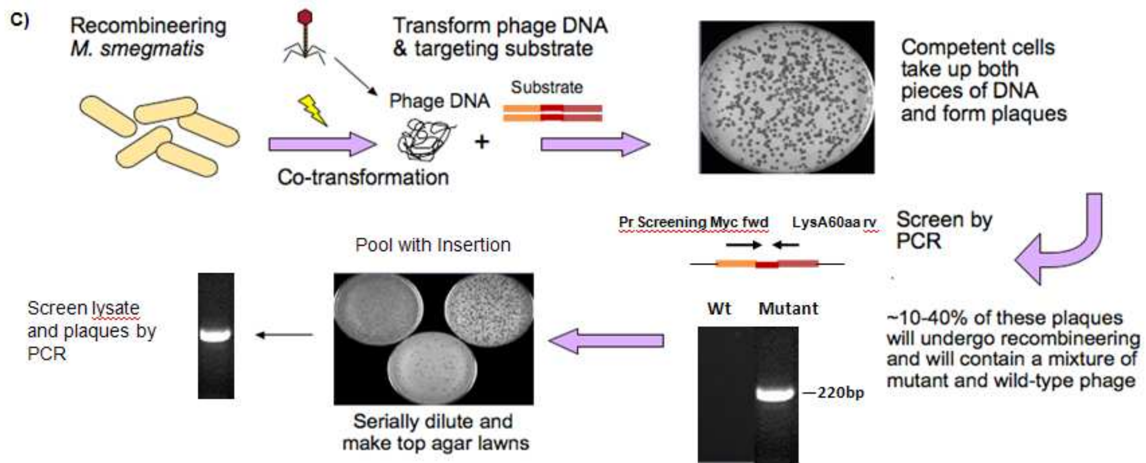


Figure 8. Representative scheme of the construction of Ms6 mutant phage. **A)** Schematic representation of the position of C-Myc tag in the end of the *gp1* gene. Nucleotides are represented in colours according to the scheme, C-Myc is represented in red. **B)** Construction of an insertion substrate (C-Myc) by PCR. **C)** Phage recombineering strategy. Adapted from Marinelli *et al.*, 2008.

Four of them (19%) contained detectable levels of the mutant phage, probably mixed with wild-type phage (Figure 9 A). To select a pure mutant phage, we performed several serial dilutions and then re-plated. The individual and secondary lysate were screened by PCR with the same primers. Since the mutant allele was still present in the lysate (Figure 9 B) the mutant was considered viable. To recover Ms6Gp1-CMyc mutants, 20 secondary plaques from the re-plating were picked and screened by PCR with the same primers mentioned above and three mutants were obtained. To verify if the mutant was pure we re-plated one secondary mutant and screened 8 plaques by PCR. 100% of homogenously pure mutant was obtained (Figure 9 C). The correct insertion of the C-Myc tag was confirmed by sequencing a PCR product amplified from the pure mutant phage with the primers, Pr LysA60aa rv and Pr Gp1A fwd, flanking the insertion region.

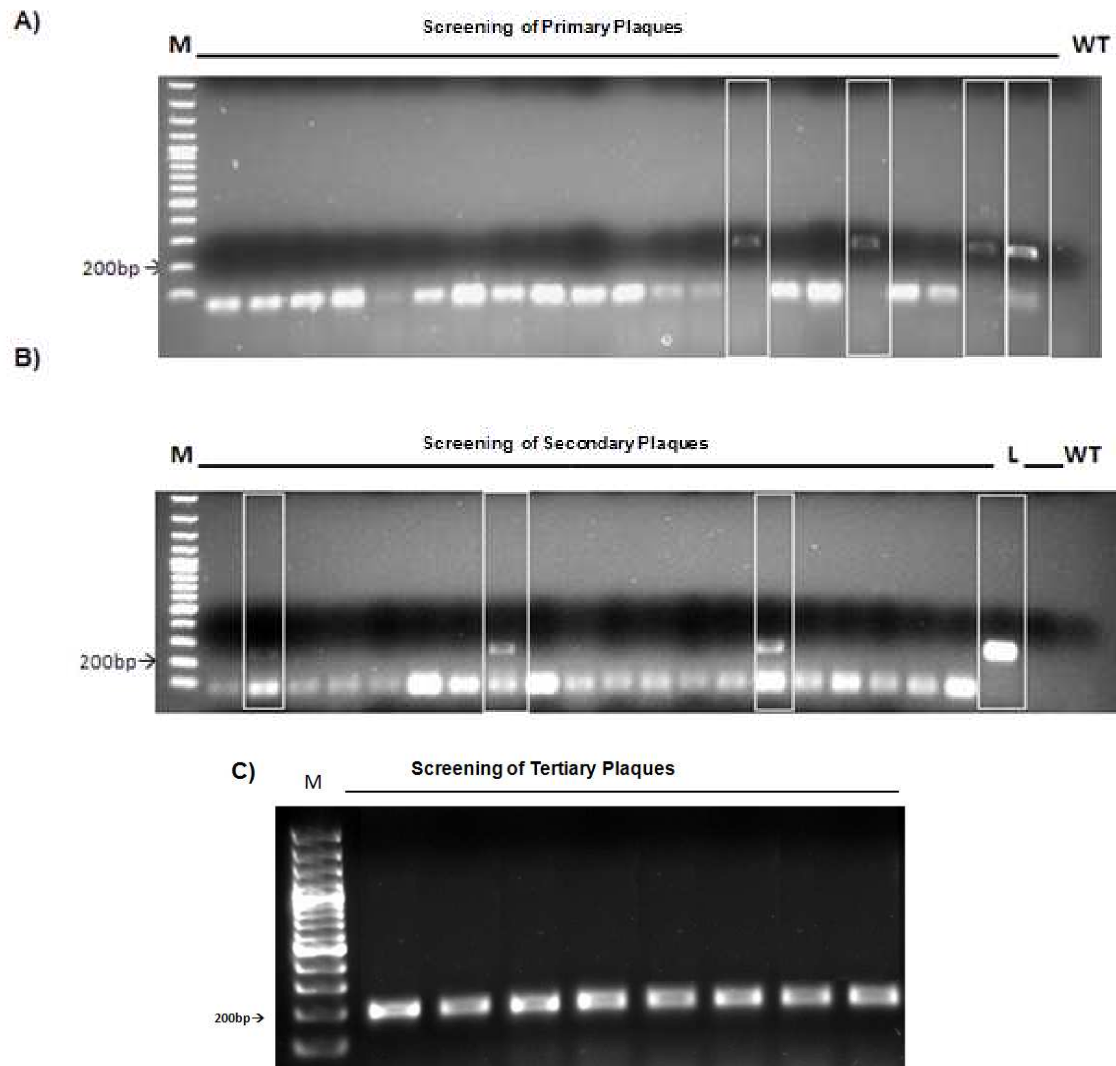


Figure 9. Selection and isolation of Ms6 derivate mutant Ms6Gp1-CMyc. **A)** Primary plaques were recovered and screened by PCR using screening primers for the insertion of the C-Myc Tag. Plaques containing mutant DNA are boxed. On the right is the wild-type Ms6. **B)** Mixed primary plaques were diluted and plated, and the lysate (on the right) and isolated secondary plaques were screened with the same primers. Ms6Gp1-CMyc mutants are boxed. **C)** One secondary plaque was diluted, plated and isolated to tertiary plaques and was screened with the same primers. M: DNA molecular weight ladder (Gene Ruler^{TM} 100 bp DNA Ladder, Fermentas).

