UNIVERSIDADE DE LISBOA
FACULDADE DE FARMÁCIA

CHARACTERIZATION AND FUNCTIONAL ANALYSIS OF THE Gp1 ACCESSORY LYSIS PROTEIN IN *MYCOBACTERIUM SMEGMATIS* INFECTION BY THE MYCOBACTERIOPHAGE Ms6

Maria João Gracias Fernandes da Costa Catalão

DOUTORAMENTO EM FARMÁCIA
MICROBIOLOGIA
2010

UNIVERSIDADE DE LISBOA
FACULDADE DE FARMÁCIA

CHARACTERIZATION AND FUNCTIONAL ANALYSIS OF THE Gp1 ACCESSORY LYSIS PROTEIN IN *MYCOBACTERIUM SMEGMATIS* INFECTION BY THE MYCOBACTERIOPHAGE Ms6

Maria João Gracias Fernandes da Costa Catalão

Research Advisors:
Madalena M. V. Pimentel, Ph.D.
José A. F. Moniz Pereira, Ph.D.

DOUTORAMENTO EM FARMÁCIA
MICROBIOLOGIA
2010

CHARACTERIZATION AND FUNCTIONAL ANALYSIS OF THE Gp1 ACCESSORY LYSIS PROTEIN IN *MYCOBACTERIUM SMEGMATIS* INFECTION BY THE MYCOBACTERIOPHAGE Ms6

CARACTERIZAÇÃO E ANÁLISE FUNCIONAL DA PROTEÍNA ACESSÓRIA DE LISE, Gp1, NA INFECÇÃO DE *MYCOBACTERIUM SMEGMATIS* PELO MICOBACTERIÓFAGO Ms6

Dissertação apresentada à Faculdade de Farmácia da Universidade de Lisboa para obtenção do grau de Doutor em Farmácia (Microbiologia)

Maria João Gracias Fernandes da Costa Catalão

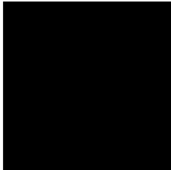
2010

The studies presented in this thesis were performed at Centro de Patogénese Molecular, Faculdade de Farmácia da Universidade de Lisboa under the supervision of Professor Madalena Pimentel and Professor José Moniz Pereira.

Maria João Gracias Fernandes da Costa Catalão was the recipient of a Ph.D. fellowship (SFRH/BD/24452/2005) from Fundação para a Ciência e Tecnologia (FCT), Lisbon, Portugal. This work was supported by grant PTDC/SAU-FCF/73017/2006 (to M. P.) from FCT, Portugal.

De acordo com o disposto no ponto 1 do artigo nº 40 do Regulamento de Estudos Pós-Graduados da Universidade de Lisboa, deliberação nº 961/2003, publicado em Diário da República – II Série nº 153 – 5 de Julho de 2003, a Autora desta dissertação declara que participou na concepção e execução do trabalho experimental, interpretação dos resultados obtidos e redacção dos manuscritos.

Aos meus Pais



Contents

Preface		ix
Acknowledgments		xiii
Summary		xv
Sumário		xvii
Abbreviations		xxiii
Chapter 1:	General Introduction	1
	Objectives	65
Chapter 2:	The Mycobacteriophage Ms6 Encodes a Chaperone-Like Protein Involved in the Endolysin Delivery to the Peptidoglycan	67
Chapter 3:	Functional Analysis of the Holin-Like Proteins of Mycobacteriophage Ms6	107
Chapter 4:	Sequence Analysis, Characterization and Lytic Activity of Lysin A, the Cell Wall Hydrolase Encoded by the Mycobacteriophage Ms6	143
Chapter 5:	Concluding Remarks	189

Preface

All the world´s a phage
c. William Shakespeare

March 24th, 1882. Robert Koch announced the discovery of the etiologic agent causative of tuberculosis, *Mycobacterium tuberculosis*. For this, was awarded the Nobel Prize in 1905. Although more than 100 years have passed, and despite the efforts of the scientific community, the introduction of many effective anti-tuberculosis agents and the use of a live attenuated vaccine (BCG), tuberculosis remains the leading infectious cause of death worldwide killing over 2 million people every year. In addition, there has been an emergence of multidrug-resistant tuberculosis (MDR-TB) cases, the ultimate result of effective drugs ineffectively administered which in turns contributes to further spread of the disease. Containment of this growing problem is one of the great challenges of the humankind, and, unless poverty and unequal development are ended throughout the world, tuberculosis will continue to be an ineradicable public health problem.

The resurgence of tuberculosis, the difficulty of treating mycobacterial disease and the emergence of drug-resistant strains of *M. tuberculosis* has stimulated the renewed interest into mycobacteriophage research in order to investigate whether they could provide a complementary mean of therapy. Mycobacteriophages are dsDNA viruses that infect mycobacteria. Following phage replication, lysis of the cell wall occurs, eventually destroying the host bacterium and releasing the progeny phage. The first mycobacteriophage was isolated in

1947, and since that time over 250 of these viruses have been identified. Mycobacteriophages have made a significant contribution to the knowledge of mycobacteria over the past sixty years, and following the development of typing techniques in the 1960s and 1970s they were widely used in epidemiological studies of tuberculosis. During the past two decades, mycobacteriophages have become important in molecular studies of mycobacteria, both in terms of studying phage biology and as tools in recombinant DNA technology, thus facilitating the investigation of mycobacterial pathogenesis. At present, their potential as diagnostics tools is also being realised with the development of exciting new techniques for rapid bacterial detection, drug susceptibility testing and mycobacterial recombineering, with the last one being an important contribution for the simple construction of potential vaccine strains of *M. tuberculosis* and the identification of its virulence determinants.

The ability of mycobacteriophages to lyse and kill mycobacteria also led to investigations of their potential as therapeutic agents against tuberculosis. Unfortunately however, early experiments on using lytic phages therapeutically during tuberculosis experimentally-infected animals not only failed to elicit cure, but in guinea pigs, had an adverse effect on their survival. Phage therapy is currently used in Eastern Europe and countries of the former Soviet Union and recent studies have confirmed that phages can be highly effective in treating many different types of bacterial infections, such as skin wounds, burns and ears infections or severe diarrhoea. However, treatment of systemic infections is the most challenging environment for bacteriophage action, with issues of compartmentation, host defense and kinetics, which are not present to the same degree in topical applications, creating real or imaginary barriers. In addition, killing intracellular pathogens such as *M. tuberculosis* presents a tough challenge as, in order to infect the target bacilli, the phages need to transverse the mammalian cell membrane, survive in the macrophage adverse intra-cellular environment

(acidification and presence of reactive oxygen species-ROS) or penetrate the granuloma.

When I started the Ph.D. program in the laboratory of Professor Madalena Pimentel and Professor Moniz Pereira, I continued the work initiated by former researchers of the lab and focused on investigating the role of a small accessory lysis protein restricted to mycobacteriophages. My first questions as a student who has just entered a new area of research, gave rise to additional new questions and to the obtainment of unexpected and exciting findings. The present work provides insight into the understanding of the molecular pathways involved in mycobacteriophage lysis, and more importantly, uncovers intriguing connections and links that warrant further investigations.

From the beginning, the purpose of my Ph.D. work, now presented and discussed in this thesis was to identify and characterize the role of the mycobacteriophage Ms6 Gp1 lysis protein during mycobacteria infection in view of the fact that, in addition to the endolysin, it is necessary for *E. coli* lysis without the requirement of holin activity to disrupt the cytoplasmic membrane. The functional analysis of the Ms6 holin-like proteins during mycobacterial infection came as a natural extension of this work. Chapter 1 provides a general, up-to-date review on the impact and the interest on mycobacteriophage research. In addition, preliminary data on the application of mycobacteriophages and the potential use of their lytic proteins in therapeutics is also discussed. We also focus on describing the mechanisms involved in bacteriophage-induced host cell lysis and the recent progress made in the field with the discovery of novel lysis models that include secretory endolysins endowed with signal peptides or signal-arrest-release (SAR) sequences and a new class of holins encoded by phages with SAR endolysins, pinholins. In Chapter 2, we provide evidence that Gp1 specifically interacts with the N-terminal region of the Ms6 endolysin and is involved in its translocation

across the cytoplasmic membrane in a holin-independent manner. In addition, construction of a mutant Ms6 phage defective for Gp1 synthesis showed that even though Gp1 is not vital for lysis is required for an efficient lysis of mycobacteria and a productive burst size. In Chapter 3, the role of the Ms6 holin-like proteins was further analysed, using a series of Ms6 mutant phages deleted for different lysis genes. Since Ms6 holins are not essential for mycobacterial lysis, their function in the Ms6 lysis model was discussed. Chapter 4 describes a remarkably finding observed during *M. smegmatis* infection by Ms6. We found that during infection, two endolysin forms are synthesized and both are required for the normal timing, progression and completion of host cell lysis. The implications of this new discovery are further examined. Finally, in Chapter 5 we integrate our overall findings, propose a general model for Ms6 lysis of mycobacteria and discuss specific future perspectives.

Even though the exact mechanism involved in Ms6 endolysin translocation in mycobacteria is still elusive, with this thesis we hope to elucidate and contribute to a better understanding of the mechanisms of mycobacteriophage lysis of mycobacteria. Ultimately, an increased knowledge of the exact pathways by which the mycobacteriophage lysis effectors are positioned next to their targets may result in the identification of mycobacteria new cell wall targets and in the promising therapeutic use of mycobacteriophage lysins.

Acknowledgments

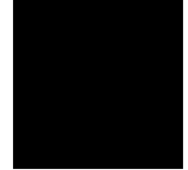
As minhas primeiras palavras de agradecimento são dirigidas aos meus orientadores: ao Professor Moniz Pereira pelo enorme entusiasmo com que me recebeu, ainda aluna do 3º ano da licenciatura no seu grupo de investigação, por ter despertado em mim o interesse pela Microbiologia e pela Ciência e por tudo o que me ensinou. Obrigada pelas palavras de ânimo perante as minhas dificuldades e pela compreensão com que me permitiu chegar até aqui e me incentivou sempre a fazer mais e melhor. À Professora Madalena que me ensinou os primeiros passos na investigação, por ter confiado nas minhas capacidades e ter estado sempre disponível para me ajudar quando precisei. Agradeço ainda terem proporcionado as condições materiais necessárias para a execução do trabalho experimental desta tese.

À Dra. Paula Resende e ao Dr. João Vital, pela preocupação constante, por tudo o que me ensinaram e por todo o carinho e amizade que sempre me deram. Aos Professores do sub-grupo de Microbiologia, Prof. Helena Lourenço, Prof. Graciete, Prof. Mané, Prof. Isabel, Prof. Zé Miguel e Prof. João Gonçalves por terem contribuído, com tanto que me ensinaram, para o meu interesse pela Microbiologia. Às Professoras Elsa Anes e Paula Leandro por terem estado sempre disponíveis para me ajudar.

À Filipa, companheira de laboratório e grande amiga, por tudo o que vivemos juntas nos últimos cinco anos, por partilhares comigo as alegrias e as tristezas, por todo o teu apoio e disponibilidade e por te preocupares comigo como

ninguém! À minha querida amiga Cátia, companheira de aventuras e de desventuras, por teres sido o meu porto-de-abrigo nos dias de tempestade...Sem o teu apoio e amizade, tudo teria sido ainda mais difícil! Aos meus amigos que acompanharam todo o meu doutoramento: Zézinha, Inês, Miguel, Jo, Ritinha, Magda, Pedro, Sandra e Helena. Obrigada por me ouvirem, pela amizade e carinho. À Inês Vinga, Francisca Monteiro, Catarina Pinho e Catarina Milho por toda a ajuda e apoio, que em algum momento destes últimos quatro anos foram muito importantes. À Mariana e ao Fred pelos ensinamentos preciosos no início da minha aprendizagem na investigação. Ao Acilino, pela companhia e boa-disposição, e pelas conversas filosóficas. À Sylvie, pela companhia fim-de-semana fora no CPM, pela paciência com que me ouviste, por confiares em mim, pelo optimismo e energia contagiantes, por te teres tornado uma grande amiga! A todos os colegas do subgrupo de Microbiologia e do subgrupo de Bioquímica pela disponibilidade, ajuda e boa-disposição.

Aos meus queridos pais pelo apoio incondicional; ao estarem sempre presentes no meu pensamento, deram-me a força necessária para terminar esta tese. Saber que posso contar sempre com o vosso carinho e compreensão ajudou-me e ajudar-me-á sempre a superar todos os obstáculos e os momentos menos bons. Aos meus irmãos Patrícia e Vasco por todo o carinho e preocupação com as minhas digressões fora de horas ao CPM. Ao Francisco e ao recém-chegado Vasquinho, por me fazerem sorrir sempre!!!



Summary

The majority of phages described to date (≈96%) is endowed with a tail and presents a double-stranded DNA (dsDNA) genome. Each infection cycle terminates with the strictly programmed and regulated lysis of the host brought about by two phage-encoded proteins, a murein-degrading enzyme, the endolysin, which is essential in achieving rapid hydrolysis of cell wall peptidoglycan, and a second membrane-embedded protein, the holin, which serves to release or activate the endolysin at a precisely defined time, thus bringing about an effective burst of the infected host. Targeting of endolysins to peptidoglycan can be achieved in two ways: through pore formation by canonical holins at a defined time or by continuous export of endolysins endowed with export signals during synthesis, assisted by the host Sec translocon. Mycobacteriophage Ms6 is a temperate bacteriophage that infects *Mycobacterium smegmatis* and possesses an unusual lytic cassette: in addition to the endolysin-holin lysis system encoded by genes *lysA* (*gp2*) and *hol* (*gp4*) respectively, the Ms6 lytic cassette comprises three accessory lysis proteins encoded by genes *gp1*, *gp3* (*lysB*) and *gp5* that are restricted to mycobacteriophages. Mycobacteria are Gram-positive bacteria that have evolved a complex cell wall, comprising a peptidoglycan-arabinogalactan polymer with covalently bound mycolic acids of considerable size, a variety of extractable lipids, and pore-forming proteins which provide an extraordinary efficient permeability barrier to noxious compounds and contribute to the high intrinsic resistance of mycobacteria to many drugs. To overcome the disadvantage that a complex cell

wall may represents for a successful infective cycle, mycobacteriophages have evolved new lysis strategies by acquiring, through their evolution, genes that likely confer a substantial selective advantage over those without them by providing faster and more complete lysis.

The present work underscores a new model for endolysin export in mycobacteriophage Ms6. Several lines of evidence indicate that *gp1* encodes a secretion chaperone-like protein that binds the endolysin, assists the export to the extracytoplasmic environment independently of holin function and is required to accomplish an efficient lysis of *M. smegmatis*. Construction of different Ms6 derivatives deleted in different regions of the lysis operon demonstrated that the gene products of *hol* and *gp5*, although nonessential for phage viability, appear to play a role in controlling the timing of lysis. Remarkably, during *M. smegmatis* infection two endolysin forms ($Lysin_{384}$ and $Lysin_{241}$) are synthesized and both enzymes were shown to be essential for the normal timing, progression and completion of host cell lysis. In conclusion, this work highlights the role of the accessory lysis protein Gp1 and revealed alternative pathways for mycobacteria lysis, demonstrating that the presence of the mycobacterium-specific lysis factor Gp1, may confer a selective advantage not only for fitness under different environmental conditions but also as an alternative to lysis exclusively holin-dependent.

Keywords: Mycobacteriophages – Molecular chaperones – Holins – Endolysins - Lysis timing - Phage therapy - Mycobacteria

Sumário

Mycobacterium tuberculosis, o agente etiológico da tuberculose, permanece como uma das principais causas de morbilidade e mortalidade no mundo. Esta realidade alarmante despertou um interesse renovado nas possibilidades terapêuticas dos bacteriófagos (vírus que infectam bactérias) e das sua proteínas líticas devido à capacidade de libertarem a progenia viral através da lise da célula hospedeira. A maioria dos fagos descritos até à data (≈96%) apresenta cauda e um genoma de DNA de cadeia dupla. Para induzirem uma lise efectiva do hospedeiro, estes bacteriófagos utilizam a estratégia "holina-endolisina" e sintetizam uma enzima, a endolisina, essencial para a hidrólise rápida do peptidoglicano e uma segunda proteína membranar, a holina, que permite a libertação ou activação da endolisina num período de tempo bem definido.

O bacteriófago Ms6 é um micobacteriófago temperado que infecta *Mycobacterium smegmatis*. Como acontece com todos os fagos de DNA de cadeia dupla, o seu operão lítico codifica para os dois genes essenciais para a lise do hospedeiro: a endolisina com actividade de amidase codificada pelo gene *lysA* (*gp2*) e a holina codificada pelo gene *hol* (*gp4*). Estes dois genes fazem parte de uma unidade de transcrição juntamente com os genes *gp1*, *gp3* (*lysB*) e *gp5*. Estudos prévios demonstraram que a expressão dos dois primeiros genes do operão lítico (*gp1* e *lysA*) na ausência da holina, é suficiente para induzir a lise de *E. coli*, um fenótipo não observado na ausência de *gp1* ou com a expressão da endolisina apenas. *gp1* é o primeiro gene do operão de lise do micobacteriófago Ms6 e

encontra-se localizado imediatamente a montante da endolisina. Codifica para uma pequena proteína de 77 aminoácidos que apresenta similaridade apenas com outras proteínas de micobacteriófagos, com função desconhecida. Apesar da inibição do crescimento de *E. coli* quando a proteína Gp1 é sobreexpressa, a estrutura secundária da sua sequência de aminoácidos não apresenta as características estruturais das holinas. A grande maioria das endolisinas descritas não possui uma sequência sinal para a secreção e dependem inteiramente das holinas para serem libertadas para o peptidoglicano. No entanto, estudos realizados revelaram a presença de um péptido sinal em N-terminal nas endolisinas de alguns fagos que infectam bactérias Gram-positivas como por exemplo, os bacteriófagos fOg44 de *Oenococcus oeni* e ϕg1e de *Lactobacillus plantarum*. Para atingirem o peptidoglicano, as endolisinas destes fagos requerem a actividade do sistema de secreção Sec de *E. coli* para a sua exportação, e não a formação de lesões membranares pelas holinas. Particularmente interessante é o caso das endolisinas do fago P1 e 21 de *E. coli* que possuem uma sequência SAR (Signal-Arrest-Release Sequence) em N-terminal que lhes permite serem transportadas pelo sistema Sec até ao peptidoglicano. A análise bioinformática da sequência de aminoácidos de LysA não previu uma sequência sinal para a secreção não estando igualmente identificadas sequências SAR em endolisinas de outros micobateriófagos, o que parece indicar que a translocação desta proteína através da membrana citoplasmática está dependente dos outros genes de lise.