Evaluation of the Ms6Gp1-CMyc growth parameters

After construction of the mutant Ms6, the next step was to determine if the insertion of a C-Myc tag could affect the phage growth parameters. As we can observe from Figure 10, the morphology of the

plaques is quite similar between the Ms6 wt phage and the Ms6Gp1-CMyc. These results indicate that insertion of the C-Myc tag does not affect plaque morphology.

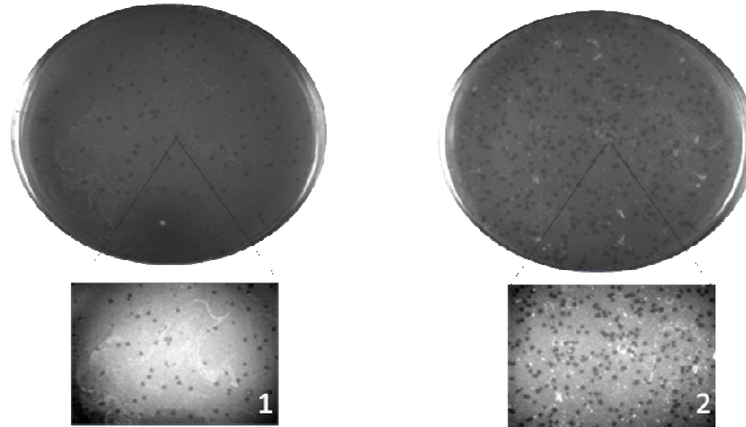


Figure 10. Morphology of the plaques in 1) Ms6 wild-type 2) Ms6Gp1-CMyc phages.

One-step growth curves were carried out to compare the phage multiplication cycles. The one-step experiment (Figure 11) showed that both phages have similar latent period and similar burst sizes. Thus, the addition of the C-Myc tag had no evident effect on the infection cycle.

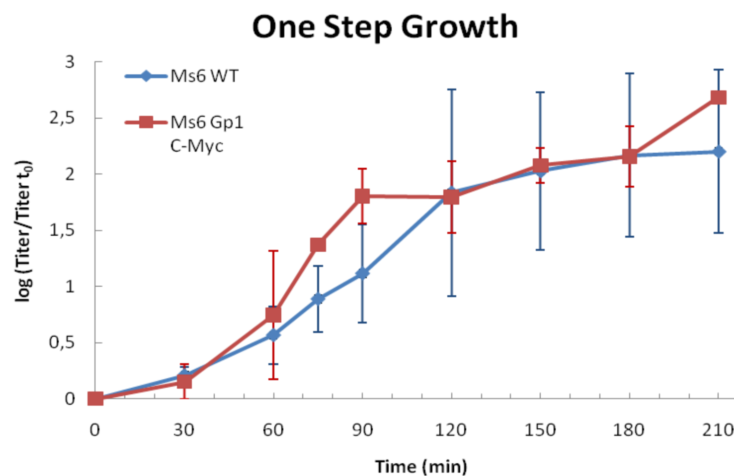


Figure 11. One step growth curves of mycobacteriophage Ms6 wild-type and Ms6Gp1-CMyc. For each curve the results are ploted as the logaritm of the measured titers divided by the titer at t_0, for normalization. The values are an average of three independent experiments.

Gp1 expression in M. smegmatis-infected cells

To follow the mycobacteriophage Ms6 Gp1 production in the course of *M. smegmatis* infection, infected cells were examined for Gp1. *M. smegmatis* was infected with Ms6Gp1-CMyc phage. As

explained above the 3' end of *gp1* gene was fused to a sequence coding for a C-Myc tag, allowing the production of a Gp1-CMyc tagged protein. Samples were collected immediately before and every 30 minutes after infection, until near the end of the Ms6 infection cycle. Protein extracts were prepared from such samples as described in the methods section and analysed for the presence of C-Myc tag proteins by immunoblotting with an anti C-Myc antibody. Although many attempts to detect Gp1-CMyc were made, we were not able to detect the Gp1-CMyc protein. In parallel to technical problems, two hypotheses may explain this unsucess: it's possible that the amount of produced Gp1 is not detectable in the assayed conditions or conformation of Gp1 in *M. smegmatis* is such that the tag is not exposed.

In order to clarify this result we searched for Gp1 in a higher amount of cells, close to the end of the infectious cycle (see below).

Subcellular localization of Gp1 in M. smegmatis-infected cells

To detect the Ms6 Gp1 upon a *M. smegmatis* infection detection in the different cell compartments, 200 mL of Ms6Gp1-CMyc infected culture at m.o.i of 10 was submitted to a cell fractionation by differential centrifugation steps. Protein extracts from soluble, membrane and cell wall fractions were separated by SDS-PAGE and analyzed by Western-blot using an anti-C-Myc antibody. As we can observe in the Figure 12, one band was detected in the soluble (cytoplasm) and membrane fractions, corresponding to the predicted molecular mass of Gp1-CMyc (≈11 kDa). As can be observed the amount of produced protein seems to be low. This indicates that the negative result obtained above may result from a low level of Gp1 production rather than a consequence of an inaccessible tag. Nevertheless, these results indicate that in *M. smegmatis* Gp1 is present in the soluble and membrane fraction, which is in agreement with the previously results obtained in *E. coli* and with the fact that Gp1 is a chaperone (Feldman *et al.,* 2003). Of note is the fact that the total extract represented in Figure 12 does not show the presence of the recombinant protein. This is because the amount of collected total extract is very low (500 µL).

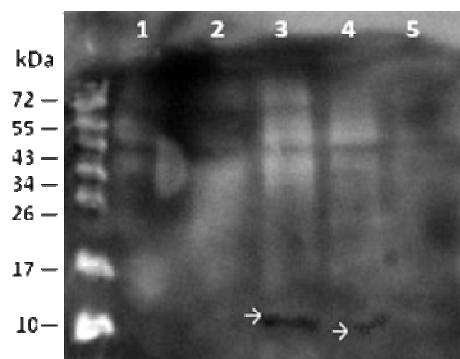


Figure 12. Subcellular localization of Gp1 in *M. smegmatis infected cells***.** Gp1-CMyc (11 kDa) is found both in the soluble and membrane fractions (indicated by the arrow). 1) total extract 2) supernatant of the 200 min infection of *M. smegmatis* 3) soluble fraction 4) membrane fraction 5) cell wall fraction. Protein was detected by Western blotting with an anti-C-Myc antibody. M correspond to a prestained protein ladder (PageRuler, Fermentas).

Expression and localization of Ms6 Gp1 in E. coli.

To complement the results from cellular fractionation of *M. semgmatis*, we analysed the expression and localization of Gp1 carrying a hexahistidine tag at the C-terminus, in *E. coli* cells containing the plasmid pMJC1 (Catalão, *et al.*, 2010), which has the *gp1* gene cloned in pET29b.

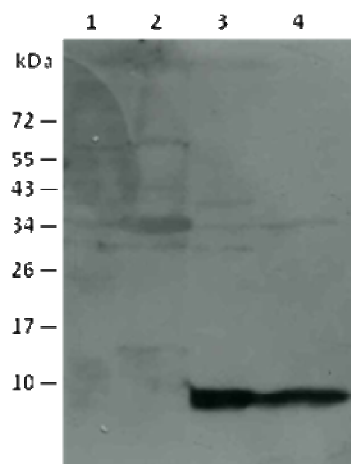


Figure 13. Immunoblot analysis of two different compartments of *E. coli* (pMJC1) *cells* expressing Gp1. Lanes 1 and 2: not induced, lanes 3 and 4 correspond to Gp1 induced with IPTG (1 mM); soluble fraction (lanes 1 and 3) and insoluble fraction (lanes 2 and 4). $Gp1His_6$ was detected with an anti-His antibody. Protein Molecular masses are indicated in kDa in the left.