No presente estudo, investigou-se o envolvimento da proteína acessória de lise Gp1 no transporte da endolisina através da membrana citoplasmática e a função das holinas Hol e Gp5 na regulação do tempo de lise em *M. smegmatis*. Foi também estudada a função da endolisina durante a infecção das micobactérias e o seu espectro de actividade lítica. Na primeira parte do trabalho, os resultados obtidos demonstraram que em *E. coli* a co-expressão das proteínas Gp1 e LysA (endolisina) é suficiente para obter um efeito lítico na ausência da holina. No

entanto, Gp1 não pertence à classe das holinas, apresentando características estruturais e bioquímicas de chaperone molecular, como um baixo peso molecular (<15 kDa), um ponto isoeléctrico acídico e a capacidade de formação de dímeros. Experiências de interação de proteínas por crosslinking em *E. coli*, demonstraram que Gp1 interage com a endolisina LysA através de ligação com os primeiros 60 aminoácidos N-terminais da enzima, sendo esta região necessária e suficiente para a interacção. De facto, durante a infecção de *M. smegmatis* pelo Ms6 esta interacção também ocorre e a oligomerização de Gp1 parece ser necessária para este processo. A delecção de *gp1* do operão de lise do bacteriófago Ms6 demonstrou que, embora Gp1 não seja essencial para a obtenção de um fenótipo de lise, é necessária para uma lise eficiente de *M. smegmatis* e para a libertação da progenia viral.

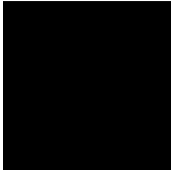
Tendo em consideração que tanto em *E. coli* como em *M. smegmatis*, a holina não é necessária para a lise, a sua função durante o ciclo de infecção foi estudada. A análise da sequência aminoacídica de Hol revelou a presença na região N-terminal da proteína de uma sequência SAR seguida de uma região transmembranar, características das pinholinas. Investigou-se ainda a função da proteína Gp5, que embora possuindo também características estruturais de holina, não é essencial para a lise de *M. smegmatis*. Utilizando a técnica de recombineering, construiram-se diferentes fagos derivados do Ms6, deleccionados em diferentes genes líticos e estudou-se o efeito dessas delecções no seu ciclo de replicação após a infecção dos hospedeiro. Observou-se que Hol e Gp5, embora não sejam essenciais para a viabilidade do bacteriófago, parecem ter um papel na regulação do tempo de lise: a delecção de *hol* causou uma lise precoce enquanto que a delecção de *gp5* retardou o tempo de lise. Por outro lado observou-se também que a presença de ambas as proteínas durante a infecção é essencial para obter a lise do hospedeiro no tempo de lise programado.

Por fim, na terceira parte do trabalho, estudou-se a função durante o ciclo lítico das duas endolisinas (Lisina$_{384}$ e Lisina$_{241}$) codificadas pelo gene *lysA*. A análise da sequência nucleotídica de *lysA* revelou a presença de um gene de lise adicional codificado na mesma fase de leitura, e precedido por sinais de tradução (um codão de iniciação GTG e um local de ligação ao ribossoma). Embora a endolisina se tenha revelado essencial para a lise de *M. smegmatis* (um derivado do Ms6 deleccionado no gene *lysA* não é viável), bacteriófagos derivados do Ms6 defectivos na síntese da lisina$_{384}$ ou da lisina$_{241}$ embora sejam viáveis, apresentam um ciclo de infecção defectivo. A proteína Gp1 demonstrou ser essencial para a síntese ou estabilidade da lisina$_{384}$ mas não da lisina$_{241}$, embora não pareça participar na função final desta proteína, nem na sua activação ou conformação. Ambas as endolisinas possuem actividade lítica em diferentes bactérias Gram-positivas, Gram-negativas e micobactérias quando adicionadas exogenamente.

Em resumo, o presente trabalho demonstra a importância da proteína acessória de lise Gp1, não só na translocação da lisina$_{384}$ através da membrana citoplasmática, mas também para uma lise eficiente de *M. smegmatis*. Por outro lado, estes estudos revelaram a existência, no micobacteriófago Ms6, de vias alternativas de lise através da presença, no seu genoma, de proteínas de lise adicionais (para além do sistema holina-endolisina) e exclusivas deste grupo de fagos. Devido à complexidade da parede celular micobacteriana e da barreira que poderá representar para um ciclo de infecção bem sucedido, estas proteínas de lise acessórias poderão conferir uma vantagem selectiva, não só em termos evolutivos durante a infecção em diferentes condições ambientais, mas também como uma alternativa viável para uma lise exclusivamente dependente da holina.

A caracterização dos mecanismos de lise das micobactérias pelos micobacteriófagos e um conhecimento das vias exactas pelas quais as enzimas efectoras da lise são posicionadas junto do seu substracto, poderá contribuir para a identificação de novos alvos terapêuticos na parede das micobactérias.

Palavras-chave: Micobacteriófagos – Tempo de lise – Chaperones moleculares – Holinas – Endolisinas – Terapia Fágica - Micobactérias



Abbreviations

µg	microgram
µl	microliter
µm	micrometer
Å	angstrom
aa	amino acid
AG	arabinogalactan
Amp	ampicillin
ATCC	American Type Culture Collection
ATP	adenosine-5'-triphosphate
β-gal	beta-galactosidase
BCG	bacille Calmette-Guérin
BCIP	5-bromo-4-chloro-3-indoxyl phosphate
BLAST	basic local alignment search tool
bp	base pair
BRED	bacteriophage recombineering of electroporated DNA
BS	burst size
BS^3	bis(sulfosuccinimidyl) suberate
BSA	bovine serum albumin
cm	centimeter
CM	cytoplasmic membrane
Cryo-EM	cryo-electron microscopy
Cys	cysteine
CWBD	cell wall binding domain
DA	dalton
DADA-PCR	deletion amplification detection assay-PCR
D-Ala	D-alanine
D-Glu	D-glutamate
DNA	deoxyribonucleic acid
ds	double-stranded
Fig	figure

FP	flanking primer
GlcNac	*N*-acetil glucosamine
Gln	glutamine
Gly	glycine
gp1	mycobacteriophage Ms6 Gp1 gene
gp5	mycobacteriophage Ms6 Gp5 gene
h	hour
His	histidine
hol	mycobacteriophage Ms6 holin gene
HRP	horse radish peroxidase
ICTV	International Committee on Taxonomy of Viruses
IM	inner membrane
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
kb	kilobase
kbp	kilobase pair
kDa	kilodalton
lacZ	β-galactosidase gene
LB	Luria-Bertani broth
LIN	lysis inhibition
Lpp	lipoprotein
LPS	lipopolysaccharide
Lys	lysine
lysA	mycobacteriophage Ms6 endolysin gene
lysB	mycobacteriophage Ms6 mycolylarabinogalactan esterase gene
MA	mycolic acids
m-DAP	meso-diaminopimelic acid
MDR-TB	multidrug-resistant tuberculosis
Met	methionine
mg	miligram
min	minute
ml	mililiter
mm	milimeter
mM	milimolar
m.o.i.	multiplicity of infection
MraY	phospho-*N*-acetylmuramoyl-pentapeptide transferase
mRNA	messenger ribonucleic acid
MurA	UDP-*N*-acetylglucosamine-enolpyruvyl transferase
MurNac	*N*-acetylmuramic
Ni-NTA	nickel-nitrilotriacetic acid
nm	nanometer

OADC	oleic acid albumin dextrose complex
OD	optical density
OM	outer membrane
ONPG	ortho-nitrophenyl-β-D-galactopyranoside
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
p.f.u.	plaque forming units
PG	peptidoglycan
PGRP	peptidoglycan recognition protein
phoA	alkaline phosphatase
pI	isoelectric point
PIM	fosfatidilmio-inositol manoside
p.m.f.	proton-motive force
PPIase	peptidyl-prolyl cis-trans-isomerase
PVDF	polyvinylidene difluoride
RBS	ribosome-binding site
RNA	ribonucleic acid
ROS	reactive oxygen species
SAR	signal-arrest-release
SD	Shine-Dalgarno
SDS	sodium dodecyl sulphate
Ser	serine
SIV	simian immunodeficiency virus
SP	signal peptide
ss	single-stranded
TB	tuberculosis
TDM	trehalose diester of mycolic acids
TMD	transmembrane domain
Tn	transposon
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
ts	temperature-sensitive
TTS	type III secretion system
Tyr	tyrosine
wt	wild type
Zn	zinc

1

General Introduction

1. The Bacteriophages: Biology and Interactions with Host Bacteria

Bacteriophages (phages) are bacterial viruses that play major roles in the ecological balance of microbial life and are the most abundant entities in the biosphere (Brüssow and Kutter, 2005). Over the course of evolution, bacteriophages have developed unique proteins that bind to and inactivate (or redirect), critical cellular proteins in bacteria, shutting off key metabolic processes to divert host metabolism to the production of progeny phages (Liu *et al.*, 2004). The vast number on earth, estimated at 10^{31} (Brüssow and Kutter, 2005), has been appreciated only recently, but it is increasingly clear that phages exert enormous influence over the microbial world (Pedulla *et al.*, 2003). The overall diversity of this population appears to be great: no genomically defined phage has been isolated more than once, and the relatively few sequenced phage genomes are highly varied and characterized by a high degree of mosaicism that likely arises from extensive horizontal genetic exchange (Hendrix *et al.*, 1999; Hendrix, 2003; Hatfull, 2008a). The phage population is dynamic, turning over rapidly through constant attrition and subsequent amplification in permissive hosts (Pedulla *et al.*, 2003). On a global scale, it is estimated that ~ 10^{25} phages initiate an infection every second (Pedulla *et al.*, 2003; Hendrix, 2005) and in each of those infections the phage encounters DNA - of bacterial or prophage origin - with which it can potentially recombine to generate new genomic arrangements (Canchaya *et al.*, 2003, Hatfull *et al.*, 2006). The great majority of these phages (≈ 96%) are endowed with a tail and all of them present a double-stranded DNA (dsDNA) genome (Ackermann, 2003). There are a variety of different morphological types of bacteriophages and taxonomy is based on their shape and size as well as on their nucleic acid. The International Committee on Taxonomy of Viruses (ICTV) presently recognizes one order (*Caudovirales*), 13 families, and 31 genera of phages (Ackermann, 2006). Virions have binary, cubic, or helical symmetry, or are pleomorphic. A few types have a lipid-containing envelope or contain lipids as part of the particle wall (Ackermann,

2005; Ackermann, 2006). The head (or capsid) is a protein shell often in the shape of an icosahedron that contains the viral genome that usually comprises dsDNA, but there are small phage groups with single-stranded (ss) DNA, ssRNA, or double-stranded (ds) RNA genomes. The tail may or may not be a contractile structure, usually connected to six tail fibres containing receptors at their tips that recognize attachment sites on the bacterial cell surface. Not all phages possess tails or tail fibres and here other attachment mechanisms are in place (Hanlon, 2007). Most of the bacteriophages of relevance belong to three families, the *Siphoviridae*, *Myoviridae* and the *Podoviridae*, comprising 15 genera. The remaining phages occupy 10 families, each with a small number of members (Ackermann, 2006).

Upon infection of the bacterial host, phages can follow several life cycles: lytic, lysogenic, pseudolysogenic and chronic infections (Weinbauer, 2004). The life cycle of a temperate bacteriophage is shown in Figure 1. The virus encounters its bacterial host during random motion and attaches via specific receptors sites that may be any of a wide variety of cell surface components, including proteins, oligosaccharides, teichoic acids, peptidoglycan and lipopolysaccharides (Guttman *et al.*, 2005; Hanlon, 2007). In some cases, the attachment sites might be present on the cell capsule, flagella or even conjugative pili. Initially the attachment is reversible but then becomes irreversible and is followed by transfer of phage genetic material into the host cell. Injection of the phage genome into the bacterial cell can occur by a variety of mechanisms, depending on the morphology of the virus, but can involve contraction of the tail and formation of a hole within the bacterial cell wall (Weinbauer, 2004; Hanlon, 2007). Many of the bases present on the phage DNA are chemically modified to confer protection against attack by cellular restriction and nuclease enzymes. The viral genome is generally transcribed by host cell RNA polymerase, producing early mRNA that has the effect of taking over the metabolic machinery of the bacterium, redirecting its metabolic processes to the manufacture of new virus components assembled into complete virions.

Following construction and assembly of new phage particles within the host cell, lysis of the bacterium generally occurs, with the release of the new phage progeny-lytic cycle (Guttman *et al.*, 2005; Hanlon, 2007). In the lysogenic cycle, temperate phages do not automatically enter a lytic cycle but instead integrate their DNA into the host cell DNA after infection (Fig. 1). The bacterial cells are then termed lysogenic. When the bacterial DNA replicates, the phage DNA replicates at the same time and each daughter cell will contain the viral DNA (prophage). The prophage directs the synthesis of a repressor protein that blocks the transcription of its own genes and also those of closely related bacteriophages (Weinbauer, 2004).

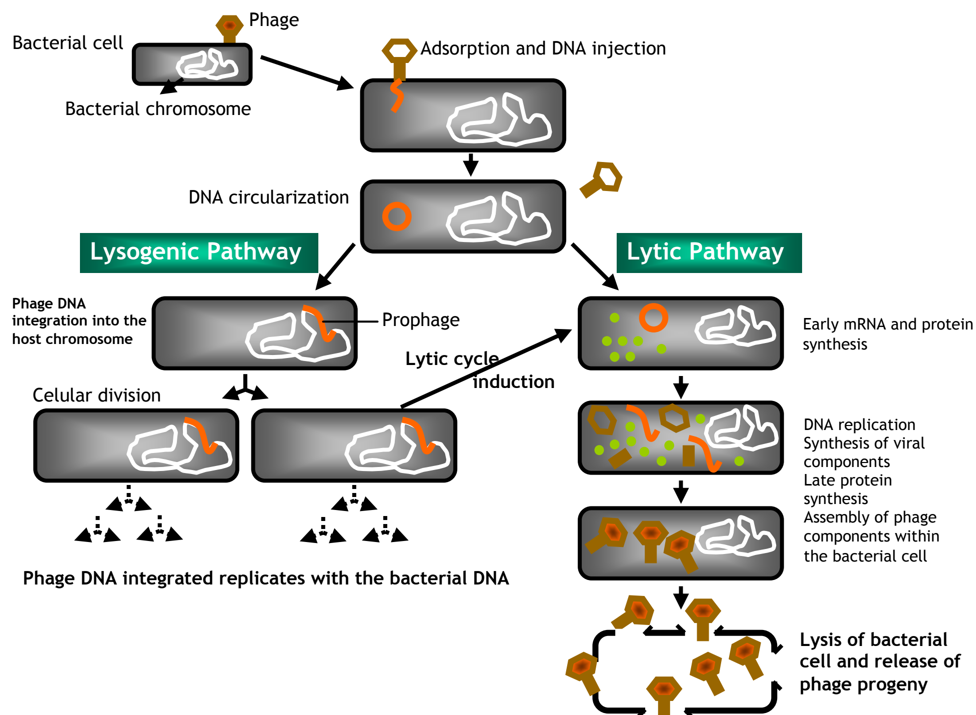


Figure 1. ***Life cycle of a temperate bacteriophage.*** See text for more complete description. Figure adapted from Hanlon, 2007.

The presence of a prophage can therefore confer upon a bacterial cell some sort of immunity to infection by other bacteriophages (Guttman *et al*., 2005; Hanlon, 2007). Cells may undergo several rounds of division but occasionally will spontaneously lyse and liberate progeny phage. Alternatively, a population of lysogenic cells may be induced to lyse by subjecting them to stress such as treatment with mutagenic agents or exposure to ultraviolet light. A chronic infection occurs, when a cell is infected and phage progeny is constantly released from the host cell by budding or extrusion without lysing it, while in a persistent infections (pseudolysogeny, phage carrier-state) phages multiply in a fraction of the population (Weinbauer, 2004).

When a prophage escapes regulation by the repressor, its DNA is cut free allowing it to embark upon a lytic cycle. However, excision of prophage DNA is often imprecise and bacterial genes adjacent to the prophage DNA may be incorporated into the infectious phage DNA and then transferred to subsequent host cells. This process is known as 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 did not transfer genes that confer selective benefit to the bacterial host as genes known to increase the survival fitness of lysogens (Brüssow *et al.*, 2004). Lysogenic bacteria may possess other advantages in terms of the acquisition of genes conferring pathogenicity or increased virulence (lysogenic conversion) (Boyd and Brüssow, 2002; Brüssow *et al.*, 2004). Many bacteriophages carry virulence genes encoding proteins that play a major role in bacterial pathogenesis. In addition, bacteriophage-bacteriophage interactions within the bacterial host cell also contribute significantly to the virulence of bacterial pathogens (Boyd *et al.*, 2001; Roucourt and Lavigne, 2009) by influencing bacterial adhesion, colonization and infection, enhancing bacterial resistance to serum and phagocytosis, altering bacterial susceptibility to antibiotics

and encoding bacterial exotoxins (Wagner and Waldor, 2002). Important toxin genes known to have been acquired by transduction include those encoding the neurotoxin of *Clostridium botulinum*, the diphtheria toxin of *Corynebacterium diphtheriae*, the Shiga toxins found in *Escherichia coli* O157 and the cholera toxin in *Vibrio cholerae* (Waldor, 1998; Wagner and Waldor, 2002). Furthermore, the advent of whole genome sequencing of bacteria has revealed that prophage or phage-like elements are abundant and not only contribute to sequence diversity but have clearly played a significant role in evolution.

1.1. The impact of Mycobacterial Phages

Mycobacteriophages are viruses of the mycobacteria. The interest in these phages derives in large part from the medical significance and biological idiosyncrasies of their hosts (Hatfull, 2000; Hatfull, 2006). Mycobacteria are acid-fast staining bacteria with characteristic waxy cell walls that can be readily divided into two groups based on their growth rate: slow-growers such as *Mycobacterium tuberculosis* and fast-growers such as *Mycobacterium smegmatis*. Several mycobacterial species are important human and animal pathogens, the most notorious being *M. tuberculosis* and *Mycobacterium leprae*, the causative agents of tuberculosis and leprosy, respectively (Hatfull and Jacobs Jr., 1994; Hatfull, 2006). The extent of tuberculosis is alarming: *M. tuberculosis* is the leading cause of human mortality from a single infectious disease and the increased prevalence of multiple drug-resistant *M. tuberculosis* strains complicates its treatment (Zhang and Telenti, 2000; Hatfull, 2005a).

Mycobacteriophages have facilitated the development of mycobacterial genetic systems, methods for mycobacterial transformation (Hatfull, 2005b) and an assortment of applications including phage-based vectors (phasmids), integration-proficient plasmids, non-antibiotic-selectable markers, gene expression systems, transposon delivery vehicles (Hatfull, 2000), gene replacement strategies using

nonreplicating vectors (Husson *et al.*, 1990), long linear DNA fragments (Balasubramanian *et al.*, 1996), incompatible plasmids (Pashley *et al.*, 2003), counterselectable markers (Pelicic *et al.*, 1996), specialized transducing shuttle plasmids (Bardarov *et al.*, 2002) and, more recently, recombineering systems for mycobacteria (van Kessel and Hatfull, 2007; van Kessel and Hatfull, 2008) and mycobacteriophages (van Kessel *et al.*, 2008; Marinelli *et al.*, 2008). They have also provided insights into viral diversity and the evolutionary mechanisms that generate them and offer potential for the development of novel methodologies for the diagnosis, prevention, and treatment of mycobacterial diseases as well as revealing interesting biological features of their unusual hosts (Hatfull, 2006; Piuri *et al.*, 2009). The first mycobacteriophages were isolated and first investigated more than 60 years ago prompted by their utility in phage-typing of clinical mycobacterial isolates (Hatfull *et al.*, 2008). Over 200 different mycobacteriophages infecting a broad variety of mycobacterial hosts have been described: some mycobacteriophages (e.g., DS6A) are specific for the *M. tuberculosis* complex, while others (e.g., Bxb1) are restricted to fast-growing strains (such as *M. smegmatis*) and still others (e.g., D29) infect both fast- and slow growing mycobacteria (Hatfull, 2006).

Currently, more than 60 complete mycobacteriophage genomes have been sequenced and all are tailed phages belonging to either the *Siphoviridae* or *Myoviridae* morphotypes; none are Podoviruses (Hatfull *et al.*, 2010). The putative gene products of these phages were grouped into "phamilies" of related sequences and the mycobacteriophages assorted into clusters and subclusters of related genomes, in an effort to provide an array of genetic tools for dissection and understanding of mycobacterial hosts and insights into viral diversity and evolution. Nine clusters of related genomes were revealed encompassing 55 of 60 genomes (clusters A-I) plus five singletons, mycobacteriophages Giles, Omega, TM4, Wildcat and Corndog (Fig. 2) (Hatfull *et al.*, 2010). The length of

mycobacteriophage genomes vary considerably, from 41 kbp to 164 kbp. There is also a broad range of GC content, varying from 56,9% to 69,1% which is similar to the host's, and genomes are typically replete with protein-coding genes with few non-coding spaces. The most common morphology is an isometric head with approximately 60 nm diameter and a long flexible tail (Hatfull *et al.*, 2008; Hatfull *et al.*, 2010). In addition to the mycobacteriophage population being large and varied, their genomes are characterized by extensive genetic mosaicism: each genome apparently contains a unique combination of individual modules, of which the majority corresponds to one or a small cluster of genes. Generation of these mosaic genomes reflects a high level of horizontal genetic exchange within the phage population, with illegitimate recombination events underlying the generation of new module boundaries (Pedulla *et al.*, 2003; Hatfull *et al.*, 2010). While the characteristic mosaic architecture of phage genomes can be explained by abundant horizontal genetic exchange events in their evolutionary history, little is known about which genomes participate in these events or the rates at which exchange occurs (Hatfull *et al.*, 2008).

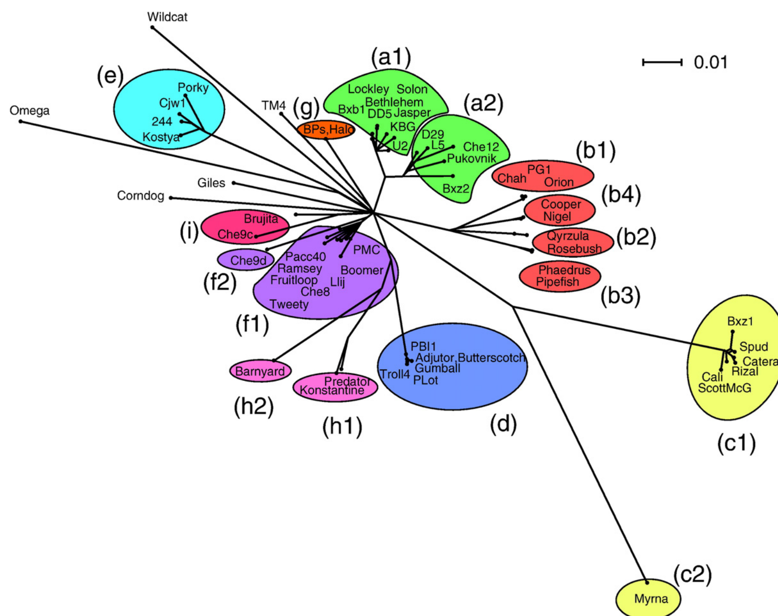


Figure 2. ***Splitstree representation of mycobacteriophage relationships.*** Mycobacteriophage predicted protein products were assorted into "phamilies" according to shared sequence similarities. Figure from Hatfull *et al.*, 2010.

While a variety of genome architectures are seen among these mycobacteriophages, most have a cluster of genes involved in virion structure and assembly. There are typically 25 to 50 genes in these clusters, which are likely to be expressed as either a single or small number of operons. Many of these genes appear to be shared by different mycobacteriophages, although the extent of sequence similarity at the amino acid level is often sufficiently low to suggest that they departed from a common ancestor. The regions of these phage genomes that do not encode structural gene products presumably provide functions which are required for DNA metabolism, regulation, metabolic optimization of viral reproduction, and lysogenic conversion of their hosts. Of the protein-coding genes, approximately 50% have no detectable sequence similarity (at the amino acid level) to other known genes (including other mycobacteriophages genes). Approximately 90% have no detectable sequence similarity to genes outside of the mycobacteriophage group, indicating that these phages are richly populated with genes that have not been previously identified (Hatfull, 2006; Hatfull *et al.*, 2008; Hatfull *et al.*, 2010). The majority of newly sequenced phage genomes, including those from mycobacteriophages, contain a large proportion of genes of unknown function, as well as many genes whose general functions are known, but that have not been previously observed in phages (Hatfull *et al.*, 2008; Hatfull *et al.*, 2010). Recently, studies attempting to understand the origin and functions of viral ORFans (ORFs with no detectable sequence similarity to any other ORF in the databases) have been reported. A first genome-wide identification and analysis of ORFans in the bacteriophage world suggested that they are exclusively or highly phage-specific. Although phage ORFans play a lesser role in horizontal transfer to

bacteria, they may be among the major players contributing to the vast phage diversity. This fact also suggests that, despite the known propensity of phages for capturing and transducing fragments of host genomes, these processes have relatively small impact on the phage gene repertoire (Liu *et al.*, 2006; Yin and Fisher, 2008). The presence of genes involved in host responses to bacterial infections as well as autoimmune diseases such as lupus suggests a more central role of bacteriophages in human diseases than previously recognized: mycobacteriophages Cjw1 gene *39* and Omega gene *61* encode close homologues of the leprosy and tuberculosis immunodominant antigen Lsr2 that is a potent stimulator of both cellular and humoral immune responses, suggesting a possible role for phages in mycobacterial virulence (Pedulla *et al.*, 2003; Hatfull *et al.*, 2005a). Phage Rosebush contains two genes (*4* and *6*) encoding homologues of enzymes involved in biosynthesis of tetrahydrobiopterin, a cofactor for a key enzyme in the host defence against mycobacterial infections, nitric oxide synthase (Hatfull *et al.*, 2005a). Bxz1 *gp220* encodes a homolog ~35% identical to the human Ro protein, a major target of the autoimmune response in Lupus and Sjogren´s diseases. The function of the Ro ribonucleoprotein is not known, but it is implicated in 5S RNA processing in *Xenopus*, dauer formation in *Caenorhabditis elegans* and resistance to ultraviolet light in *Deinococcus radiodurans*; the presence of a Ro homolog in Bxz1 raises the possibility that bacteriophages could act in concert with their hosts to stimulate autoimmunity (Pedulla *et al.*, 2003; Hatfull *et al.*, 2005a). Conversely, toxin genes have not been identified in the sequenced mycobacterial or mycobacteriophage genomes and consequently, have not been implicated in mycobacterial virulence. The role of phages in the virulence of mycobacterial pathogens remains unclear and the genomes of *M. tuberculosis* and *M. leprae* contain no full-length prophages. However, both of the two sequenced genomes of *M. tuberculosis* contain two small prophage-like elements, ϕRv1 and ϕRv2, at least one of which (ϕRv1) has an active integration system; both ϕRv1 and

ϕRv2 contain capsid genes and could in principle form virus-like particles. Nevertheless, while all clinical isolates of *M. tuberculosis* appear to contain at least one of these elements, it is uncertain whether they play any role in the physiology of the host (Pedulla *et al.*, 2003; Hatfull *et al.*, 2005a; Hatfull, 2008b).

In summary, the mycobacteriophages are a remarkably diverse group of viruses whose characterization has provided helpful insights into the mosaic nature of bacteriophage genomes and the evolutionary mechanisms that give rise to them. There are huge numbers of uncharacterized mycobacteriophage genes that may have specific utilities such as the development of new tools for the genetic manipulation of mycobacteria or by providing new insights and approaches necessary to overcome the humankind's most deadly microbial enemy, *M. tuberculosis*. Indeed, the genes encoding for the lysis proteins of mycobacteriophages represent an attractive direction for further work on their potential as natural anti-mycobacterial compounds.

2. Bacteriophage-Induced Host Cell Lysis

In general, bacteriophages must lyse their host cell to liberate the newly assembled progeny virions to the extracellular milieu (Young and Wang, 2006). Thus, lysis is a programmed event like all the other steps in the infectious cycle and of major importance regarding phage survival and ecological fitness (Wang, 2006). A sharply time-defined and efficient release of phage progeny is crucial to maximize both the burst-size and the opportunity to infect new hosts (Wang, 2006; São-José *et al.*, 2007). The main barrier to host lysis is the continuous meshwork of peptidoglycan, a strong, stable structure that allows the bacterial envelope to withstand internal osmotic pressure (Young *et al.*, 2000): compromising the cell wall is thus the fundamental goal for lytic processes. With the exception of filamentous phages that, as a result of their unique morphology and morphogenesis

can extrude through the envelope without fatal consequences for the host, all other phages must either degrade or otherwise compromise the peptidoglycan to cause lysis (Young *et al*., 2000; São-José *et al*., 2007). The filamentous phages constitute a large family of single-stranded DNA viruses that infect Gram-negative bacteria using pili as receptors. In contrast to other bacterial viruses, filamentous phage particles do not accumulate in the cytoplasm and cell lysis does not occur during phage progeny release. Instead, filamentous phages are produced by a concerted mechanism of assembly and secretion across both membranes of the Gram-negative cell (Russel, 1995; São-José *et al*., 2007).

From a mechanistic point of view, bacteriophages lyse the bacteria they infect by two essentially distinct modes:

1. Small single-stranded nucleic acid phages, with very limited genomes under 6 kb, have a single phage protein to elicit lysis, presumably because of their restricted coding capacity (Bernhardt *et al*., 2002a). This protein with no apparent muralytic activity causes lysis by acting as a specific inhibitor of an enzyme in the multi-step pathway of murein biosynthesis. As this protein acts as a "protein antibiotic", inhibiting cell wall synthesis and promoting a septal catastrophe as the cell attempts division, lysis is dependent on cell growth (Bernhardt *et al*., 2002a; Young and Wang, 2006). The only well characterized phages of this class are coliphages and among them there are three different prototypical lysis proteins: i) the E protein of the single-stranded (ss)-DNA bacteriophage ϕX174 (*Microviridae*), ii) the L protein of the ss-RNA bacteriophage MS2 (*Leviviridae*), and iii) the A_2 protein of the ss-RNA bacteriophage Qβ (*Alloleviviridae*) (Young *et al*., 2000; Bernhardt *et al.*, 2002a). For ϕX174 (Bernhardt *et al*., 2000) and Qβ (Bernhardt *et al*., 2001a), it has been demonstrated unequivocally that the single lysis protein produced, for which it was proposed the term amurin (Bernhardt *et al*., 2002a) acts as an inhibitor of a specific step in murein biosynthesis. However, the inhibited steps are different for the two phages: ϕX174 E and Qβ A_2 inhibit separate

enzymes, MraY and MurA, respectively, of the murein synthesis pathway (Young, 2005). Although the regulation of expression of the L protein of phage MS2 has been well studied and evidences indicate that interference with the stability of a large translation regulatory hairpin affects lysis timing, the L target remains unknown (São-José *et al*., 2007). The lytic capacity of E requires a host protein, SlyD (Roof *et al*., 1997) that is a peptidyl-prolyl cis-trans-isomerase (PPIase) or rotamase of the FKBP family (Roof *et al*., 1994; Hottenrott *et al*., 1997) that although is not the target of E lytic function, is required for protein stabilization (Bernhardt *et al*., 2000, Bernhardt *et al*., 2002b). The molecular target for the bacteriolytic E protein from bacteriophage ϕX174 is known to be the enzyme phospho-*N*-acetylmuramoyl-pentapeptide transferase (MraY), an integral membrane protein (Bernhardt *et al*., 2001b; Mendel *et al*., 2006) that catalyses the transfer of the MurNac pentapeptide from a UDP-MurNac pentapeptide to the polyisoprenoid carrier undecaprenylphosphate originating lipid I, the first membrane-bound murein precursor (Zheng *et al*., 2008). Phage Qβ A_2 is a dual functional protein: it is simultaneously the Qβ capsid protein responsible for adsorption to the host pilus and the lysis effector (Karnik and Billeter, 1983; Bernhardt *et al*., 2001a). Genetic and biochemical characterization has identified its target as UDP-*N*-acetylglucosamine enolpyruvyl transferase (MurA), a soluble enzyme that catalyzes the first committed step of the murein precursor biosynthesis pathway (Bernhardt *et al*., 2001a).

2. Phages with double-stranded nucleic acid genomes use the "holin-endolysin" strategy to achieve host cell lysis. The phage elaborates a soluble murein degrading enzyme, an endolysin, specifically dedicated to degrade the host cell wall, and a second membrane embedded protein, the holin, which serves to release or activate the endolysin at a precisely defined time (Young, 1992; São-José *et al*., 2003). Moreover, the holin function is responsible for the crucial regulation of lysis timing as detailed below.

Genes controlling lysis are typically late-expressed and must be transcriptionally regulated (Young, 1992). Endolysins and holins encoded by different phages can be extremely diverse. Interestingly, pairwise combinations of different endolysins with related holins or similar endolysins with distinct holins are frequently found in the natural phage population. Still, it is apparent that phages have evolved additional means of fine-tuning the lytic schedule, through the synthesis of holin antagonists, referred to as antiholins. Thus, what characteristics would be appropriate for a system evolved to effect host lysis at a defined time? First, the lysis system should be as saltatory as possible and should not affect the productivity of the infected host, in terms of virion assembly, until the programmed time of lysis. Second, the lysis system should be very efficient and rapid once the infective cycle is terminated. Finally, the timer should be capable of being over-ridden in real time, should circumstances change during an infection. All these characteristics are exhibited by the holin-endolysin system of lysis and are mostly due to the properties of holins alone (Young and Wang, 2006).

2.1. Holins: Saltatory lethal membrane permeabilization

Holins are much more diverse and frequently unique with respect to their primary sequence (Young, 2002) and they may be defined as a single gene, encoding a putative small protein with at least one transmembrane domain (TMD) found in the vicinity of the endolysin gene (Wang *et al*., 2000; Young, 2002). Many holins genes have potential dual-start motifs which allow for the production of a shorter holin and a longer antiholin (Young, 2005). During the late phase of phage development, holins progressively accumulate in the cytoplasmic membrane of the host and while the proton-motive force (p.m.f.) is maintained, they assemble into oligomers and rafts of intrinsic stability (Young and Wang, 2006). Holins can be prematurely triggered by membrane depolarization with energy poisons such as cyanide and dinitrophenol (Gründling *et al*., 2001; Young, 2005). This observation

led to a model for holin timing in which at a precise time programmed into its primary structure, the holin suddenly causes disruption of the membrane with non-specific hole formation and collapse of the membrane potential which sets the time of lysis by allowing the destruction of the cell wall by the released or activated phage encoded muralytic enzymes, the endolysins (São-José *et al.*, 2003; Young and Wang, 2006).

2.1.1. The phage λ paradigm

In lambdoid phages, all late genes are transcribed from a single promoter, designated P_R´in bacteriophage λ (Young, 1992). The first genes of the late operon are the phage lysis genes: *S*, *R*, *Rz* and *Rz1* (Fig. 3) (Young, 1992). The S gene encodes the holin (S_{105}) and the antiholin (S_{107}), as a result of translational initiation sites defined by codons 3 and 1 (Bläsi *et al.*, 1989; Bläsi *et al.*, 1990), respectively. The *R* gene encodes the endolysin which is a 18 kDa soluble murein transglycosylase that cleaves the glycosidic bond between *N*-acetylglucosamine and *N*-acetylmuramic acid residues, forming a cyclic product (Bieńkowska-Szewczyk *et al.*, 1981). The DNA sequence in the lysis cassette contains two additional genes, designated *Rz* (Young *et al.*, 1979) and *Rz1*, the later fully embedded in *Rz* in a +1 frame (Hanych *et al.*, 1993; Kedzierska *et al.*, 1996). *Rz*/*Rz1* gene products constitute a protein complex that somehow attacks the outer membrane (OM), cleaving the oligopeptide linkages between the peptidoglycan and the outer membrane lipoprotein (Lpp) (Young and Wang, 2006; Berry *et al.*, 2008).