The bands detected in Figure 13 lanes 3 and 4, correspond to the presence of $Gp1His_6$ protein which has a predicted molecular mass of 10kDa, in the soluble and insoluble fraction, respectively. This result is in agreement with the results obtained in *M. smegmatis* and as mentioned before with the fact that Ms6 Gp1 is a chaperone protein.

To complement the obtained results of Gp1 analysis, we used two reporter proteins: bacterial alkaline phosphatase (PhoA) as a reporter of periplasmic localization (Manoil *et al.*, 1990) and green fluorescent protein (GFP) as a reporter of cytoplasmic localization (Phillips *et al.,* 2001).

Analysis of a Ms6 Gp1 fusion to Alkaline Phosphatase

E. coli alkaline phosphatase (PhoA) is widely used as a reporter protein for studying protein export in bacteria. PhoA exhibits an export-dependent activity due to its requirement for the disulfide bond-forming enzymatic system which is localized in the periplasm. This enzyme is only active after its export to the periplasm, which provides the best quantitative reporter for protein translocation studies (Belin, 2010). We used this activity dependence to further assess the ability of Gp1 to promote the translocation of a signal sequence-less form of PhoA. For this purpose we used $pAS3_{GP1\text{-}PhoA}$ and pMJC50 (Catalão, unpublished) plasmids, derivatives of vectors pQE30 and pET29a(+) respectively, which allows expression of Gp1 fused to the mature form of PhoA under the control of a T5 and T7 promoter (Figure 15). Two positive controls derivatives of pQE30 and pET29a(+) were constructed, by fusing the signal sequence of the Lys44 endolysin from bacteriophage fOg44 (São-José, 2002), generating plasmids $pAS1_{SP\text{-}PhoA}$ and $pAS2_{SP\text{-}PhoA}$ (Figure 15).

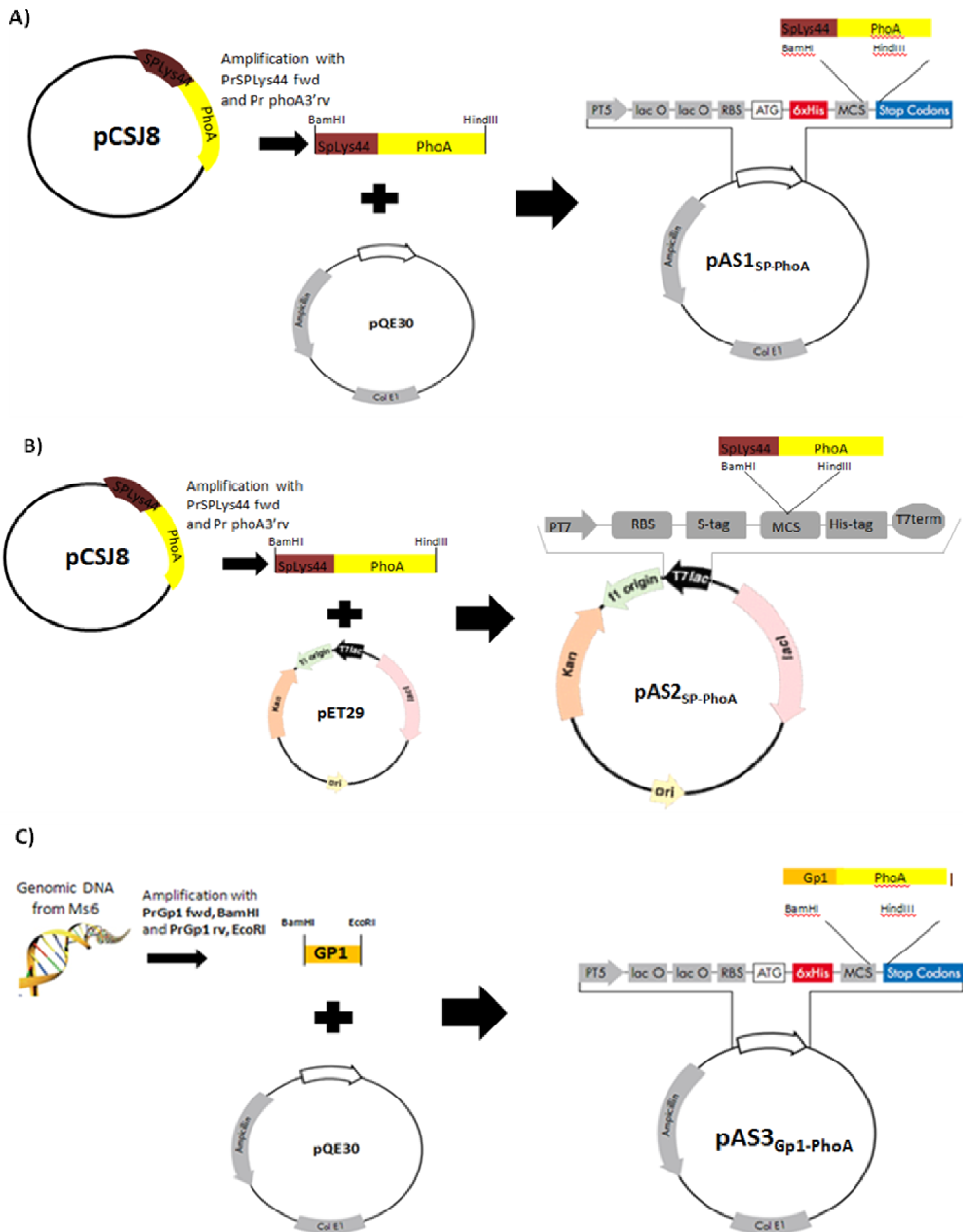


Figure 15. Representative scheme of the three recombinant plasmids constructed in this work: **A)** pAS1$_{SP\text{-}PhoA}$, **B)** pAS2$_{SP\text{-}PhoA}$ and **C)** pAS3$_{GP1\text{-}PhoA}$.

Plasmids pAs1 and pAs3 were introduced in *E.coli* JM109 and plasmids pMJC50 and pAS2 in *E.coli* BL21 (DE3). The recombinant cells were than plated on LB agar supplemented with the appropriate antibiotic, IPTG and with the chromogenic substrate 5-bromo-4-chloro-3-indoxyl phosphatase (BCIP) to search for alkaline phophatase activity. As expected cells expressing the SPLys44-PhoA fusion gave blue colonies (Figure 16, panel A and B). Cells harbouring plasmids pMJC1 (Figure 16, panel E) and pMJC49 (Figure 16, panel F), carrying Gp1 and PhoA without its signal peptide respectively, used as negative controls, produced white colonies. The cells that express Gp1-PhoA (figure 16 Panel C and D) gave colonies slightly blue - more visible when the fused protein is expressed from the stronger promoter T7 - meaning that there is alkaline phosphatase activity. Although the generated color is not so intense as the positive control, the existence of a blue color suggests that Gp1 may allow the translocation of PhoA across the cytoplasmic membrane.

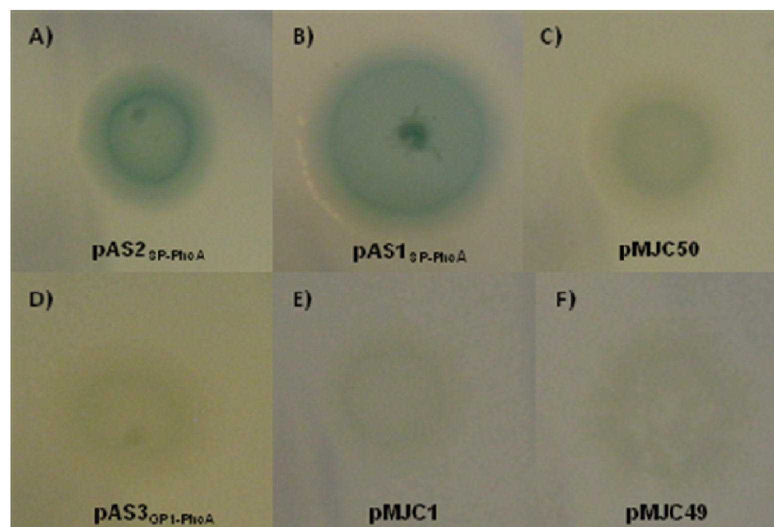


Figure 16. Detection of alkaline phosphatase activity in *E. coli* strains carrying plasmids $pAS2_{SP\text{-}PhoA}$ $pAS1_{SP\text{-}PhoA}$, , pMJC50, $pAS3_{GP1\text{-}PhoA}$, pMJC1 or pMJC49, after plating on LB agar supplemented with BCIP and IPTG. Blue colonies indicate the activity of the alkaline phosphatase.

Quantification of PhoA activity was carried out with the same recombinant *E. coli*, using the 4-nitro-phenil-phosphate (PNPP) as substrate, according to the protocol described by Brickman (1975) Activity of Alkaline Phosphatase generates p-nitrophenol, a yellow product that can be quantified in a spectrophotometer at 405nm. Several attempts to quantify the phosphatase activity were made but the results gave a high variation. However cells harbouring plasmids $pAS2_{SP\text{-}PhoA}$ and pMJC50 produced an intensive yellow color (data not shown).

In order to confirm expression of the fused proteins, protein extracts of the induced cells were analysed by SDS-PAGE followed by coomassie blue staining and western blot using an antibody against a His6 tag (Figure 17).

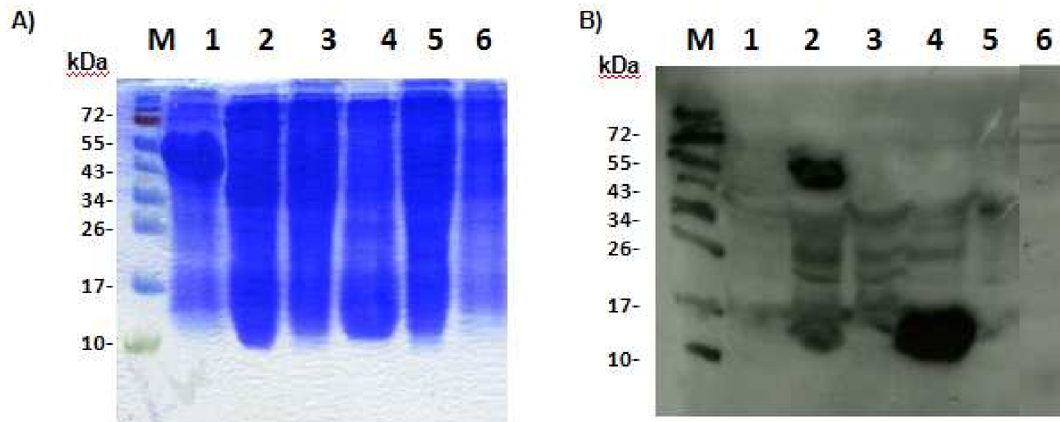


Figure 17. Detection of the of the fusion proteins in *E.coli* extracts after SDS-PAGE A) Coomassie-blue stainind and B) Western-blot using an antibody against His_6-tag. Lanes: 1) BL21(DE3) (pMJC49); 2) BL21(DE3)($pAS2_{SP\text{-}PhoA}$);3) BL21(DE3) (pMJC50); 4) BL21(DE3) (pMJC1), 5) JM109 ($pAS1_{SP\text{-}PhoA}$); 6) JM109 ($pAS3_{Gp1\text{-}PhoA}$). M correspond to a prestained protein ladder (PageRuler, Fermentas).

The results indicate that expression of the PhoA protein (Figure 17A, lane 1) could be detected in the coomassie staining but not in the western-blot, because the vector didn´t have a tail of His6. Expression of the SPLys44-PhoA fusion although not clear in the coomassie staining could be detected in the western blot (Figure 17B lane 2) with an approximated molecular mass of 50 kDa (45kDa of PhoA + 5 kDa of the signal of Lys44), which corresponds to the predicted molecular mass of SP-PhoA (Figure 17B, lane 2). Expression of the Gp1 protein (Figure 16B, lane 4) with 10 kDa could also been detected (Figure 17B lane 4). In the case of SP-PhoA in pQE30, $pAS1_{SP\text{-}PhoA}$ (Figure 17B, lane 5) and In the $pAS3_{Gp1\text{-}PhoA}$(Figure 17B, lane 6) the proteins were not detected in both coomassie and western blot. This is because in these constructions the His Tag is fused at the N-terminus and, while the SpLys44-PhoA is cleaved and the signal peptide is lost, in the Gp1-PhoA fusion the His tag in not detected at the N-terminus, a result that was previously observed in constructions where the sole Ms6Gp1 is fused to a hexahistidine tag at the N-terminus, as it happens in pWE30 derivative plasmids.

Analysis of a Ms6 Gp1 fusion to Green-fluorescent protein

To complement the ability of Gp1 to translocate proteins across the cytoplasmic membrane we

used another reporter gene, Green-fluorescent protein (GFP), a protein encoded by a gene originally isolated from the jellyfish, *Aequorea victoria*. For this experiment we used a modified form of the GFP gene, which greatly enhance fluorescence of the protein. GFPuv is a reporter molecule, with a molecular mass of 27 kDa, which can monitor gene expression and protein localization in living organisms. This protein produces a bright green fluorescence when viewed with UV light and remains active as long the protein remains in the cytoplasm (Feilmeier *et al*., 2000).

We constructed two recombinant plasmids, $pAS4_{GFP}$, which contains the GFPuv gene, amplified by PCR from plasmid pGLO (BioRad) and cloned in pQE30, and $pAS5_{GP1\text{-}GFP}$, containing Ms6 *gp1* fused with *gfp*uv cloned in the same vector (Figure 18)..