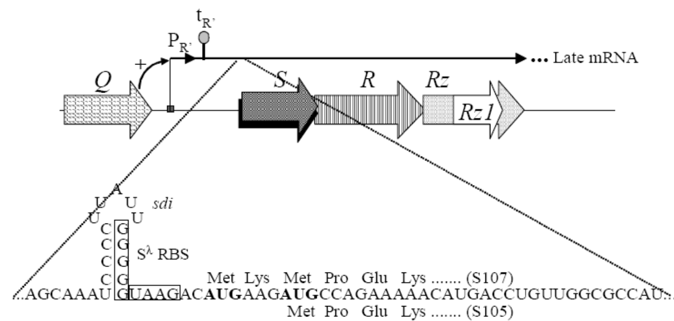


Figure 3. ***Features of the lysis gene region in bacteriophage lambda*** (see text for further details). Figure from São-José *et al.*, 2003.

The phage λ S_{105} Holin

The S holin or S_{105}, is a 105-residue integral cytoplasmic membrane protein with a three transmembrane domains (TMD) topology (N-out, C-in) (Gründling *et al.*, 2000a). Many S_{105} mutants have been isolated (Raab *et al.*, 1986; Raab *et al.*, 1988), including not only unconditional lysis defectives but also a plethora of alleles which accelerate or delay the onset of lysis. A particularly sensitive microdomain appears to be in the middle of TMD_2. Both λS_{A52V} and λS_{A52G} are non-plaque-formers, but for opposite reasons. S_{A52V} is unconditionally lysis-defective, whereas S_{A52G} supports catastrophically early lysis (Raab *et al.*, 1986; Raab *et al.*, 1988; Johnson-Boaz *et al.*, 1994). Similarly, at Cys51, substitution to a Ser accelerates lysis and replacement with a Tyr abolishes lysis. Thus, every potential TMD surface of S appears to be involved in setting the lysis clock (Gründling *et al.*, 2000a; Gründling *et al.*, 2000b). Although the sequences of the short periplasmic N-terminal domain and the cytoplasmic C-terminal domain are not essential for lysis, functional assembly of the λS holin requires periplasmic localization of its N-terminus (Graschopf and Bläsi, 1999a) and distribution of amino-terminal charged amino acids as well as the total amino-terminal net charge of S_{105} and S_{107} influence their lethal potential (Steiner and Bläsi, 1993). Furthermore, the C-terminal sequence of the λS holin constitutes a cytoplasmic regulatory domain which affects the timing of lysis although is not involved in the formation of the lethal membrane lesion nor in the “dual-start regulation” conserved in lambdoid holins (Bläsi *et al.*, 1999). Strangely, however, although a $G_2H_6G_2$ tag inserted near the membrane interface in the C-terminal cytoplasmic domain resulted in a functional lytic S protein, which has been the source of purified holin (Smith and Young, 1998), the simple addition of a hexahistidine tag

in the C-terminal cytoplasmic domain disrupts holin function (Smith *et al*., 1998). Recently, it has been demonstrated that the N-terminus of S_{105} retains its formylated Met residue but that the N-terminus of S_{107} is fully deformylated. This supports the model that in S_{105}, TMD_1 inserts into the membrane very rapidly but that in S_{107}, it is retained in the cytoplasm. Further, it reveals that, compared to S_{105}, S_{107} has two extra positively charged moieties, Lys2 and the free N-terminal amino group, to hinder its penetration into an energized membrane. Moreover, an allele, $S_{105\Delta TMD1}$, with TMD_1 deleted, was found to be defective in lysis, insensitive to membrane depolarization, and dominant to the wild-type allele, indicating that the lysis defective, antiholin character of S_{107} is due to the absence of TMD_1 from the bilayer rather than to its ectopic localization at the inner face of the cytoplasmic membrane (White *et al*., 2010). λS holin has been shown by crosslinking studies, to form oligomers in the inner membrane of *E. coli*, but the ultimate degree of oligomerization in unknown (Zagotta and Wilson, 1990). Oligomerization does not depend on disulfide bridge formation (Gründling *et al*., 2000b) and the hydrophilic C-terminal part of the lambda S holin is non-essential (Rietsch *et al*., 1997). Oxidative dimer formation between S variants with single cysteines in the hydrophobic core of the TMD_2 revealed that positions 48 and 51 are on a dimer interface (Gründling *et al*., 2000b). Furthermore, it has been suggested that the ability of S molecules to dimerize is not sufficient for the lytic step in holin function: oligomerization and a concerted conformational change which is equivalent to triggering of hole formation are also required: some lysis-defective alleles appear to be blocked at the monomer, dimer and oligomers stages (Gründling *et al*., 2000b).

Until recently, nothing was known about the nature of the λS membrane lesions. To calibrate the scale of the S-hole, a C-terminal fusion of the R endolysin with full-length β-galactosidase (β-gal) was constructed and the hybrid R-β-gal

product, an active tetrameric protein greater than 480 kDa in mass, was fully functional in lysis mediated by the S holin (Wang *et al.*, 2003). This result suggested that at least some of the lesions created by the triggering of S are of size comparable to the large-scale lesions, in excess of 30 nm diameter (Wang *et al.*, 2003; Young and Wang, 2006). In this study a holin lesion model was formulated in which holin protomers accumulate in the cytoplasmic membrane, oligomerize and ultimately form large two dimensional protein aggregates, designated as 'death rafts', from which lipid molecules are largely excluded by intimate interaction between individual holins via their transmembrane domains. Opening of an aqueous channel, with consequent local depolarization of the cytoplasmic membrane, triggers conformational change in the holins and subsequent dispersion of the subunits into the protein-bounded lesion in the cytoplasmic membrane (Wang *et al*., 2003). Latter, Savva and collaborators (2008) made use of electron microscopy and single-particle analysis to characterize the structures formed by λS holin *in vitro*. Purified S_{105} forms rings of at least two size classes, the most common having inner and outer diameters of 8.5 and 23 nm, respectively, and containing approximately 72 S_{105} monomers. The height of these rings, 4 nm, closely matches the thickness of the lipid bilayer (Krupovič and Bamford, 2008; Savva *et al*., 2008). More recently, cryo-electron microscopy (cryo-EM) analysis revealed that the scale of the holes is at least an order of magnitude greater than any previously described membrane channel, with an average diameter of 340 nm, some exceeding 1µm. The large holes can be viewed as supporting the notion that at the time of lethal triggering, the S_{105} holin exists in such large aggregates, leading to one or a small number of holes rather than many smaller holes distributed throughout the membrane (Dewey *et al*., 2010) as initially thought.

The phage λ S_{107} Antiholin

The only essential genes for phage λ-mediated lysis are the S_{105} holin and the R endolysin but the presence of S_{107} delays lysis onset, allowing for a larger burst size and increasing effective hole formation after triggering (Gründling *et al*. 2000a). The λS_{107} antiholin and the λS_{105} holin are encoded in frame in the same *S* gene, and the proportion of S_{105} and S_{107} is for wild-type S, ~2:1 by the time of lysis (Chang *et al*., 1995). The differential expression of S_{107} and S_{105} is due to a structure directed initiation loop (*sdi*) overlapping the Shine-Dalgarno (SD) sequence in the S mRNA which hinders initiation of the S_{107} start codon (Bläsi and Young, 1996). The λ*S* gene is characterized by a dual-start motif: the S_{105} holin and S_{107} antiholin share the same 105 amino acid sequence but S_{107} has an extra Met and Lys residues in the N-terminus (Fig. 3). These extra residues in S_{107} confer two extra positive charges comparing to S_{105}, one charge by deformylation of Met_1 and another by the Lys_2 residue (Bläsi and Young, 1996): extra charges are known to hinder the translocation of the first TMD in an energized-membrane resulting in the altered topology of S_{107} compared to S_{105} (Graschopf and Bläsi, 1999b). The S_{105} holin has a three TMD topology whereas in the antiholin the first hydrophobic segment is unlikely to span the membrane. Furthermore, S_{107} exerts its inhibitory effect by dimerizing with S_{105}, creating heterodimers which are either non-functional or of reduced functional capacity (Gründling *et al*., 2000c). The dissipation of membrane proton-motive force triggers the translocation of the first TMD of S_{107} which then becomes a topological homolog of S_{105} with similar hole/lesion-forming properties (Graschopf and Bläsi, 1999b).

Rz and Rz1 genes

The *Rz* and *Rz1* genes and their equivalents are unique in biology: they are the only genes that share the same DNA in different reading frames but associated with the same phenotype (Hanych *et al*., 1993; Zhang and Young, 1999).

The *Rz* gene of bacteriophage λ belongs to the late operon which encompasses more than 25 genes for host cell lysis, phage DNA packaging and phage morphogenesis (Hanych *et al*., 1993). This third lysis gene was revealed by Tn*903* mutagenesis of the λ phage which confers Mg^{2+}-dependent lysis defect (Young *et al*., 1979). Cells infected with λ*Rz*:*Tn903* in the presence of 10 mM $MgCl_2$ do not undergo lysis; instead form spherical cells which are mechanically fragile and gradually lose refractility (Young *et al*., 1979). Until recently (Berry *et al*., 2008) nothing was known about *Rz* function. It was suggested that *Rz* might encode a murein-specific endopeptidase detected in lambda lysates (Taylor, 1971; Young *et al*., 1979; Bieńkowska-Szewczyk, 1980) which cleaves the oligopeptide crosslinks between outer membrane protein and the peptidoglycan (Zhang and Young, 1999). *Rz* encodes a 153 amino acid polypeptide with a hydrophobic N-terminus that is predicted to be either a secretory signal or N-terminal transmembrane domain by sequence analysis algorithms (Berry *et al*., 2008). Hanych *et al*. (1993), in an attempt to clone the *Rz* gene under a highly active promoter system, generated constructs in which an internal portion of the *Rz* gene was expressed. An unexpected 6.5 kDa polypeptide (Rz1) was observed to accumulate in the membrane fraction, which was found to result from the use of an internal, out-of-frame translational initiation site within the *Rz* gene. The sequence of the predicted Rz1 protein spans only 60 codons and contains a signal peptidase II cleavage site at Cys20 (Hanych *et al*., 1993; Zhang and Young, 1999). Rz1 processing was blocked by the signal peptidase II inhibitor globomycin, and by labelling with palmitate. The sequence of the mature Rz1 lipoprotein is very unusual containing almost 25% proline residues and the Rz1 protein is located almost exclusively in the outer membrane of *E. coli* (Kedzierska *et al*., 1996). It was demonstrated that, at physiological levels of expression, Rz and Rz1 are localized to inner membrane (IM) and outer membrane (OM), respectively, with their C-terminal domains predicted to lie in the periplasm (Berry *et al*., 2008).

Furthermore, there is some genetic and phylogenetic evidence that the Rz and Rz1 proteins interact. First, the *Rz/Rz1* equivalents from phage P2, *lysB*/*lysC*, complement defects in the lambda genes, but only as a cognate pair (Markov *et al*., 2004). In addition, yeast two-hybrid analysis of a library of phage T7 genes found multiple positives between clones with the last 10 codons of 18.7, the T7 *Rz1* equivalent, and the last 50 codons of 18.5, the *Rz* equivalent (Bartel *et al*., 1996). These data suggest that Rz and Rz1 interact in a C-terminus to C-terminus fashion, which may account for the architectures of the embedded and overlapped *Rz*/*Rz1* genes, in that these unusual arrangements minimize the likelihood of recombinational separation of the interacting domains (Summer *et al*., 2007). Finally, recently it was demonstrated that Rz and Rz1 proteins form a complex that improves the effectiveness of endolysin degradation of the murein layer and that function of this complex is lost if either protein is improperly localized. The OM is covalently attached to the cell wall by oligopeptide linkages between Lpp and the murein; a complex that spans the periplasm could conceivably push the IM away from the murein layer which might make endolysin-mediated murein degradation more efficient. In this study, a model for Rz/Rz1 function in host cell lysis was proposed: the first step is the temporally programmed permeabilization of the cytoplasmic membrane by the holin, resulting in the release of a cytoplasmic endolysin or the activation of a signal-anchor-release (SAR) endolysin. The second stage is the endolysin-dependent degradation of the murein layer followed by a third stage involving the fusion of the IM and OM mediated by the Rz-Rz1 complexes. The three steps of phage lysis mediated by holins, endolysins and Rz-Rz1 complexes form a sequential pathway in which holin function is required for endolysin function, which is in turn required for Rz-Rz1 function, although they are mechanistically independent (i.e. do not require heterotypic interactions with each other) (Berry *et al*., 2008).

A comprehensive bioinformatic search found Rz/Rz1 equivalents in nearly all phages of Gram-negative but not Gram-positive hosts (Summer *et al*., 2007). The diversity of Rz/Rz1 equivalents was striking: 37 unrelated sequence families, including eight families with the embedded structure found in lambda, but also families in which Rz1 extends beyond Rz (overlapped structure; 23 families) and others where the two genes are completely separated (6 families) (Fig. 4). The physical association of *Rz* and *Rz1* gene pairs, as exemplified by the embedded and overlapped classes, suggests that there is strong selective pressure for the linkage of these genes. The widespread occurrence of these genes among phages of Gram-negative hosts suggests that they confer an advantage in nature that is not apparent under laboratory conditions (Summer *et al*., 2007). The most surprising outcome of the global search for phage *Rz/Rz1* equivalents was the discovery of a new class of proteins, the spanins, which are functionally comparable to Rz-Rz1 pairs and provided new insights into Rz-Rz1 function. Spanins have a lipoprotein signal peptide as well as a C-terminal TMD and, thus, should provide a physical connection between the inner and outer bacterial membranes (Summer *et al*., 2007).

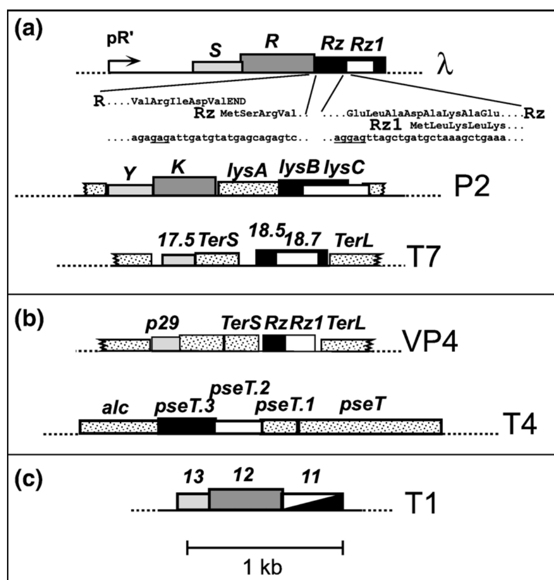


Figure 4. *Rz/Rz1 and spanin gene arrangements.* In each map, light grey, dark grey, black and white boxes indicate holin, endolysin, *Rz* and *Rz1* genes, respectively.
(a) The prototype embedded (λ, T7), or overlapped (P2) *Rz*/*Rz1* genes.
(b) Separated *Rz* and *Rz1* genes (VP4 and T4). For phage T4, the start codon for the *Rz1* (pseT.2) gene overlaps the stop codon of the *Rz* (pseT.3) gene, although the protein coding sequences are separated.
(c) The T1 spanin gene, shown as a box with half black and half white filling. Figure from Summer *et al*., 2007.

The ability of the T1 spanin gene to complement the *Rz/Rz1* lysis phenotype provides further support to the notion that Rz and Rz1 interact and that complexes span the periplasm and connect the inner and outer membranes (Summer *et al.*, 2007). Rz and Rz1 equivalents were also identified in the tailless membrane containing dsDNA virus PRD1 (*Tectiviridae*) that infects a variety of Gram-negative hosts. It was suggested that Rz-Rz1 interact not only with each other but also, as a complex with the holin, to transmit the mechanical stress of the holin mediated lesion in the cytoplasmic membrane to the outer membrane, resulting in its disintegration (Krupovič *et al.*, 2008).

2.2. Diversity in Holin-Endolysin Systems

Holin and Antiholin Diversity

Holins are extremely diverse, found in many unrelated sequence families with at least three membrane topologies suggesting that they may have evolved from multiple distinct origins (Wang *et al.*, 2000) to allow precisely scheduled, efficient lysis and rapid adjustment of the lysis time, either on the basis of genetic selection, or in some cases, in real time in response to environmental changes. Holins are small membrane proteins and are currently grouped into three classes based on their membrane topology (Fig. 5). The two main classes are class I, with three TMDs (N side out, C side in), and class II, with 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 (encoded by *gpt)* and its relatives in T-even phages, defines class III with its relatively large C-terminal periplasmic domain and only one TMD (N side in, C side out) (Young, 2005; Young and Wang, 2006). In many cases, two holin-like open-reading frames are found in vicinity of the endolysin gene which led to the suggestion that one of the genes could encode a holin effector and the other the antiholin. In 20 out of 46 cases of phages infecting Gram-positive hosts in

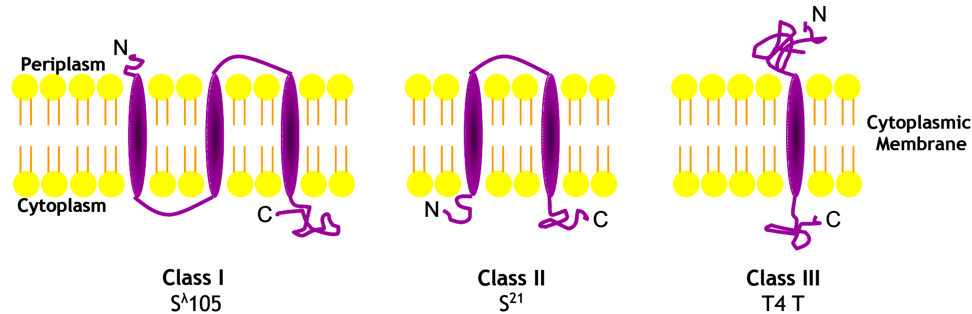


Figure 5. ***Schematic representation of known topologies of described phage holins.*** Adapted from São-José *et al*., 2003.

which a typical holin-lysin cassette was reported, a second holin-like gene, located immediately upstream of the *hol-lys* pair, could be detected (São-José *et al*., 2003). Also, it can not be excluded the situation where the holin functional unit is built from a complex of different polypeptides, rather than from a single protein. In *Bacillus subtilis* defective prophage PBSX (Wood *et al*., 1990) the lysis cassette consists of four genes, *xepA*, *xhlA*, *xhlB*, *xlyA*, organized as a late expressed operon (Longchamp *et al*., 1994). The putative product of *xhlA* is a small protein with a hydrophilic N-terminus and a putative transmembrane helix located at the C-terminus, suggesting that XhlA may be membrane associated. *xhlB* and *xlyA* encode a predicted class II holin and the endolysin, respectively. From expression analysis of different PBSX derivatives deleted in different regions of the lysis operon it was demonstrated that only co-expression of *xhlB* and *xhlA* with the endolysin would efficiently promote lysis. Neither the putative class II holin nor the single TMD polypeptide could trigger lysis in pairwise combinations with the endolysin. Expression of both XhlA and XhlB is necessary to effect host cell lysis of *B. subtilis* and the proteins have thus been suggested to associate in the membrane in order to allow endolysin release (Krogh *et al*., 1998; São-José *et al*., 2003).