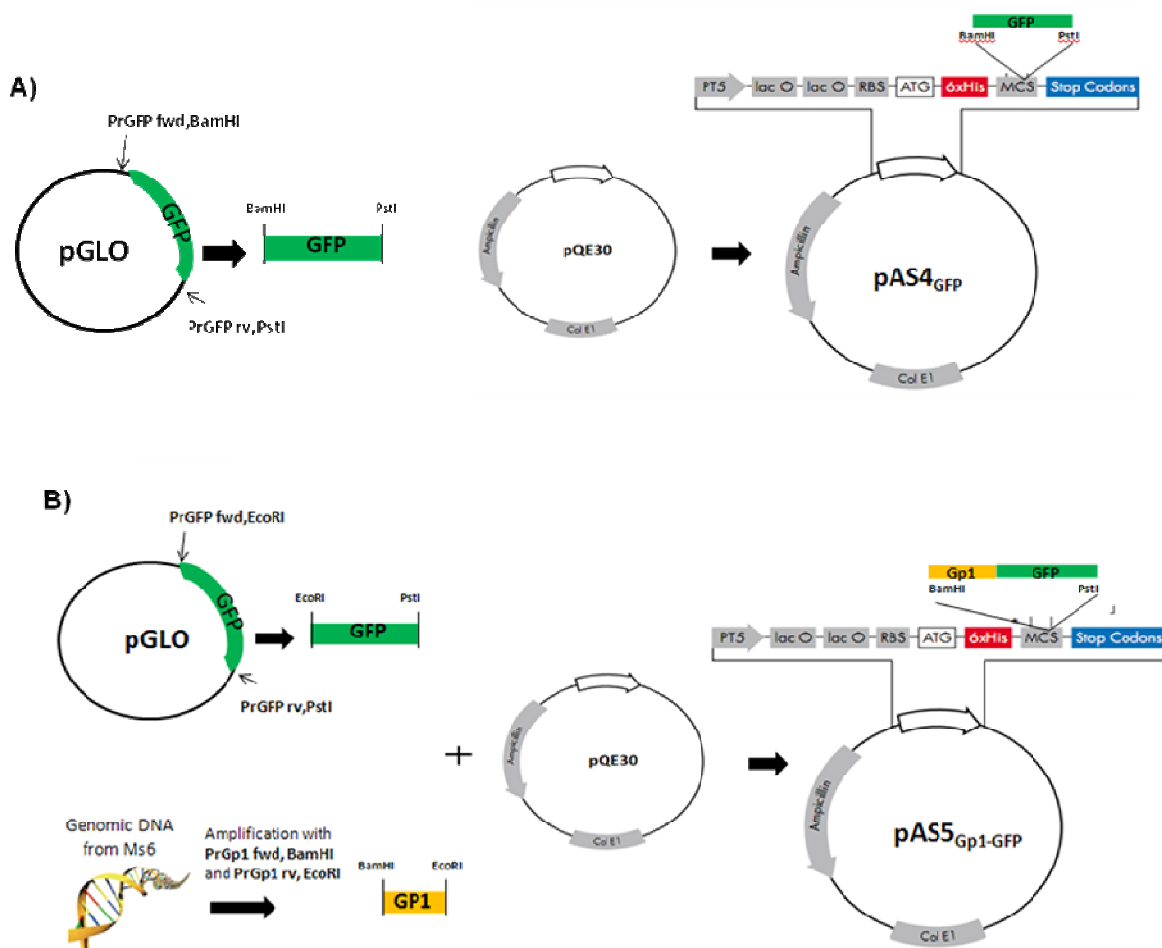


Figure 18. Construction of $pAS4_{GFP}$ and $pAS5_{Gp1\text{-}GFP}$.

After transformation of *E coli* JM109 with the constructed plasmids, the recombinant cells were platted on LB agar with ampicillin and 1 mM of IPTG to induce expression of the proteins. After 24h of incubation at 37ºC, the LB plaques were placed under UV light and checked for the emission of fluorescence. The results show that *E. coli* cells expressing GFPuv produced an intense fluorescence

under irradiation with UV light. In contrast, the fusion protein Gp1-GFPuv did not fluoresce in the same conditions (Figure 18). The emission of fluorescence could be observed on both agar plates and in a fluorescence microscope. These results suggest that Gp1-GFPuv is not localized in the cytoplasm, meaning that Gp1 removes GFPuv from the cytoplasm strengthening the role of Ms6 Gp1 in protein translocation.

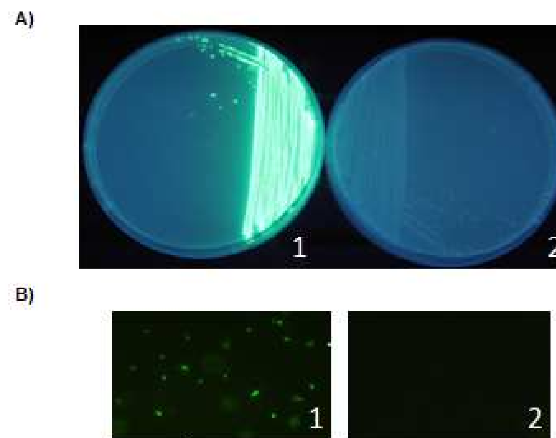


Figure 18. Visualization of 1) GFP and 2) Gp1-GFP in A) LB agar and B) live Images in fluorescence microscopy under UV light.

In order to locate the fusion Gp1-GFP protein in *E. coli*, we proceeded with the separation of the soluble and insoluble fraction of the cellular extracts. The different fractions were analysed by SDS-PAGE followed by Coomassie blue staining (Figure 19).

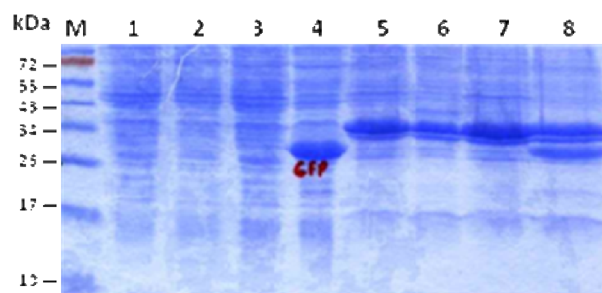


Figure 19. Detection of Gp1-GFP fusion in *E. coli* extracts after SDS-PAGE and Coomassie blue analysis. The lanes 1 to 4 correspond to the soluble fraction (cytoplasmic proteins) and the lanes 5 to 8 correspond to the insoluble fraction (membrane proteins).: lane 1) control, pQE30+IPTG (3h); lane 2) $pAS5_{GP1GFP}$+IPTG (1h), lane 3) $pAS5_{GP1\text{-}GFP}$+IPTG (3h), lane 4) $pAS4_{GFP}$+IPTG (3h), respectively and also

to the soluble fraction (cytoplasmic proteins); The lanes 5 until 8, correspond to the same plasmids recombinants but in other fraction.

The results obtained are not clear. Although we can see GFPuv, with 27 kDa in the soluble fraction, a band with the same molecular mass can also be seen in the insoluble fraction. Although GFPuv is a soluble protein, the presence in the insoluble fraction may be a result of an overexpression causing precipitation of the protein (Figure 19 lanes 4 and 8). Surprisingly we could not identify in the Coomassie staining the band corresponding to the GP1-GFP fusion. Although a band with a similar molecular mass (37 kDa: 10 kDa of GP1 + 27 kDa of GFP) is observed in the insoluble fraction (Figure 19 lanes 6 and 7), we can not identify it as the Gp1-GFP fusion, since a similar band is observed in *E. coli* cells carrying the empty vector pQE30 (Figure 19 lane 5), suggesting that this a cellular protein. However it's noteworthy that the fusion is also not detected in the soluble fraction raising the possibility that Gp1 removes GFP from the cytoplasm. It's clear that further assays must be performed to clarify this hypothesis.

Protease accessibility experiments of LysA in E. coli

As already mentioned Ms6 Gp1 behaves as a chaperone like protein and it was previously suggested that LysA is exported in a holin independent manner. Thus in the presence of Gp1, Ms6LysA should be translocated across the cytoplasmic membrane and exposed in the periplasm whereas without Gp1, LysA should remain in the cytoplasm. Thus accessibility of LysA to externally added proteinase K was used to check its presence in the periplasm compartment. For that we used two pET29 derivative plasmids, pMJC3 and pMJC4 (Catalão *et al.,* 2010), containing *gp1* and *lysA* genes or *lys*A, respectively. *E. coli* cells expressing only $LysAHis_6$ (pMJC4) or both Gp1 and $LysAHis_6$ (pMJC3) were subjected to a Proteinase K treatment and the presence of LysA was checked by western blot with an anti His antibody. The results of this experiment are represented in Figure 14.

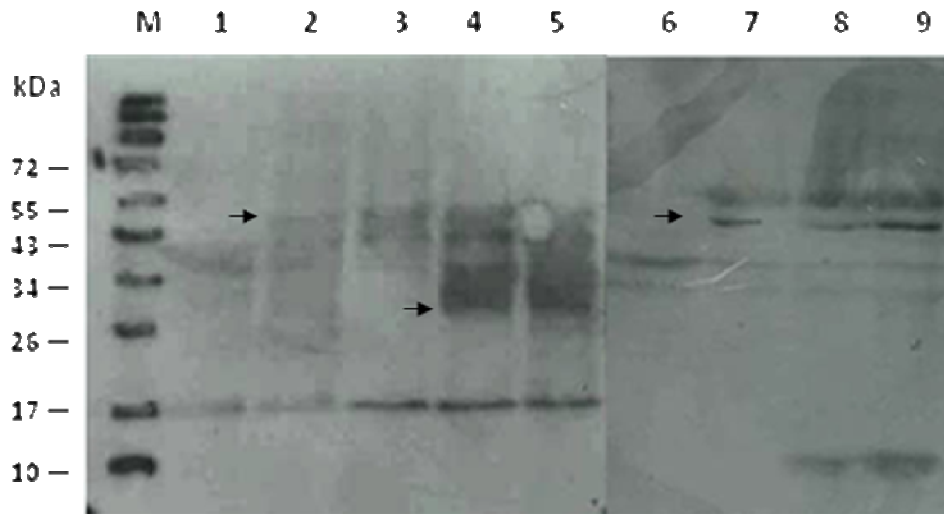


Figure 14. Immunoblot analysis after proteolysis of the periplasmic proteins with proteinase K (PK) in *E.coli.*, containing the recombinant plasmids pMJC3 and pMJC4. Lanes: 1) pMJC3 uninduced 2) pMJC3 induced 1h with 1 mM IPTG 3) pMJC3 induced and treated with lysozyme 4) pMJC3 induced and treated with lysozyme and PK (1 mL of culture) 5) pMJC3 induced and treated with lysozyme and PK (2 mL of culture) 6) pMJC4 uninduced 7) pMJC4 induced 1h with 1 mM IPTG 8) pMJC4 treated with lysozyme 7) pMJC4 treated with lysozyme and PK. M correspond to a prestained protein ladder (PageRuler, Fermentas).

The presence of LysAHis$_6$, with a predicted molecular mass of 43 kDa, can be detected in lanes 2 and 3. In the lanes 4 and 5, although the same band is still present we observe additional bands between 26 kDa and 43 kDa in samples treated with proteinase K. Although we were expecting the disappearance of LysA or a reduction in the detected amount of LysA the appearance of bands with lower molecular mass led us to question if only part of LysA was degraded by proteinase K.

CONCLUSION

At the beginning of the present work, the intention was to contribute to a better understanding of the Ms6 Gp1 in export of the endolysin LysA, investigating whether Gp1 only assists the LysA until reaching the cytoplasmic membrane or whether it is translocated together with LysA to the periplasm. Although in this work it was not possible to answer this question, several achievements were reached.

The first step of this work was to construct a Ms6 derivative mutant by fusing the 3' end of *gp1* with a sequence coding for a C-Myc tag that would allow the detection of the protein with an antibody since no anti Gp1 antibodies were available in the laboratory. To achieve this objective we used the Bacteriophage Recombineering of Electroporated DNA (BRED) system developed by Marineli (2008). Successfully it was possible to obtain an Ms6 mutant that does not affect the phage growth parameters as demonstrated by the production of phage plaque morphologically identical to the wild type and by a very similar one step growth curve. Subsequently, we tried to follow the production of Gp1-CMyc during the infection of *M. smegmatis* with the mutant phage. After several attempts to obtain results for this experiment, we failed to detect the Gp1-CMyc, possibly because it is expressed in a low amount. Several variations to the methodology were tested but still without results. However when we did the fractionation of *M. smegmatis* infected cells with MS6GP1-CMYC we obtained some results. In this experiment we detected that Gp1 was found in both the cytoplasmic and membrane fraction. This is in agreement with the results obtained with *E. coli* cells expressing a Gp1 protein tagged with a hexahistidine at the C-terminus. This is not surprisingly since chaperones are small proteins that to exert their function distribute in different cell compartments and the Gp1 characteristics fit the ones of molecular chaperones (Page *et al.*, 2002; Feldman *et al.,* 2003).The Gp1 protein was found to be both associated with the cytoplasmic membrane and free in the soluble fractions in *M. smegmatis* and in *E. coli*.

With these results in mind, the next step was to determine if Ms6 Gp1 delivers proteins to the extracytoplasmic environment and for that, protein fusions with the periplasmic reporter phosphatase alkaline or the cytoplasmic green fluorescent protein were constructed.

The results obtained showed an alkaline phosphatase activity when Ms6 Gp1 was fused to PhoA without its own signal peptide. This indicates that Gp1 promotes the translocation of PhoA' to the periplasm, since no activity was observed without Gp1. With the Gp1-GFP fusion, although we could not detect the protein in a coomassie blue staining, the absence of a fluorescence under UV light (Figure 18 and B 2) may suggest that GFP is not in the cytoplasm and could thus be translocated by Gp1. It has been shown by Phillips *et al.,* 2001) that GFP only fluoresce when localized in the cytoplasm. Nevertheless this result needs to be confirmed.

Regarding the question: is LysA localized to the periplasm in the presence of Gp1? Since we were unsuccessful with LysA localization assays, we tested the accessibility of Ms6 LysA to proteinase K in *E. coli*. Although these results need further confirmation, they suggest a proteinase K activity on Gp1 LysA tagged with a hexahistidine sequence.

Thus, in conclusion the results obtained in this work show that Gp1 can translocate proteins across the cytoplasmic membrane both in *E. coli* and *M. smegmatis* and suggests that Ms6 LysA is periplasmic localized in *E. coli* in the presence of Gp1.

Taking in consideration what is known about export of endolysins to the extracytoplasmic environment it is clear that the model of Mycobacteriophage is different from the ones described so far and opens new opportunities for new studies in mycobacterial secretion pathways..

REFERENCES

Ackermann, H-W. (1997) Félix d'Hérelle, decouvreur des bacteriophages. Les Selécion de Médicine Sciences, 8: 3-6.

Ackermann, H-W. (2003) Bacteriophage observations and evolution. Res Microbiol 154, 245–251.

Ackermann, H-W. (2009) Phage Classification and Characterization- cap 13. *In*, Bacteriophages: Methods and Protocol*,* Human Press.

Adams, M.H. (1959) Methods of study of bacterial viruses. *In Bacteriophages*. New York: Interscience Publishers, Inc., pp. 443–522.

Arora, Kriti ; Whiteford, Danelle C.; Lau-Bonilla, Dalia; Davitt, Christine M.; Dahl, John L. (2008) Inactivation of *lsr2* Results in a Hypermotile Phenotype in *Mycobacterium smegmatis*_†. Journal of bacteriology, 190: 4291–4300.

Belin, Dominique. (2010) In vivo analysis of protein translocation to the *Escherichia coli* Periplasm. *In*, Protein Secretion, Methods in Molecular Biology 619. A. Economou (ed.).

Bernhardt, Thomas G.; Wang, Ing-Nang; Struck Douglas K.; Young, Ry. (2002) Breaking free: "Protein antibiotics" and phage lysis. Research Microbiol 153:493-501.

Borysowski, Jan; weber-Dabrowska, Beata; Górski, Andrzej. (2006) Bacteriophage Endolysins as a Novel Class of Antibacterial Agents. Exp Biol Med, 231:366–377.