Many holin genes have potential dual-start motifs which would allow for the production of a shorter holin and a longer antiholin as exemplified by the λS holin. However, antiholins are not always encoded by a dual-start motif. In fact, antiholins have been found to have widely different topologies and molecular features: i) a dual-start holin-antiholin system was described for the *Listeria* phage A118. The *hol118* gene features a dual-start motif similar to λ*S*. However, in contrast to λS, inactivation of AUG-1 or AUG-2 showed no significant influence on lysis timing. Toeprinting assays on *hol118* mRNA revealed an unexpected translational start codon (AUG-3) at nucleotide position 40. The intragenic, in-frame translated Hol118 (83) product, which is devoid of its first transmembrane domain, acts as a functional inhibitor and constitutes a key part of the lysis clock of A118 (Vukov *et al*., 2003); ii) in phage T4 the antiholin gprI is predicted to be either a type II signal anchor protein with an N-terminal TMD, or a processed, secreted periplasmic protein (Young, 2005). Because of the asynchrony of both adsorption and host lysis, waves of phage particles arrive to infect already infected cells, and the secondary infections cause the imposition of the lysis inhibition (LIN) state. Phenotypic analysis of deletion and point mutations in *gpt* holin gene, indicates that the periplasmic domain of T is the major determinant of the timing mechanism and is involved in the LIN response (Ramanculov and Young, 2001a; Ramanculov and Young, 2001b). Somehow, a signal provided by the superinfecting phage particles cause gprI to inhibit gpt (Paddison *et al*., 1998; Tran *et al*., 2005). Superinfection somehow shifts the steady state of *rI* synthesis and degradation in favour of accumulation, allowing the formation of gprI-gpt inhibition complexes (Ramanculov and Young, 2001c). The N-terminal sequence of RI comprises a SAR domain, which causes secretion of RI in a membrane-tethered form that is subsequently released into the periplasm without proteolytic processing. Moreover, the SAR domain confers both functional lability and DegP-mediated proteolytic instability on the released form of RI (Tran *et al*., 2007); iii)

genetic or physiological evidence for other independent antiholin genes is available for the coliphages P1 and P2. In P1, although the known lysis genes are all under control of the late gene activator, the holin and antiholin genes, *lydAB*, are adjacent but unlinked to the endolysin gene, *lyz*. Remarkably, the antiholin LydB is essential to obtain a productive burst; in its absence, lysis occurs catastrophically early (Walker and Walker Jr., 1980) which indicates that LydB antagonizes the P1 LydA holin (Łobocka *et al*., 2004). In phage P2, the lysis cassette is comprised of genes *Y*, *K*, *lysA*, *lysB* and *lysC*, encoding, respectively, a class I holin, an orthologue of the λR endolysin, an antiholin and functional analogues of λ Rz/Rz1. LysA appears to be an integral membrane protein with four TMDs, is predicted to have two basic residues at its cytoplasmic N-terminus reminiscent of the antiholins of the class II holin genes and is Y-specific antiholin (Ziermann *et al*., 1994; Schmidt *et al.*, 1996).

Endolysin Diversity

Endolysin is a rather generic term used to describe a range of dsDNA bacteriophage-encoded peptidoglycan hydrolases, which are synthesized in phage-infected cells at the end of the multiplication cycle. These enzymes are also known as phage lysozymes, lysins, or muralytic enzymes. They are uniformly characterized by their ability to directly target bonds in the peptidoglycan (PG) layer of the bacterial cell wall; the result of this activity is degradation of the rigid murein layer and release of newly assembled virions by way of lysis (Loessner, 2005, Borysowski *et al*., 2006). The term endolysin was first used to designate a probably proteinaceous lytic substance synthesized in bacterial cells during phage multiplication and acting on the cell wall from inside the cell (Jacob and Fuerst, 1958). Lysins should, therefore, be clearly distinguished from the lytic enzymes, which, in some phages, are an integral component of the virion and that locally digest the cell wall from the outside to enable the phage genome to be injected into

the host cell; an example of such an enzyme is the tail lysozyme of bacteriophage T4 (Arisaka *et al*., 2003; Kanamaru *et al*., 2005). The bacterial cell wall protects the cell protoplast from mechanical damage and osmotic rupture (lysis), and is, therefore, essential to bacterial viability. The main constituent of the bacterial cell wall is peptidoglycan. Peptidoglycan is composed of the repeat polymer of the amino sugars *N*-acetylglucosamine and *N*-acetylmuramic acid, linked together by β-1, 4 glycosidic bonds, and tetrapeptide side chains attached to the lactyl group of the muramic acid by amide bonds. Adjacent tetrapeptides may be cross-linked by an interpeptide bond (in Gram-negative bacteria) or by an interpeptide bridge (in Gram-positive bacteria) (Borysowski *et al*., 2006). In Gram-positive bacteria, the cell wall is thick (15–80 nm) and consists of several layers of peptidoglycan associated with teichoic acids (Schleifer and Kandler, 1972; Beveridge, 2000). In contrast, the cell wall of Gram-negative bacteria is relatively thin (10 nm) and is composed of a single layer of peptidoglycan surrounded by the outer membrane (Costerton *et al*., 1974; Beveridge, 1999). When analyzing the products encoded in a phage genomic sequence, perhaps the only lysis function which can be unequivocally identified is the one corresponding to the endolysin (São-José *et al*., 2003). Almost all endolysins are late proteins, produced during the late phase of gene expression in the phage infection cycle; a notable exception is T7 gp3.5, which is an early protein and has an important function as a specific inhibitor of T7 RNA polymerase (Moffatt and Studier, 1987; Cheng *et al*., 1994). Phage endolysins distribute between five major functional types; i) *N*-acetylmuramidases (lysozymes); ii) endo-β-*N*-acetylglucosaminidases, which all hydrolyze the β-1,4 glycosidic bonds in the murein; iii) transglycosylases, which attack the same bonds but form a cyclic 1,6-anhydro-*N*-acetylmuramic acid product; iv) *N*-acetyl-muramoyl-L-alanine amidases, which hydrolyze the amide bond between the *N*-acetylmuramic acid and L-alanine residues in the oligopeptide crosslinking chains and v) endopeptidases, which attack the peptide bonds in the same chains (Fig. 6)

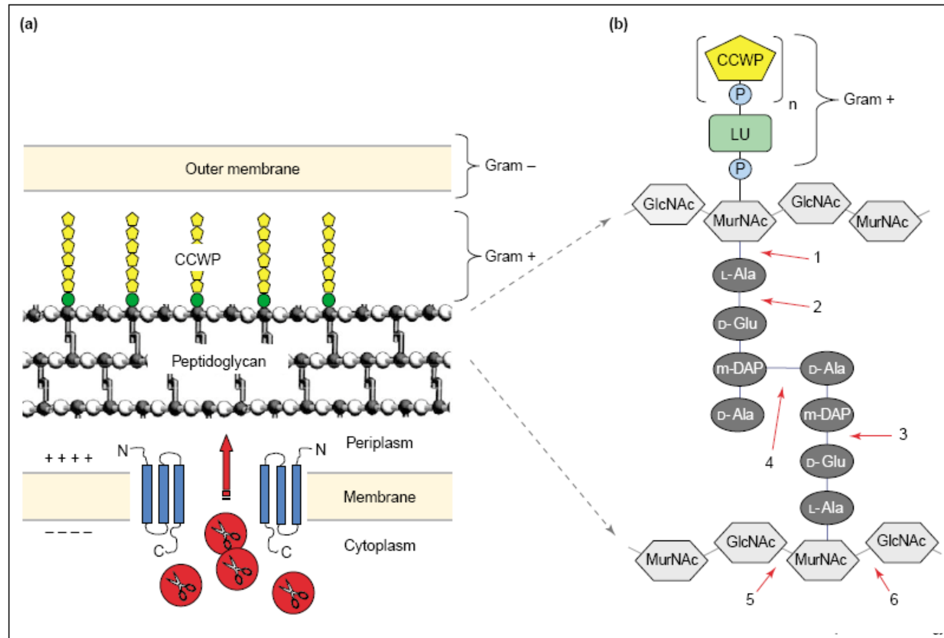


Figure 6. ***Bacterial cell wall structure and endolysin targets.*** **(a)** Schematic representation of the bacterial cell wall, and one way of phage endolysins gain access to their substrate. **(b)** The bonds potentially attacked by endolysins of different enzymatic specificities are indicated by 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-β-D-glucosaminidase; and 6) *N*-acetyl-β-D-muramidase (also known as muramoylhydrolase and 'lysozyme') and lytic transglycosylase. 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. Figure from Loessner, 2005.

(Loessner, 2005). Typically, one endolysin displays only one kind of muralytic activity. However, at least four bifunctional lysins have also been reported, enzymes harbouring two independent muralytic activities. Examples are the endolysins encoded by *Streptococcus agalactiae* bacteriophage B30 (muramidase and endopeptidase) (Pritchard *et al.*, 2004), *Staphylococcus aureus* phage ϕ11 (endopeptidase and amidase) (Navarre *et al.*, 1999), *S. agalactiae* phage NCTC 11261 (endopeptidase and muramidase) (Cheng *et al.*, 2005), and *Staphylococcus warneri* M phage ϕWMY (endopeptidase and amidase) (Yokoi *et al.*, 2005).

With few predicted exceptions, the endolysins encoded by phages of Gram-positive bacteria present a conserved modular structure: they are composed of at

least two clearly separated functional domains: the N-terminal domain(s) generally harbour the enzymatic activity, whereas the cell wall binding domains (CWBDs) located at the C-terminal direct the enzymes to their substrates and may restrain the enzyme lytic action to a particular type of cell wall (Loessner, 2005; São-José *et al*., 2007).

2.3 The other paradigm: sec-Mediated Lysis

One of the hallmarks of lysis regulation was that phage endolysins would not reach the cell wall before the formation of holin lesions (Fig. 7A). The possibility that phage endolysins could be endowed with secretion signals was seldom contemplated (São-José *et al*., 2003; Young and Wang, 2006). Exceptions were suggested for XlyA (Longchamp *et al*., 1994) and Ply21 (Loessner *et al*., 1997), two endolysins encoded by *Bacillus subtilis* defective phage PBSX and *B. cereus* phage TP21, respectively.

A paradigm shift began with studies by Santos and colleagues on the phage fOg44, which grows on the Gram-positive bacterium *Oenococcus oeni* (Parreira *et al*., 1999). This phage encodes an endolysin (Lys44) with a bona-fide signal peptide (SP) and muramidase activity and is processed by the leader peptidase during infection, presumptive evidence that the endolysin is exported by the *sec* machinery. The results were clearly suggestive of a continuous targeting of the endolysin to the cell wall in its active, SP-processed form, from the moment of its synthesis (Fig. 7B) (São-José *et al*., 2000; São-José *et al*., 2004). The authors proposed that if endolysins can use host endogenous pathways to reach their substrate, a phage holin-related function would seem expendable. The only requirements for a successful and productive infection would be a delayed expression of the endolysin gene, a low rate of enzyme synthesis or a combined effect of transcriptional and translational regulatory mechanisms (São-José, 2002). However, an intriguing observation is that all phages which synthesize secreted

endolysins or are proposed to do so also appear to encode a holin-like protein. Indeed, fOg44 apparently has a holin gene, *hol44* (Parreira *et al*., 1999), and its co-expression with *lys44* in the heterologous *E. coli* environment leads to more efficient lysis (São-José *et al*., 2000). The ubiquitous presence of holin genes in phages where the endolysins are apparently Sec-targeted to the wall compartment, suggests that also in these cases the lysis clock is regulated by holin-mediated membrane disruption. Membrane damage brought about by the holins would somehow activate the exported endolysins rather than allowing their egress. Maintaining secreted endolysins in a less-than optimal configuration with respect to activity could result either directly from their association with the energized membrane or indirectly by some parameter related to the membrane potential (São-José *et al*., 2000). It was demonstrated that bacterial membrane proton-motive force regulates the lytic activity of the secreted endolysin Lys44 from *O. oeni* fOg44. Lys44 activity is not the result of cell metabolic arrest (leading to loss of viability) and requires permeabilization of the cytoplasmic membrane: cytoplasmic membrane voltage dissipation is necessary but not sufficient to full sensitization of cells to Lys44 (Nascimento *et al*., 2008).

A survey of orthologous endolysins from other phages of Gram-positive bacteria suggested that some, but not all, of these endolysins have N-terminal sequences resembling secretory signals, although in every case an adjacent holin gene orthologous to the putative *hol44* was also present (Young and Wang, 2006). Five proteins showed a notable similarity with the fOg44 lysin sequence in the SP-corresponding region: the lysin of bacteriophage ϕ10MC (also infecting *O. oeni*) and the highly related lytic enzymes from the temperate *Lactococcus lactis* phages Tuc2009, ul36, TP901-1, ϕLC3, ϕAM2 and TPW22 (São-José *et al*., 2000; São-José *et al*., 2004). In addition, the endolysin of *Lactobacillus plantarum* phage ϕg1e had also been reported to be processed in *E. coli* by the Sec machinery (Kakikawa *et al*., 2002).

Lysis systems involving secreted endolysins may therefore uncover a new and more general function for holins: to activate the endolysin, irrespective of the nature of its enzymatic activity or its subcellular localization prior to activation (Fig. 7B) (São-José *et al*., 2000; Young and Wang, 2006).

2.4. Echos in Coliphage: Secreted Endolysins and Pinholins

Sequence comparison also revealed that a number of endolysins from phages of Gram-negative bacteria also have N-terminal sequences that could engage the *sec* system. The length of the hydrophobic N-terminal extension on the predicted endolysins, homologous of T4 E, from coliphage P1 and from the lambdoid phages 21, P22, PS119 and PS3, and the distribution of flanking charged residues were, however, indicative of an N-terminal TMD, a “signal anchor” or “uncleaved signal sequence”, rather than of a cleavable signal peptide (Xu *et al*., 2004; Young and Wang, 2006). A particularly interesting case are the endolysins of phages P1 (Lyz) and 21 (R) that differ significantly from the fOg44 secretory endolysin in not having a secretory signal sequence although they were found to cause lysis of the host without a holin. Instead of requiring a holin, export is mediated by an N-terminal transmembrane domain and requires host *sec* function (Xu *et al*., 2004; Sun *et al*., 2009). The N-terminal domain of Lyz^{P1} is both necessary and sufficient not only for export of this endolysin to the membrane but also for its release into the periplasm. The unusual N-terminal domain, rich in residues that are weakly hydrophobic, functions as a signal-arrest-release (SAR) sequence, which first acts as a normal signal-arrest domain to direct the endolysin to the periplasm in a membrane-tethered form and then allows it to be released as a soluble active enzyme in the periplasm (Xu *et al*., 2004).

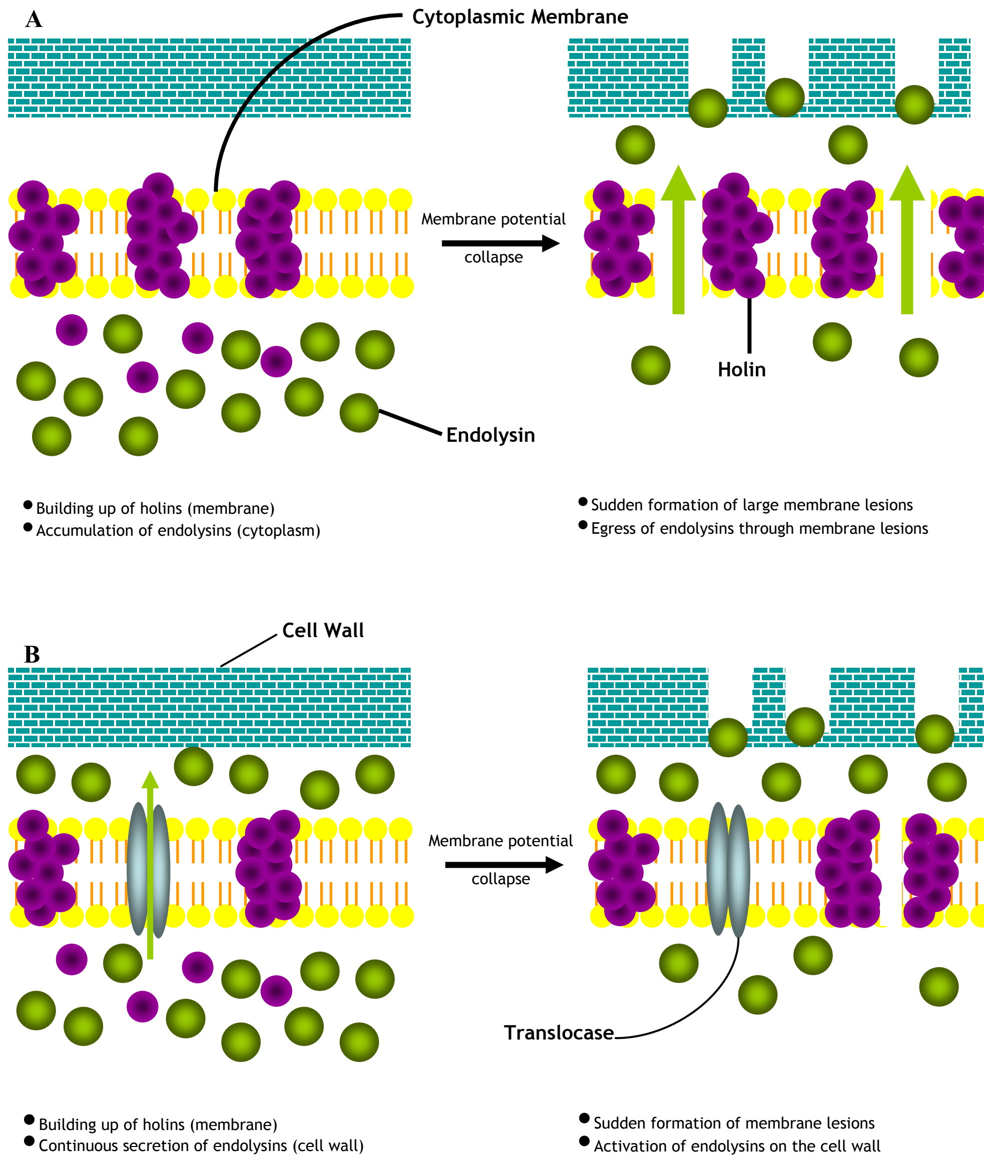


Figure 7. ***Model representation of host-lysis strategies of phages*** producing non-secreted endolysins **(A)** or signal peptide-bearing endolysins **(B)**. See text for further description. Figure adapted from São-José *et al*., 2003.