Boyer, Mickaël; Haurat, Jacqueline; Samain, Sylvie; Segurens, Béatrice; Gavory, Frédérick; González, Víctor; Mavingui, Patrick; Rohr, René; Bally, René; Dyé ,Florence Wisniewski. (2008) Bacteriophage Prevalence in the Genus *Azospirillum* and Analysis of the First Genome Sequence of an *Azospirillum brasilense* Integrative Phage. Applied and environmental microbiology, 74: 861–874.

Bradley, David E. (1967) Ultrastructure of Bacteriophages and Bacteriocins. Bacteriol reviews, 31: 230-314.

Brickman, Edith; Beckwith, Jon. (1975) Analysis of the Regulation of Escherichia coli Alkaline Phosphatase Synthesis using Deletions and Ø80 Transducing Phages. J Mol Biol 96, 307-316.

Brüssow, Harald ; Canchaya, Carlos; Hardt, Wolf-Dietrich. (2004) Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. Microbiology and molecular biology reviews, 68: 560–602.

Brüssow, Harald. (2005) Phage therapy: the Escherichia coli experience. Microbiol, 151, 2133–2140.

Canchaya, Carlos; Fournous, Ghislain; Chibani-Chennoufi,Sandra; Dillmann, Marie-Lise; Brüssow, Harald. (2003) Phage as agents of lateral gene transfer. Current Opinion in Microbiology , 6:417–424.

Catalão, Maria João; Gil, Filipa; Pereira, José Moniz; Pimentel, Madalena. (2010) The mycobacteriophage Ms6 encodes a chaperone-like protein involved in the endolysin delivery to the peptidoglycan. Mol Microbiol 77(3), 672-686.

Catalão, M.J.; Milho, C.; Gil, F.; Moniz-Pereira ,J.; Pimentel, M. (2011a) A Second Endolysin Gene Is Fully Embedded In-Frame

with the lysA Gene of Mycobacteriophage Ms6. PLoS ONE 6(6): e20515.

Catalão, MJ; Gil ,F; Moniz-Pereira ,J.;, Pimentel ,M. (2011b) Functional analysis of the holin-like proteins of mycobacteriophage Ms6. J Bacteriol 193: 2793-2803.

Cheng, Q.; Nelson, D.; Zhu, S.; Fichetti, V.A. (2005) Removal of group B streptococci colonizing the vagina and oropharrynx of mice with bacteriophage lytic enzymes. Antimicrob Agents Chemother 49(1):87-110.

Deschavanne, Patrick; DuBow, Michael S; Regeard, Christophe. (2010) The use of genomic signature distance between bacteriophages and their hosts displays evolutionary relationships and phage growth cycle determination. Virology Journal, 7:163.

Feldman, M.F.; Cornelis G.R. (2003) The multitalented type III chaperons: all you can do with 15 kDa. Fems Microbiol Lett 219:151-158.

Feilmeier , Bradley J.; Iseminger, Ginger; Schroeder, Diane; Webber, Hannali; Phillips, Gregory J (2000) Green Fluorescent Protein Functions as a Reporter for Protein Localization in *Escherichia coli.* J Bacteriol, 182: 4068–4076.

Fischetti, Vincent A. (2005) Bacteriophage lytic enzymes: novel anti-infectives. Trends in Microbiology, 13 (10).

Freitas-Vieira, Acilino; Anes, Elsa; Moniz-Pereira José. (1998) The site-specific recombination locus of mycobacteriophage Ms6 determines DNA integration at the $tRNA^{Ala}$ gene of *Mycobacterium* spp. Microbiol , 144, 3397-3406.

Garcia, Miguel: Pimentel, Madalena; Moniz-Pereira, José. (2002) Expression of Mycobacteriophage Ms6 Lysis Genes Is Driven by Two _70-Like Promoters and Is Dependent on a Transcription Termination Signal Present in the Leader RNA. J Bacteriol 184:3034-3043.

Gil, Filipa; Catalão, Maria João; Moniz-Pereira, José; Leandro, Paula; McNeil, Michael; Pimentel, Madalena. (2008) The lytic cassette of mycobacteriophage Ms6 encodes an enzyme with lipolytic activity. Microbiol, 154, 1364–1371.

Gil, Filipa; Grzegorzewicz, Anna E.; Catalão, Maria João; Vital, João; McNeil, Michael R.; Pimentel, Madalena. (2010) Mycobacteriophage Ms6 LysB specifically targets the outer membrane of Mycobacterium smegmatis. Microbiol, 156, 1497–1504.

Guttman Burton; Raya, Raul; Kutter, Elizabeth. (2004) Cap 3: Basic phage biology. *In*, Bacteriophages: biology and applications. Kutter E.; Sulakvelidze A. (Eds.) Boca raton, FL: CRC Press p.29-66.

Hanlon G.W. (2007) Bacteriophages: an appraisal of their role in the treatment of bacterial infections. Int J Antimicrob Agents 30:118-128.

Hanahan D. (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol. 1983 166(4):557-80.

Hatfull, Graham F.; Jacobs, W.R. (1994) Mycobacteriophages: cornerstones of mycobacterial research. *In*, Tubercolosis: pathogenesis, protection and control. Bloom B.R. (Ed.) ASM, Washinton D.C., p.165-183.

Hatfull, Graham F. (2000) Molecular genetics of mycobacteriophages. *In,* Molecular genetics of micobactéria. Hatfull and Jacobs (Eds.) ASM, Washinton D.C., p.3-54.

Hatfull, Graham F. (2006) Mycobacteriophages. *In,* The bacteriophages. Calendar R. (Ed.) 2nd ed. Oxford University Pres, New York, p602-620.

Hatfull, Graham F.; Jacobs-Sera, Deborah; Lawrence, Jeffrey G.; Pope, Welkin H.; Russell, Daniel A.; Ching-Chung Ko, *et al.* (2010) Comparative genomic analysis of sixty mycobacteriophage genomes: Genome clustering, gene acquisition and gene size. *JMol Biol.*, 397(1): 119–143.

Hendrix, R.W (2005) Bacteriophage evolution and the role of phages in host evolution. In, Phages: Their Role in Bacterial Pathogenesis and Biotechonology. Waldor M. K., Fiedman D.I. and S. Adhya (Eds.) ASM Press, Washinton D.C., p.55-65.

Hoffmann, Christian; Leis, Andrew; Niederweis, Michael; Plitzko, Jürgen M.; Engelhardt, Harald. (2008) Disclosure of the mycobacterial outer membrane: Cryo-electron tomography and vitreous sections reveal the lipid bilayer structure . PNAS, 105: 3963–3967.

van Kessel, J. C. and G. F. Hatfull, (2007) Recombineering in Mycobacterium tuberculosis. Nat Methods 4: 147-152.

Kutter, Elizabeth; Sulakvelidze, Alexander. (2004) Bacteriophages- Biology and Applications. CRC PRESS 2000 N.W.

Lagos, R.; Villnueva, J.E., Monasteiro, O. (1999) Identification and proprietis of the genes enconding Microcin E492 and its immunity protein. J Bacteriol 181: 212-217.

Loessner, Martin J. (2005) Bacteriophage endolysins: current state of research and applications. Current Opinion in Microbiology 8:480–487.

Manoil, C., J. J. Mekalanos, and J. Beckwith. (1990). Alkaline phosphatase fusions: sensors of subcellular location. J. Bacteriol. 172:515–518.

Navarre, W.W.; Ton-That, H.; Faull, K.F.; Schneewind, O. (1999) Multiple enzymatic activities of the murein hydrolase from staphylococcal phage Φ11. Identification of D-alanyl-glycine endopeptidase activity. J Biol Chem 274(22):15847-15856.