A model for triggering of lysis with SAR endolysins was proposed in which the SAR endolysin is initially tethered in an inactive form to the energized membrane, in which the holin protein accumulates without affecting the proton-motive force. At the programmed lysis time, the holin triggers, disrupting the membrane sufficiently to abolish the p.m.f., and perhaps also to assist the liberation of the endolysin from the membrane, which results in activation of the endolysin (Xu *et al*., 2004). The molecular details of the activation event were further dissected. When the P1 lysozyme Lyz is secreted to the periplasm of *E. coli* and accumulates in an inactive membrane-tethered form, a disulfide bond is formed between two cysteine residues near the active site. This keeps the enzyme inactive in two ways, first by sterically blocking and distorting the active site cleft and second by keeping the distal cysteine residue, which is required for the sulphydryl state for catalysis, oxidized. When released from the bilayer, Lyz is activated by an intramolecular thiol-disulfide isomerization, which requires a cysteine in its N-terminal SAR domain: the SH side chain of Cys13, located within the SAR sequence attacks the disulfide, resulting in unblocking of the active site and the reduction of the distal Cys residue. Crystal structures confirmed the alternative disulfide linkages in the two forms of Lyz and revealed dramatic conformational differences in the catalytic domain. Thus, the exported P1 endolysin is kept inactive by three levels of control -topological, conformational, and covalent- until its release from the membrane is triggered by the canonical P1 holin (Xu *et al*., 2005): this functional regulation is essential to avoid premature lysis of the infected host. More recently, the regulation of the muralytic enzyme of phage 21 was also dissected. Both genetic and crystallographic studies show that the SAR domain of R^{21}, once extracted from the bilayer, refolds into the body of the enzyme and effects muralytic activation by repositioning one residue of the canonical lysozyme catalytic triad: besides controlling the topology of protein, the SAR domain in R^{21} plays a specific and more integral role in the catalytic activity of the enzyme. Once

synthesized, SAR endolysins require both strict post-secretory negative regulation and the means to become activated in a timely manner, because they are not sequestered from their substrate by the membrane. Inactive Lyz^{P1} has two levels of negative control— covalent inactivation of its active-site cysteine and an N-terminal catalytic domain with radical conformational disability—whereas R^{21} seems to lack only the correct placement of its catalytic glutamate (Sun *et al.*, 2009).

An alternative and remarkably different class of holin-endolysin systems has emerged recently (Park *et al.*, 2006; Park *et al.*, 2007). This class, represented by the lambdoid bacteriophage 21, utilizes endolysins having N-terminal secretory signals SAR domains and pinholins. For phages encoding SAR endolysins, the holin proteins need only to produce lesions large enough to allow the passage of ions and depolarize the cytoplasmic membrane in order to fulfil their role in controlling the timing of lysis. Indeed, unlike lesions formed by the λ holin, lesions formed by the phage 21 holin do not allow the passage of λ endolysin (Park *et al.*, 2007). The term ‘‘pinholin’’ has been proposed to differentiate the small-hole (pinhole) forming character of the phage 21 holin from the canonical holins that form large, nonspecific holes (Park *et al.*, 2007). Like the λ holin *S* gene, the phage 21 holin gene, S^{21} encodes two proteins, $S^{21}71$ and $S^{21}68$, by virtue of alternate translational starts; the shorter product ($S^{21}68$) is required for lysis and the longer gene product ($S^{21}71$), like that of lambda, has inhibitory character of $S^{21}68$, and thus, is an antiholin (Barenboim *et al.*, 1999). Analysis of $S^{21}68$, an allele that produces only the holin, has shown that despite its small size of 68 residues, the product has two functional domains corresponding to the two predicted transmembrane domains (TMDs). TMD_1 is not only dispensable for hole-formation and lysis, but in fact, is itself a SAR domain that must exit the bilayer in order for TMD_2 to be competent for hole-formation. The departure of TMD_1 from the bilayer coincides with the lethal triggering of the holin and is accelerated by membrane

depolarization (Park *et al*., 2007). In addition, TMD_1 acts in *trans* as an inhibitor of the lethal function of TMD_2 (Pang *et al.*, 2010). The structure of the prototype pinholin S^{21} was examined by negative-stain transmission electron-microscopy, cysteine-accessibility, and chemical crosslinking, as well as by computational approaches. Together, the results suggested that the pinholin forms symmetric heptameric structures with the hydrophilic surface of one transmembrane domain lining the surface of a central channel ~15 Å in diameter (Pang *et al*., 2009).

It was suggested that the S^{21}/R^{21} gene pair, encoding a pinholin and a SAR endolysin, may represent an intermediate stage in the evolution of holin-endolysin systems. The most primitive dedicated lysis system could have consisted of a SAR endolysin alone. This would provide a lysis delay because of the gradual release and activation of the membrane-tethered endolysins. In addition, due to their sensitivity to membrane depolarization, the SAR endolysins would provide a sentinel function to effect immediate lysis in the event of any condition which disrupted the integrity of the membrane. However, a lysis system employing a SAR endolysin alone would be inherently inferior to canonical holin-endolysin systems for two reasons. First, because canonical holins function with cytoplasmic endolysins, the muralytic activity elaborated during the infection cycle can be produced in great excess. Not only does this mean that once the holin triggers, host lysis occurs in a matter of seconds, reducing the dwell time in the dead, non-productive host to a minimum, but also it means that lysis timing is completely dependent on the holin. Secondly, it has been shown that most missense changes in holin proteins alter the timing of lysis, unpredictably advancing or retarding the instant of triggering. The canonical holins thus have a selective advantage not only for fitness, in terms of the mechanistic advantages of holin function, but also because they can function with either cytoplasmic endolysins or SAR endolysins,

whereas the pinholin genes can function only with SAR endolysins (Park *et al*., 2007).

3. Mycobacteriophage Ms6 and Its Lysis Operon

3.1. Genome overview and general features

Ms6 is a temperate mycobacteriophage that infects *Mycobacterium smegmatis*, was spontaneously induced from a culture of *M. smegmatis* strain HB5688 and has the ability to form stable lysogens (Portugal *et al*., 1989; Anes *et al*., 1992). Electronic microscopy studies revealed that phage particles are composed by an isometric polyhedral head with 80 nm in diameter, hexagonal in shape, and a long non-contractile tail with 210 nm long. The nucleic acid of the phage is composed by a dsDNA molecule and the length of the genome is about 48 kbp with a GC content of 62%. The morphological characteristics of the phage allowed its classification in the *Siphoviridae* group (Portugal *et al*., 1989; Ackermann, 2006). When the kinetic of adsorption of Ms6 to sensitive, lysogenic and resistant strains was examined, the phage was shown to adsorb to the lysogenic but not to the resistant strain. This data indicates that bacteria surviving an infection by Ms6 are immune to a superinfection (Portugal *et al*., 1989) and the Ms6 Pin protein seems to be involved in the superinfection exclusion of the phage. It was demonstrated that expression of this protein does not prevent adsorption but blocks some early step following infection, at a stage between adsorption of the phage and DNA injection and according to the known phage-encoded resistance mechanisms, the protein Pin acts by a superinfection exclusion mechanism (Pimentel, 1999).

Recently, Hatfull and collaborators (2010) grouped 60 mycobacteriophage genomes into clusters by gene content comparison and the putative gene products of these phages were grouped into “phamilies” of related sequences. The mycobacteriophage Ms6 lysis module is closely related to those of phages

belonging to cluster F, subcluster F1 that includes phages PMC, Llij, Che8, Boomer, Fruitloop, Pacc40, Ramsey and Tweety. Genomes of subcluster F1 phages present a high diversity, with the leftmost regions of the genomes containing the virion structure and assembly genes being better conserved that the rightmost genomic segments (Hatfull *et al*., 2010).

3.2. Integration

In 1998, the genetic elements involved in the site-specific integration event between bacteriophage Ms6 DNA and the mycobacterial genome required for chromosomal integration were characterized (Freitas-Vieira *et al*., 1998). DNA sequence analysis of an *attP* containing fragment revealed an ORF encoding a protein of 372 amino acid residues with a C-terminus similar to other conserved C-terminal regions typical of the phage tyrosine-integrase family. Comparison of the sequences of *attP, attB* and bacteria-prophage junctions *attL* and *attR* showed a 26 bp common core sequence, where recombination takes place, near the 5' end of the integrase gene. Nucleotide sequence analysis of the *attB* chromosomal region showed that the core site overlaps the 3' end of the $tRNA^{Ala}$ gene and seems to be conserved in fast- and slow-growing mycobacteria. An integration-proficient plasmid vector was constructed, efficiently inserted at the $tRNA^{Ala}$ genes of *M. smegmatis, M. vaccae, M. bovis* BCG and *M. tuberculosis* H37Ra (Freitas-Vieira *et al*., 1998) and allowed the high and long-lived expression of heterologous genes [the early regulatory *nef* and the structural *gag* (p26) genes from the simian immunodeficiency virus (SIV)] in recombinant BCG strains (Méderlé *et al*., 2002).

A BLASTn search of the *M. smegmatis* genome for sequences similar to the Ms6 *attP* core region revealed three possible *attB* sites at the 3′ ends of three $tRNA^{Ala}$ genes: $tRNA^{AlaU}$, $tRNA^{AlaV}$, $tRNA^{AlaT.}$ The three $tRNA^{Ala}$ genes of *M. smegmatis* and BCG are highly similar and conserved, but not identical. Small changes in the 7 bp T-loop *attP* core of Ms6 to give a perfect match with the other

$tRNA^{Ala}$ T-loops allowed the integration of multiple vectors into three different $tRNA^{Ala}$ genes albeit with a lower integration efficiency, which identifies this spot as a possible site of strand exchange (Catalão, unpublished results; Vultos *et al*., 2006). Furthermore, construction of two Ms6 vectors to integrate two reporter genes into the $tRNA^{AlaU}$ and $tRNA^{AlaV}$ T-loops of the same BCG chromosome provided the basis for the development of recombinant *M. bovis* strains expressing several reporter genes inserted into different $tRNA^{Ala}$ genes (Vultos *et al*., 2006).

3.3. Lysis

The genetic organization and some transcriptional control elements of the mycobacteriophage Ms6 lysis functions were described in 2002 (Garcia *et al*., 2002). A strong promoter region (P_{lys}) was identified and isolated by using transcriptional fusions with the *lac*Z reporter gene. Two tandem σ^{70}-like promoter sequences (P1 and P2) that are recognised by the host RNA polymerase were found in this region by genetic analysis. Transcription of the lysis genes is dependent on the P_{lys} promoter located about 6 kb away from the integration locus (Fig. 8A), which is positioned in the middle of the Ms6 DNA genome (Garcia, 2001; Garcia *et al*., 2002). The transcription initiated from this promoter goes towards the integration elements and the mRNA transcript can not be longer than 5.5 kb. DNA sequencing of the P_{lys} downstream region revealed a 214 bp leader sequence in which the first intrinsic transcription termination signal reported in mycobacteriophages was identified. These data suggested that an antitermination mechanism might be involved in the regulation of Ms6 lysis genes transcription and that during the early stages of the lytic cycle, an antiterminator factor should be synthesized in order to allow transcription to proceed beyond this termination signal (Garcia *et al*., 2002).

The leader sequence is followed by five adjacent genes of 231 bp (*gp1*), 1152 bp (*gp2*), 996 bp (*gp3*), 231 bp (*gp4*) and 372 bp (*gp5*) encoding for proteins involved in cell lysis (Fig. 8B) (Garcia, 2001; Garcia *et al*., 2002).

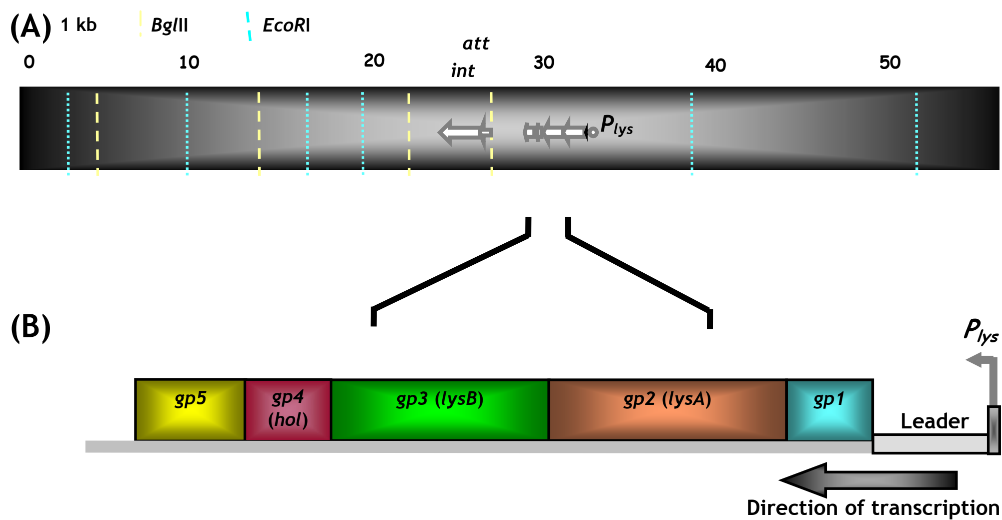


Figure 8. (A) Schematic representation of P_{lys} on the 57-kb Ms6 DNA. **(B)** Genetic organization of the Ms6 lysis *locus*. Figure adapted from Garcia *et al*., 2002.

gp1, of 231 bp is separated from the +1 position by the 214 bp leader sequence and has the potential to encode a 77-amino acid polypeptide (Gp1). A BLASTp search using the Ms6 Gp1 deduced amino acid sequence identified a number of mycobacteriophage putative proteins with a high degree of sequence identity whose functions are currently unknown. Several genes encoding these putative proteins are localized within the lysis cassette of mycobacteriophages genomes upstream of the *gp2* (*lysA*) gene and, they have been recently grouped in a mycobacteriophage gene family Pham1480 representing 32 members (Hatfull *et al*., 2010). This gene family was identified in phages belonging to subcluster A1 (Bethlehem, Bxb1, U2, DD5, Jasper, KBG, Lockley and Solon), phages included in

cluster E (244, Cjw1, Kostya and Porky), cluster F, subcluster F1 (Boomer, PMC, Llij, Che8, Tweety, Fruitloop, Ramsey and Pacc40) that infect *M. smegmatis*, in phages Che12, Pukovnik, L5 and Bxz2 of subcluster A2 (except in D29) and in the singleton phage TM4 that efficiently infect *M. tuberculosis* (Hatfull *et al*., 2010). Recently it was also identified in Ardmore a mycobacteriophage isolated from a soil sample in Ireland that also infects *M. smegmatis* (Henry *et al*., 2010). The Phams present in only a subset of the genomes correspond to genes that are in greatest evolutionary flux. The lack of full representation could result from loss of a gene from one or more genomes, or alternatively, from recent acquisition by genetic horizontal genetic exchange (Hatfull *et al*., 2010)

The highest identity (96%) was observed with the predicted amino acid sequence of Gp28 from mycobacteriophage Fruitloop, but it also produced significant alignments with Gp29 of phage Llij, Tweety and Pacc40, Gp30 of phage Ramsey and Boomer and Gp31 of phage Che8 (78-94%). Sequence identity with gene family Pham1480 of the subcluster A1 was significantly lower (45-54%) as well as with phages of the subcluster A2, cluster E and TM4. Furthermore, while in some phages of subcluster A1 (Bethlehem, Bxb1, Jasper, Lockley and U2), of subcluster F1 (Che8, Tweety, PMC, Llij, Fruitloop, Pacc40) and in phages Pukovnik and Bxz2 of subcluster A2 this gene family is overlapped with the predicted lysin gene, in phages Che12 and TM4 and in phages included in cluster E it is apart from the lysin gene and distant in the genome. In phages DD5, KBG and Solon (subcluster A1) and in phages Boomer and Ramsey (subcluster F1) Pham1480 is separated from the lysin gene by one intervening gene that code for putative homing endonuclease HNH motifs (Hatfull *et al*., 2010)

***gp2* (*lysA*)** of 1152 bp starts at a GTG codon that overlaps the *gp1* TGA stop codon, in a different reading frame. It encodes a 384 amino acid polypeptide and revealed significant similarity to mureinolytic enzymes (lysins) of several bacteriophages with *N*-acetyl-muramoyl-L-alanine amidase activity and was

designated *lysA* to indicate that it represents the Ms6 lysin gene (Garcia *et al*., 2002). Different types of cell wall hydrolases seem to be produced by mycobacteriophages. While phages Ms6 and TM4 encode an atypical enzyme with an “Amid-2” type domain, others such as D29 and Bxb1 employ hydrolases with lysozyme-like activity to bring about host cell lysis. Ms6 LysA holds a peptidoglycan recognition protein (PGRP) super-family conserved domain (cl02712) with its 3 sites: amidase catalytic site, a Zn binding site and a substrate binding site. PGRPs are pattern recognition receptors that bind, and in certain cases, hydrolyze peptidoglycan of bacterial cell walls. This family includes Zn-dependent *N*-acetylmuramoyl-L-alanine amidases (EC:3.5.1.28), which cleave the amide bond between *N*-acetylmuramoyl and L-amino acids, preferentially D-lactyl-L-Ala, in bacterial cell walls. The purified protein was shown to cleave the bond between L-Ala and D-muramic acid and to release up to 70% of the diaminopimelic acid present in isolated mycobacterial cell walls which confirmed the amidase activity of the enzyme (Piechota *et al*., unpublished results). Peptidoglycan hydrolysis was also demonstrated for other three mycobacteriophage LysA proteins (gp32 Che8, gp236 Bxz1 and gp69 Corndog) in zymograms (Payne *et al*., 2009) and LysA of mycobacteriophage Giles is essential for lytic growth (Marinelli *et al*., 2008).