Pang, Ting; Savva, Christos G.; Fleming, Karen G.; Struck, Douglas K.; Young, Ry. (2009) Structure of the lethal phage pinhole. PNAS, 106: 18966–18971.

Payne, Kimberly; Sun, Qingan; Sacchettini, James; Hatfull, Graham F. (2009) Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. Mol Microbiol, 73(3): 367–381.

Park, Taehyun; Struck, Douglas K. ; Deaton, John F.; Young Ry. (2006) Topological dynamics of holins in programmed bacterial lysis. PNAS, 103: 9713–19718.

Park, Taehyun; Struck, Douglas K.; Dankenbring, Chelsey A.; Young, Ry. (2007) The Pinholin of Lambdoid Phage 21: Control of Lysis by Membrane Depolarization. J Bacteriol, 189: 9135–9139.

Pedulla, Marisa L; Ford Michael E; Houtz, JM; Karthikeyan, Tharun; Wadsworth, Curtis; Lewis, JA; Jacobs-Sera, Debbie; Falbo, Jacob; Gross J, Pannunzio NR, Brucker W, Kumar V, Kandasamy J, Keenan L, Bardarov S, Kriakov J, Lawrence JG, Jacobs WR Jr, Hendrix RW, Hatfull, Graham F. (2003) Origins of Highly Mosaic Mycobacteriophage Genomes. Cell Press, 113: 171–182.

Phillips, G. J. (2001). Green fluorescent protein—a bright idea for the study of bacterial protein localization. FEMS Microbiol. Lett. 204:9–18.

Portugal, Isabel; Anes, Elsa; Pereira, Moniz. (1999) Temperate Mycobacteriophage from *M.smegmatis* ; Acta Leprologica 7 (supplement 1): 243-244.

Prescott, L. M., J. P. Harley, and D. A. Klein. 1993. *Microbiology,* 2nd ed. New York: William C Brown Pub.

Pritchard, D.G.; Dong, S.; Baker, J.R.; Engler, J.A. (2004) The bifunctional peptidoglycan lysin of *Streptococcus agalactiae* bacteriophage B30. Microbiology 150(Pt7):2079-2087.

Sambrook, Joseph; Russelll, David W. (2001) Molecular Cloning- A Laboratory Manual. 3TH edition, Cold Spring Harbor Laboratory Press, vol 3, New York.

São-José C.;Parreira R.;Vieria, Graça; Santos, Mário A. (2000) The N-Terminal Region of the *Oenococcus oeni* Bacteriophage fOg44 Lysin Behaves as a Bona Fide Signal Peptide in *Escherichia coli* and as a *cis*-Inhibitory Element, Preventing Lytic Activity on Oenococcal Cells. Journal of Bacteriology, 182: 5823–5831.

São-José C. (2002) Genome analysis and gene expression in Oenophage fOg44: Evidence for a new strategy of bacteriophage-indunced host lysis. PhD Thesis. Faculdade de Ciências da Univ. Lisboa.

São-José, C.; Parreira, R.; Santos, Mário A. (2003) Triggering of host-cell lysis by double-stranded DNA bacteriophages: fundamental concepts, recent developments and emerging applicatrions. Recent Res Dev Bacteriol 1:103-130.

São-José C, Nascimento J, Parreira R, Santos M (2007) Release of progeny phages from infected cells. In: McGrath S and van Sinderen D, editors. Bacteriophage: genetics and molecular biology. Caister Academic Press. pp. 309-336.

Schaechter, Moselio; Ingraham, John L.; Neidhardt, Frederick C. (2006) Microbe. ASM Press, American Society for Microbiology, Washington, D.C.

Shao, Yongping; Wang, Ing-Nang (2008) Bacteriophage Adsorption Rate and Optimal Lysis Time. Genetics 180: 471–482.

Marinelli, Laura J.; Piuri, Mariana; Swigonová , Zuzana; Balachandran, Amrita; Oldfield, Lauren; Kessel, M. Julia C. van; Hatfull, Graham F. (2008) BRED: A Simple and Powerful Tool for Constructing Mutant and Recombinant Bacteriophage Genomes. PLoS ONE, Volume 3, Issue 12 , e3957.

Short, S. M. ; Suttle, C. A. (1999) Use of the polymerase chain reaction and denaturing gradient gel electrophoresis to study diversity in natural virus communities. *Hydrobiologia* 401: 19–32.

Snapper, S. B., Melton, R. E:, Mustafa, S., Kieser, T., Jacobs Jr., W. R. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*, Mol. Microbiol. 4: 1911– 1919.

Summer, Elizabeth J.; Berry, Joel; Tran, Tram Anh T.; Struck, Lili Niu Douglas K.; Young, Ry. (2007) Rz/Rz1 Lysis Gene Equivalents in Phages of Gram-negative Hosts. J. Mol. Biol., 373, 1098–1112.

Sun, Qingan; Kuty, Gabriel F.; Arockiasamy, Arulandu; Xu, Min; Young, Ry; Sacchettini , James C. (2009) Regulation of a muralytic enzyme by dynamic membrane topology. Nature Structural & Molecular Biology; 16(11): 1192-1194.

Wang, Ing-Nang, Smith, David L.;Young, Ry. (2000) Holins: The Protein Clocks of Bacteriophage Infections. Annu Rev Microbiol 54:799–825.

Wang, Ing-Nang; Deaton, John; Young, Ry. (2003) Sizing the Holin Lesion with an Endolysin–β-Galactosidase Fusion. Journal of Bacteriology, 185: 779–787.

Wang, Ing-Nang. (2006) Lysis Timing and Bacteriophage Fitness. Genetics 172(1): 17–26.

Weinbauer, Markus G. (2004) Ecology of prokaryotic viruses. Federation of European Microbiological Societies. FEMS Microbiology Reviews 28 (2004) 127–181.

Xu, Min; Arulandu, Arockiasamy; Struck, Douglas K.; Swanson, Stephanie; Sacchettini,. James C.; Young ,Ry. (2005) Disulfide Isomerization After Membrane Release of Its SAR Domain Activates P1 Lysozyme. Science Vol 307.

Xu, Min; Struck, Douglas K.; Deaton, John; Wang, Ing-Nang; Young, Ry. (2004) A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. PNAS, 101(17): 6415–6420.

Young, Ry. (1992) Bacteriophage Lysis: Mechanism and Regulation. Microbiological Reviews, 56: 430-481.

Young, R., Wang I. N. e Roof, W. D. (2000) Phages will out: strategies of host cell lysis. Trends in Microbiology 8: 120-128.

Young, Ry. (2002) Bacteriophage Holins: Deadly Diversity. J. Mol. Microbiol. Biotechnol, 4: 21-36.

Young, Ry. (2005) Phage lysis. *In*, Phages: their role in bacterial pathogenesis and biotechnologie. Waldor M.K., Friedman D. I. and Adhya S. (Eds.) AMS Press, Washington D.C., p. 92-127.

Young, R.; Wang I.N. (2006) Phage lysis. *In*, The bacteriophages. Calendar R. 2nd edition, Oxford University Press, New York, p. 104-125.

Zheng, Yi; Struck Douglas K.; Dankenbring Chelsey A.; Young, Ry. (2008) Evolutionary dominance of holin lysis systems derives from superior genetic malleability. Microbiology, 154, 1710–1718.

Zuber, Benoît ; Chami, Mohamed; Houssin, Christine; Dubochet, Jacques; Griffiths, Gareth; Daffé, Mamadou. (2008) Direct Visualization of the Outer Membrane of Mycobacteria and Corynebacteria in Their Native State_†. Journal of bacteriology, 190: 5672–5680.