All of the mycobacteriophages sequenced so far encode a putative *lysA*-like gene that have been recently grouped in mycobacteriophage gene family Pham66-1 (Payne *et al*., 2009; Hatfull *et al*., 2010) representing 67 members, although this group of enzymes is not restricted to phages that infect mycobacteria (Hatfull *et al*., 2010). The Lysin A (LysA) family of proteins appears to be a particularly highly diverse and interesting group of lytic enzymes composed of subgenic modules with reasonably defined boundaries (Hatfull *et al*., 2006) and containing amidase, glycosidase or peptidase motifs (Payne *et al*., 2009). The extensive sequence divergence between these and other endolysins may reflect the peculiar

composition and structural features of the mycobacterial cell wall. Remarkably, some mycobacteriophages also encode a second cell envelope lytic enzyme Lysin B (LysB) which does not resemble endolysins and has no identifiable homologues in bacteriophages other than those that infect mycobacterial hosts (Hatfull *et al.*, 2010) with mycolic acid outer membranes (Payne *et al.*, 2009).

***gp3* (*lysB*)** with 996 bp, starts at an ATG codon that overlaps, in a different reading frame, the TGA stop codon of *gp2* and encodes a 332-amino acid protein. Similarity to proteins with cutinase-like (serine esterase) domains (São-José *et al.*, 2003) is consistent with a role in cell lysis and indicates that mycobacteriophages have evolved additional lytic functions to allow degradation of the cutin-like layer that covers the peptidoglycan/arabinogalactan-rich mycobacterial cell wall (São-José *et al.*, 2003). Analysis of the deduced amino acid sequence of Ms6 LysB revealed the presence of a conserved pentapeptide motif (Gly-Tyr-Ser-Gln-Gly) which matches the Gly-X-Ser-X-Gly characteristic of lipolytic enzymes. Further characterization of the gene product of *lysB* demonstrated its lipolytic activity against a wide range of chain length substrates (C_4-C_{18}), although it showed a higher affinity for ρ-nitrophenyl esters of longer chain length (C_{16} and C_{18}) (Gil *et al.*, 2008). Ms6 LysB was shown to hydrolyze the mycolic acids from the mycolyl-arabinogalactan-peptidoglycan complex where the mycolates of the inner leaflet of the outer membrane are covalently attached to an arabinosyl head group. In addition, LysB treatment of the extractable lipids from *Mycobacterium smegmatis*, *Mycobacterium bovis* BCG and *M. tuberculosis* H37Ra, showed that TDM, a trehalose diester of two mycolic acid molecules that is generally believed to be localized on the outer membrane (perhaps mostly on the outer leaflet), was hydrolyzed by the enzyme (Gil *et al.*, 2010). Even though LysB is not essential for lysis as indicated by viability of a mycobacteriophage Ms6 Δ*lysB* mutant, is important for the normal timing of host cell lysis (Gil F., unpublished results).

Recently, Payne *et al*. (2009) proposed that mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase that cleaves the mycolylarabinogalactan bond to release free mycolic acids and although is not essential to accomplish cell lysis, is necessary for the normal timing, progression and completion of host cell lysis. These authors also suggested that LysB acts at a late stage of infection to facilitate lysis by compromising the integrity of the mycobacterial outer membrane linkage to the arabinogalactan-peptidoglycan layer and a model for mycobacteriophage host cell lysis was proposed (Fig. 9).

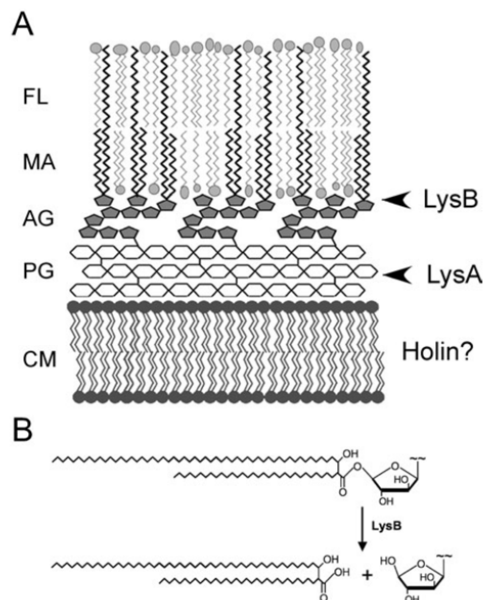


Fig. 9. ***A model for mycobacteriophage lysis of mycobacteria.***
A. Mycobacterial cell walls are unusual in that the cytoplasmic membrane (CM) is surrounded by a peptidoglycan layer (PG) to which a network of arabinogalactan (AG) is covalently attached. A mycobacterial outer membrane consisting of 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 or by other unknown mechanism– 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.
B. Diagram illustrating LysB cleavage of the ester bond linking mycolic acids and arabinogalactan. Figure from Payne *et al*., 2009.

Interestingly, homologues of this protein are found adjacent, or very close to *lysA* in 56 of the 60 completely sequenced mycobacteriophage genomes, are located downstream of *lysA* and separated from it by no more than four intervening genes. Some of the intervening genes encode putative holins and exhibit holin-like function (e. g. D29 gene *11*), whereas others (e. g. Omega gene *51*) code for putative homing endonuclease HNH motifs, or have homologues elsewhere in

other mycobacteriophage genomes (e.g. Troll4 gene *37*) (Payne *et al*., 2009). The *lysB* genes have been grouped in gene family Pham67 with 56 members (Hatfull *et al*., 2010). Sequence alignment of the family of LysB proteins shows that they are highly diverse, and only three residues are completely conserved. Although the proteins vary in length and there are many gaps throughout the alignment, these proteins no not have modular constructions as seen in the LysA proteins (Payne *et al*., 2009).

gp4 **(*hol*)** begins with an ATG codon located 10 nucleotides downstream of the *lysB* stop codon and encodes a 77-amino acid polypeptide related to the holin protein of *Lactococcus lactis* r1t bacteriophage (Garcia *et al*., 2002). It shares some structural characteristics with class II holins, which are usual hydrophobic in nature and small in size, with a highly hydrophilic carboxy-terminal domain and two potential transmembrane domains and was designated *hol* to indicate that it represents the Ms6 holin gene. Holin function was demonstrated by its ability to complement a λ phage S mutant and the lethal phenotype when overexpressed in *E. coli*, explained by the introduction of non-specific lesions in the cytoplasmic membrane. However, unlike some holins as the lambda S holin it lacks a dual start motif (Garcia, 2001; Garcia *et al*., 2002). According to the sequence similarity to putative holin genes from mycobacteriophages of subcluster F1, the Ms6 holin was included in Pham95. Putative holin genes were also identified in mycobacteriophages from clusters A2 (*gp11* from phages D29 and L5 and *gp12* from phages Che12 and Pukovnik), E (*gp33* from phages Cjw1 and Porky, *gp34* from phage Kostya and *gp35* from phage Porky) and I (*gp27* from phage Che9c and *gp31* from phage Brujita), all belonging to Pham1981 (Hatfull *et al*., 2010). It is interesting to note that in mycobacteriophage genomes where a holin gene has been assigned (including Ms6) it is closely linked to *lysA* in an endolysin-holin gene organization, contrary to what is usually found in the majority of the dsDNA bacteriophages (holin-endolysin), with the holin gene mapping downstream and not

upstream of the lysin gene. At least three other phages appear to have an identical organization, the *Bacillus subtilis* phage SPP1, the *Oenococcus* phage ϕ10MC and fOg44 (São-José *et al.*, 2003) and also some mycobacteriophages. However, additional holin genes have yet to be identified bioinformatically or by functional studies as in the majority of the mycobacteriophage genomes sequenced there is no closely linked and easily recognizable holin gene (Payne *et al.*, 2009; Hatfull *et al.*, 2010).

Finally, ***gp5*** with 372 bp starts at an ATG codon that overlaps the *gp4* TGA stop codon in a different reading frame. It has the potential to encode a 124-amino acid protein and exhibits similarity exclusively with other mycobacteriophages proteins with unknown functions (Garcia, 2001, Garcia *et al.*, 2002). BLASTp search using the Ms6 Gp5 deduced amino acid sequence identified a number of putative proteins restricted to the mycobacteriophage with a high degree of sequence identity. Owing to their related sequences, genes encoding these proteins have been recently grouped in a mycobacteriophage gene family, Pham96, representing 15 members (Hatfull *et al.*, 2010). This gene family was identified in phages belonging to subcluster A2 (Che12, Pukovnik, L5, Bxz2 and D29), subcluster F1 (Boomer, PMC, Llij, Che8, Tweety, Fruitloop, Ramsey and Pacc40), cluster I (Brujita) and in the singleton phage Omega (Hatfull *et al.*, 2010).

Although mycobacteria are considered more similar to Gram-positive bacteria, mycobacteriophages have to overcome the complexity of the mycobacterial outer membrane a second barrier analogous to the outer membrane of Gram-negative bacteria (Fig. 10) (Hoffmann *et al.*, 2008; Zuber *et al.*, 2008). Phages that infect Gram-negative bacteria have the help of spanin proteins and their Rz/Rz1 counterparts to span the periplasmic space and link the inner and outer membrane, enhancing the completion of lysis (Summer *et al.*, 2007; Berry *et al.*, 2008; Krupovič *et al.*, 2008).

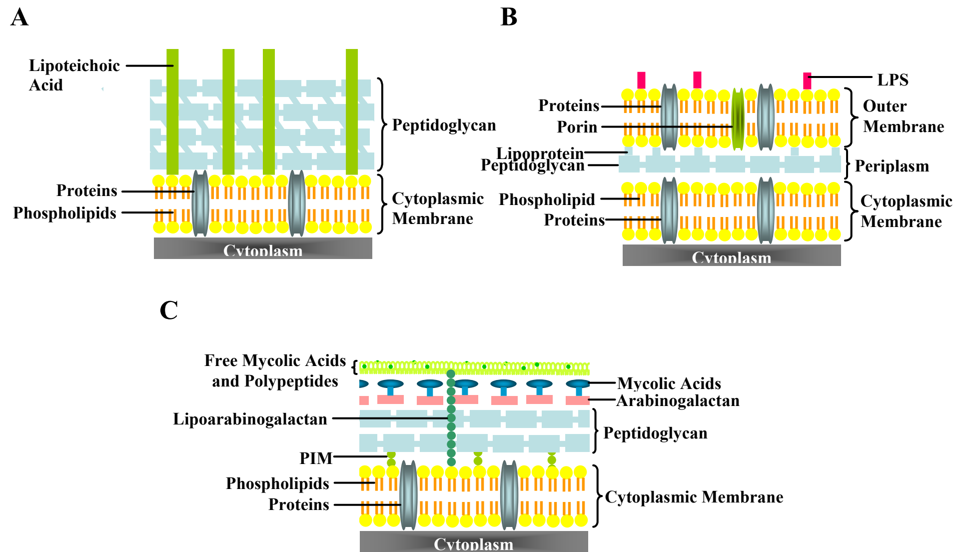


Figure 10. ***Cell envelopes of bacteria.*** **(A)** Gram-positive bacteria; **(B)** Gram-negative bacteria; **(C)** Mycobacteria. Abbreviations: PIM, fosfatidilmio-inositol manoside; LPS, lipopolysaccharide. Adapted from Nikaido *et al*., 1994.

Although not mandatory for lysis, acquisition of additional lysis genes by mycobacteriophages throughout their evolution, likely confers a substantial selective advantage over those without it by providing faster and more complete lysis: mycobacteriophages have developed an alternative solution to compromise the mycobacterial outer membrane by means of LysB activity and others that are yet to be described. The products encoded by *gp1*, *lysB* and *gp5* of Ms6 are then mycobacterium-specific lysis factors restricted to mycobacteriophages (Garcia *et al*., 2002; Payne *et al*., 2009; Hatfull *et al*., 2010) that may confer host lysis benefits.

The Ms6 lysis genes are expressed by their own promoter region (P_{lys}) and are not expressed as a single operon with the virion structure and assembly protein genes as it happens in bacteriophage λ. In this phage, the lysis genes are the first

genes of the late operon clustered with the structural proteins of the phage particle and their transcription is initiated from the same strong promoter, P_R'. The separation of lysis and structural genes in two different operons indicates more complex transcriptional control of the late functions in bacteriophage Ms6 (Garcia *et al*., 2002). At this time, it is unknown how LysA and LysB are localized to their substrates, the peptidoglycan and the mycobacterial outer membrane, as no signal peptide motifs have been identified in these enzymes (Gil *et al*., 2008; Payne *et al*, 2009). Even though it has not yet been described, mycobacteriophage endolysins could contain SAR domains or may employ other mechanisms that allow secretion of the endolysin into the periplasmic space thus rendering them holin-independent (Xu *et al*., 2005; Payne *et al*., 2009). Identification and characterization of the mycobacteriophage lysis effectors and the exact pathways by which they are positioned next to their targets is of interest not only for understanding the mechanisms and timing of lysis but also because of the possible therapeutic use of mycobacteriophage lysins; exogenously applied phage-encoded endolysins have been shown to have effective antimicrobial activity (Fischetti, 2008) against Gram-positive bacterial pathogens including *Streptococcus pneumoniae* (Loeffler *et al*., 2001; Loeffler and Fischetti, 2003; Loeffler *et al*., 2003) and *Bacillus anthracis* (Schuch *et al*., 2002).

4. Bacteriophage Therapy

4.1. Mycobacteriophage and Their Application to Disease Control

The application of bacteriophage to treat non-mycobacterial diseases (Fig. 11) has been extensively reviewed (Summers, 2001; Sulakvelidze *et al*., 2001; Matsuzaki *et al*., 2005; Hanlon, 2007; O′Flaherty *et al*., 2009). Phage therapy is currently used in Eastern Europe and countries of the former Soviet Union where it has been applied to infections caused by *Staphylococcus*, *Pseudomonas*, *Escherichia coli*, *Klebsiella* and *Salmonella* (O′Flaherty *et al*., 2009). The

mycobacteriophage therapy, however, have received less attention. Nevertheless, the resurgence of mycobacterial disease has prompted fresh appraisal of their potential as therapeutic tools. Mycobacterial infections are difficult to treat; they are naturally resistant to many antibiotics and require multiple drug therapy for extended periods of time (O'Brien and Nunn, 2001). The minimum treatment for tuberculosis disease is 6 months; 2 months with four drugs, followed by 4 months of two drugs (Frieden *et al.*, 2003). Failure to maintain adequate therapy may result in the development of drug resistance. Strains of multi-drug resistant tuberculosis (MDR-TB) that can not be cured by standard treatment regimes are a serious threat to control the disease (Nachega and Chaisson, 2003). The difficulty of treating mycobacterial disease and the emergence of drug resistance has encouraged scientists to investigate whether mycobacteriophage could provide a complementary means of therapy. Early attempts to treat laboratory animals infected with *M. bovis* BCG and *M. tuberculosis* were not successful (McNerney, 1999; McNerney and Traoré, 2005). Killing intracellular pathogens such as *M. tuberculosis* presents a tough challenge as in order to infect the target bacilli the phage need to transverse the mammalian cell membrane and survive in adverse intra-cellular environments such as reduced pH. Novel phage delivery systems are required and a possible strategy has recently been identified where bacteriophage are transported into macrophages via nonvirulent-carrier-bacteria (Broxmeyer *et al.*, 2002; Danelishvili *et al.*, 2006). Macrophage cell lines infected with *M. tuberculosis* or *M. avium* were treated by the addition of *M. smegmatis* infected with mycobacteriophage TM4. Following ingestion and destruction of the *M. smegmatis* bacilli, TM4 phages were released within the macrophage. A significant reduction in infection was observed in both the *M. tuberculosis* and *M. avium* experiments suggesting that TM4 bacteriophage had successfully infected and destroyed pathogenic bacteria within the macrophage (Broxmeyer *et al.*, 2002). Another study showed that infected-mycobacteria macrophages could phagocyte

phage D29. D29 was effective in lysis of intracellular mycobacteria without any deliver system (Peng *et al*., 2006) contrary to phage TM4 that can not kill intracellular mycobacteria alone due to incapacity of coming into macrophages (Broxmeyer *et al*., 2002).

A second challenge to successful phage therapy is the presence of granulomas which might prove impenetrable to the bacteriophage. These are often observed in mycobacterial diseases and might prevent complete clearance of the bacteria (Bowman *et al*., 1972). Phage therapy might be more readily applicable to less visceral mycobacterial diseases such as *M. marinum* or Buruli ulcer where the site of infection is accessible. However, no such studies have so far been reported. Further research is needed to overcome the existing technical barriers and enable the development of effective therapeutic tools. The emergence of drug resistant disease as a serious public health problem has sparked considerable scientific and public interest in this area of research but a role for mycobacteriophage in the treatment of tuberculosis and other mycobacterial infections has yet to be established (McNerney and Traoré, 2005).

4.2. Application of Phage Lysis Proteins as Therapeutics

The increased research into the utilization of phage lysins as therapeutics is evident. As these enzymes break down the cell wall, they have the potential to be used as therapeutic agents in their own right (Fig. 11). The continuing emergence of phage genome sequences enables the putative identification of many lysins. Fischetti and colleagues have exploited these enzymes, which they have termed ‘enzybiotics’, to kill a variety of Gram-positive pathogens and reported the prophylactic use of a phage lysin in an *in vivo* model. In this study, phage lysin encoded by the C1 phage was utilized, which is specific for groups A, C and E streptococci (Nelson *et al*., 2001). The use of phage lysins to control *S. pneumoniae* was also studied by Fischetti’s group (Loeffler *et al*., 2001; Loeffler and Fischetti,

2003). These included the use of the purified lysin Pal, which is active against 15 common serotypes of pneumococci In a second study, lysin Cpl-1 was also shown to be effective against *S. pneumoniae* in a mouse model of infection as a topical application and when injected into the bloodstream (Loeffler *et al*., 2003). A combination of Pal and Cpl-1 lysins resulted in an increased killing effect *in vitro* against *S. pneumoniae* (Loeffler and Fischetti, 2003). Phage lysin has also been utilized for the detection and elimination of *Bacillus anthracis* (Schuch *et al*., 2002) and *S. aureus* (Rashel *et al*., 2007).

Exogenously applied phage-encoded endolysins have been shown to possess effective antimicrobial activity against Gram-positive bacterial pathogens, as described above. However, they are ineffective against Gram-negative bacteria because the outer membrane blocks access to the peptidoglycan targets. The mycobacteria are likely to be similar intractable to exogenously added endolysins because of their mycolic-acid-rich outer membrane (Payne *et al*., 2009) and the complexity of the cell wall; enzymes that attack the mycobacterial cell envelope from the outside are needed as a novel therapeutic method to kill mycobacteria. Lysin B is able to hydrolyze the mycolate-arabinogalactan and the mycolate-trehalose bonds (Payne *et al*., 2009; Gil *et al*., 2010). However, when Ms6 LysB was added from without it was unable to disrupt the outer membrane sufficiently to kill the mycobacteria (Gil *et al*., 2010).

Importantly, lytic enzymes originating from phages have huge potential from a therapeutic perspective as these enzymes show no adverse reactions during *in vivo* trials (Jado *et al*., 2003) and no resistance to them has been discovered (Loeffler *et al*., 2001; Nelson *et al*., 2001; Schuch *et al*., 2002). Another important feature of phage lysins is the capability to produce engineered lysins by domain swapping to obtain lytic enzymes with multiple lytic activities and/or multiple

binding domains, which have the potential to increase the therapeutic potential of phage lytic enzymes (O´Flaherty *et al*., 2009).

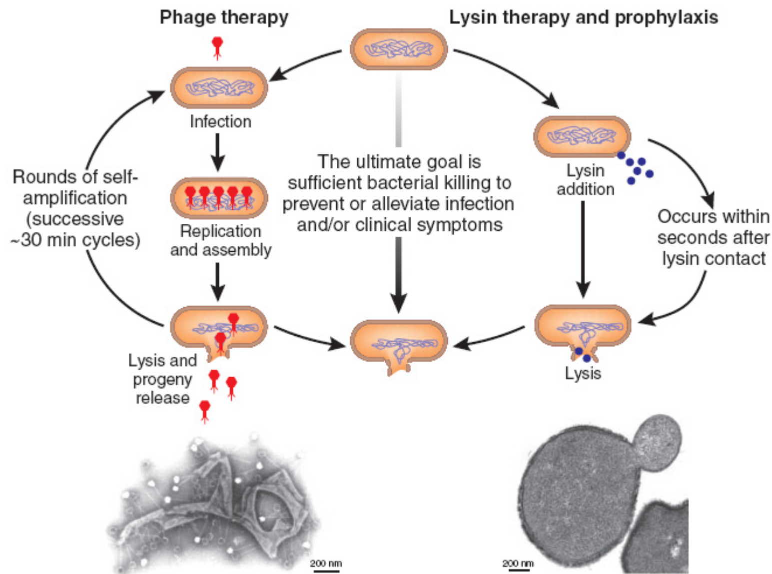


Figure 11. *Steps to bacterial lysis in phage and lysin therapy.* Phage therapy (on the left) exploits a natural phage lytic cycle, which occurs over 30 min and is divided into three major steps, including the release of new virions (in red) into the environment. Subsequent infection of new hosts illustrates the process of self-amplification. The electron micrograph depicts phage particles adhering to the debris of a lysed streptococcal cell. In comparison, lysin therapy and prophylaxis (on the right) is defined by only two steps, in which purified lysin (in blue) binds to, and rapidly kills through osmotic lysis, the target pathogen. The electron micrograph depicts a cross-section of *Bacillus anthracis* treated with the purified PlyG lysin showing an externalized cytoplasmic membrane just before lysis. Figure from Fischetti *et al*., 2006.

References

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

Ackermann H.-W. (2005) Bacteriophage classification. *In*, Bacteriophages: biology and applications. Kutter E., Sulakvelidze A. (Ed.) Boca Raton, FL: CRC Press, p. 67-89.

Ackermann H.-W. (2006) Classification of bacteriophages. *In*, The Bacteriophages. R. Calendar (Ed.), 2nd ed. Oxford University Press, New York, p. 8-16.

Anes E., Portugal I. and Moniz-Pereira J. (1992) Insertion into the *Mycobacterium smegmatis* genome of the *aph* gene through lysogenization with the temperate mycobacteriophage Ms6. *FEMS Microbiol Lett* **74(1):**21-25.

Arisaka F., Kanamaru S., Leiman P. and Rossmann M. G. (2003) The tail lysozyme complex of bacteriophage T4. *Int J Biochem Cell Biol* **35(1):**16-21.

Balasubramanian V., Pavelka M. S. Jr, Bardarov S. S., Martin J., Weisbrod T. R., McAdam R. A., Bloom B. R. and Jacobs W. R. Jr. (1996) Allelic exchange in *Mycobacterium tuberculosis* with long linear recombination substrates. *J Bacteriol* **178:**273–279.

Bardarov S., Bardarov S. Jr., Pavelka M. S. Jr., Sambandamurthy V., Larsen M., Tufariello J., Chan J., Hatfull G. and Jacobs W. R. Jr. (2002) Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG and *M. smegmatis*. *Microbiology* **148:**3007–3017.

Barenboim M., Chang C. Y., dib Hajj F. and Young R. (1999) Characterization of the dual-start motif of a class II holin gene. *Mol Microbiol* **32(4):**715-727.

Bartel P. L., Roecklein J. A., SenGupta D. and Fields S. (1996) A protein linkage map of *Escherichia coli* bacteriophage T7. *Nat Genet* **12(1):**72-77.

Bernhardt T. G., Roof W. D. and Young R. (2000) Genetic evidence that the bacteriophage ϕX174 lysis protein inhibits cell wall synthesis. *Proc Natl Acad Sci USA* **97(8):**4297-4302.

Bernhardt T. G., Wang I.-N., Struck D. K. and Young R. (2001a) A protein antibiotic in the phage Qβ virion: diversity in lysis targets. *Science* **292(5525):**2326-2329.

Bernhardt T. G., Struck D. K. and Young R. (2001b) The lysis protein E of ϕX174 is a specific inhibitor of the MraY-catalyzed step in peptidoglycan synthesis. *J Biol Chem* **276(9):**6093-6097.

Bernhardt T. G., Wang I-N., Struck D. K. and Young R. (2002a) Breaking free: “protein antibiotics” and phage lysis. *Research Microbiol* **153:**493-501.

Bernhardt T. G., Roof W. D. and Young R. (2002b) The *Escherichia coli* FKBP-type PPIase SlyD is required for the stabilization of the E lysis protein of bacteriophage ϕX174. *Mol Microbiol* **45(1):**99-108.

Berry J., Summer E. J., Struck D. K. and Young R. (2008) The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. *Mol Microbiol* **70(2):**341-351.

Beveridge T. J. (1999) Structures of Gram-negative cell walls and their derived membrane vesicles. *J Bacteriol* **181(16):**4725-4733.

Beveridge T. J. (2000) Ultrastructure of Gram-positive cell walls. *In*, Gram-Positive Pathogens. V. A. Fischetti (Ed.) Washington D. C., ASM Press, p. 3-10.

Bieńkowska-Szewczyk K. and Taylor A. (1980) Murein transglycosylase from phage lambda lysate. Purification and properties. *Biochim Biophys Acta* **615(2):**489-496.

Bieńkowska-Szewczyk K., Lipinska B. and Taylor A. (1981) The *R* gene product of bacteriophage lambda is the murein transglycosylase. *Mol Gen Genet* **184(1):**111-114.

Bläsi U., Nam K., Hartz D., Gold L. and Young R. (1989) Dual translational initiation sites control function of the lambda S gene. *EMBO J* **8(11):**3501-3510.

Bläsi U., Chang C. Y., Zagotta M. T., Nam K. B. and Young R. (1990) The lethal lambda S gene encodes its own inhibitor. *EMBO J* **9(4):**981-989.

Bläsi U. and Young R. (1996) Two beginnings for a single purpose: the dual-start holins in the regulation of phage lysis. *Mol Microbiol* **21(4):**675-682.

Bläsi U., Fraisl P., Chang C. Y., Zhang N. and Young R. (1999) The C-terminal sequence of the lambda holin constitutes a cytoplasmic regulatory domain. *J Bacteriol* **181(9):**2922-2929.

Borysowski J., Weber-Dabrowska B. and Gorski A. (2006) Bacteriophage endolysins as a novel class of antibacterial agents. *Exp Biol Med* (Maywood) **231(4):**366-377.

Bowman B. U., Amos W. T. and Geer J. C. (1972) Failure to produce experimental sarcoidosis in guinea pigs with *Mycobacterium tuberculosis* and mycobacteriophage DS6A. *Am Rev Respir Dis* **105(1):**85-94.

Boyd E. F., Davis B. M. and Hochhut B. (2001) Bacteriophage-bacteriophage interactions in the evolution of pathogenic bacteria. *Trends Microbiol* **9(3):**137-144.

Boyd E. F. and Brüssow H. (2002) Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol* **10(11):**521-529.

Broxmeyer L., Sosnowska D., Miltner E., Chacón O., Wagner D., McGarvey J., Barletta R. G. and Bermudez L. E. (2002) Killing of *Mycobacterium avium* and *Mycobacterium tuberculosis* by a mycobacteriophage delivered by a nonvirulent mycobacterium: a model for phage therapy of intracellular bacterial pathogens. *J Infect Dis* **186:**1155-1160.

Brüssow H., Canchaya C. and Hardt W.-D. (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* **68(3):**560-602.

Brüssow H. and Kutter E. (2005) Phage ecology. *In*, Bacteriophages: biology and applications. Kutter E., Sulakvelidze A. (Ed.) Boca Raton, FL: CRC Press, p. 129-163.

Canchaya C., Fournous G., Chibani-Chennoufi S., Dillmann M.-L. and Brüssow H. (2003) Phage as agents of lateral gene transfer. *Curr Opin Microbiol* **6:**417-424.

Chang C. Y., Nam K. and Young R. (1995) *S* gene expression and the timing of lysis by bacteriophage lambda. *J Bacteriol* **177(11):**3283-3294.

Cheng X., Zhang X., Pflugrath J. W. and Studier F. W. (1994) The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proc Natl Acad Sci USA* **91(9):**4034-4038.

Cheng Q., Nelson D., Zhu S. and Fischetti V. A. (2005) Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. *Antimicrob Agents Chemother* **49(1):**111-117.

Costerton J. W., Ingram J. M. and Cheng K. J. (1974) Structure and function of the cell envelope of Gram-negative bacteria. *Bacteriol Rev* **38(1):**87-110.

Danelishvili L., Young L. S. and Bermudez L. E. (2006) *In vivo* efficacy of phage therapy for *Mycobacterium avium* infection as delivered by a nonvirulent mycobacterium. *Microb Drug Resist* **12(1):**1-6.

Dewey J. S., Savva C. G., White R. L., Vitha S., Holzenburg A. and Young R. (2010) Micron-scale holes terminate the phage infection cycle. *Proc Natl Acad Sci USA* **107(5):**2219-2223.

Fischetti V. A., Nelson D. and Schuch R. (2006) Reinventing phage therapy: are the parts greater than the sum? *Nat Biotechnol* **24(12):**1508-1511.

Fischetti V. A. (2008) Bacteriophage lysins as effective antibacterials. *Curr Opin Microbiol* **11(5):**393-400.

Freitas-Vieira A., Anes E. and Moniz-Pereira J. (1998) The site-specific recombination *locus* of mycobacteriophage Ms6 determines DNA integration at the tRNA(Ala) gene of *Mycobacterium* spp. *Microbiology* **144(Pt 12):**3397-3406.

Frieden T. R., Sterling T. R., Munsiff S. S., Watt C. J. and Dye C. (2003) Tuberculosis. *Lancet* **362(9387):**887-899.

Garcia M. (2001) Regulação da expressão genética em micobactérias: controlo da transcrição dos genes líticos do micobacteriófago Ms6. PhD Thesis. Univ. of Lisbon.

Garcia M., Pimentel M. and Moniz-Pereira J. (2002) Expression of mycobacteriophage Ms6 lysis genes is driven by two sigma (70)-like promoters and is dependent on a transcription termination signal present in the leader RNA. *J Bacteriol* **184:**3034-3043.

Gil F., Catalão M. J., Moniz-Pereira J., Leandro P., McNeil M. and Pimentel M. (2008) The lytic cassette of mycobacteriophage Ms6 encodes an enzyme with lipolytic activity. *Microbiology* **154(Pt 5):**1364-1371.

Gil F., Grzegorzewicz A. E., Catalão M. J., Vital J., McNeil M. and Pimentel M. (2010) The mycobacteriophage Ms6 LysB specifically targets the outer membrane of *Mycobacterium smegmatis*. *Microbiology* **156:**1497-1504.

Graschopf A. and Bläsi U. (1999a) Functional assembly of the lambda S holin requires periplasmic localization of its N-terminus. *Arch Microbiol* **172(1):**31-39.

Graschopf A. and Bläsi U. (1999b) Molecular function of the dual-start motif in the lambda S holin. *Mol Microbiol* **33(3):**569-582.

Gründling A., Bläsi U. and Young R. (2000a) Biochemical and genetic evidence for three transmembrane domains in the class I holin, lambda S. *J Biol Chem* **275(2):**769-776.

Gründling A., Bläsi U. and Young R. (2000b) Genetic and biochemical analysis of dimer and oligomer interactions of the lambda S holin. *J Bacteriol* **182(21):**6082-6090.

Gründling A., Smith D. L., Bläsi U. and Young R. (2000c) Dimerization between the holin and holin inhibitor of phage lambda. *J Bacteriol* **182(21):**6075-6081.

Gründling A., Manson M. D. and Young R. (2001) Holins kill without warning. *Proc Natl Acad Sci USA* **98:**9348-9352.

Guttman B., Raya R. and E. Kutter (2005) 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.

Hanych B., Kedzierska S., Walderich B., Uznański B. and Taylor A. (1993) Expression of the Rz gene and the overlapping Rz1 reading frame present at the right end of the bacteriophage lambda genome. *Gene* **129(1):**1-8.

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

Hatfull G. F. (2000) Molecular genetics of mycobacteriophages. *In*, Molecular genetics of mycobacteria. Hatfull G. F. and Jacobs W. R. Jr. (Eds.) ASM Press, Washington D. C., p. 3-54.

Hatfull G. F. (2005a) Mycobacteriophages and tuberculosis. *In*, Tuberculosis and the tubercle bacillus. Stewart T. Cole *et al*. (Ed.) ASM Press, Washington D. C., p. 203-218.

Hatfull G. F. (2005b) Mycobacteriophages: pathogenesis and applications. *In*, Phages: Their role in bacterial pathogenesis and biotechnology. Waldor M. K., Friedman D. I. and S. Adhya (Eds.) ASM Press, Washington D.C., p. 238-255.

Hatfull G. F. (2006) Mycobacteriophages. *In*, The bacteriophages. Calendar R. (Ed.), 2nd ed. Oxford University Press, New York, p. 602-620.

Hatfull G. F., Pedulla M. L., Jacobs-Sera D., Cichon P. M., Foley A., Ford M. E., Gonda R. M., Houtz J. M., Hryckowian A. J., Kelchner V. A., Namburi S. Pajcini K. V., Popovich M. G., Schleicher D. T., Simanek B. Z., Smith A. L., Zdanowicz G. M., Kumar V., Peebles C. L., Jacobs W. R. Jr., Lawrence J. G. and Hendrix R. W. (2006) Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genetics* **2:**0835-0847.

Hatfull G. F. (2008a) Bacteriophage genomics. *Curr Opin Microbiol* **11:**447-453.

Hatfull G. F. (2008b) What can mycobacteriophages tell us about *Mycobacterium tuberculosis*? *In*, National Institute of Allergy and Infectious Diseases, NIH. Georgiev V., Western K. A., McGowan J. J. (Ed.), Volume I, Frontiers in Research. Humana Press, Totowa, N. J., p. 67-76.

Hatfull G. F., Cresawn S. G. and Hendrix R. W. (2008) Comparative genomics of the mycobacteriophages: insights into bacteriophage evolution. *Res Microbiol* **159(5):**332-339.

Hatfull G. F., Jacobs-Sera D., Lawrence J. G., Pope W. H., Russell D. A., Ko C. C., Weber R. J., Patel M. C., Germane K. L., Edgar R. H., Hoyte N. N., Bowman C. A., Tantoco A. T., Paladin E. C., Myers M. S., Smith A. L., Grace M. S., Pham T. T., O'Brien M. B., Vogelsberger A. M., Hryckowian A. J., Wynalek J. L., Donis-Keller H., Bogel M. W., Peebles C. L., Cresawn S. G. and Hendrix R. W. (2010) Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. *J Mol Biol* **397(1):**119-143.

Hendrix R. W., Smith M. C. M., Burns R. N., Ford M. E. and Hatfull G. F. (1999) Evolutionary relationships among diverse bacteriophages and prophages: All the world´s a phage. *Proc Natl Acad Sci USA* **96:**2192-2197.

Hendrix R. W. (2003) Bacteriophage genomics. *Curr Opin Microbiol* **6:**506-511.

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

Henry M., O'Sullivan O., Sleator R. D., Coffey A., Ross R. P., McAuliffe O. and O'Mahony J. M. (2010) *In silico* analysis of Ardmore, a novel mycobacteriophage isolated from soil. *Gene* **453(1-2):**9-23.

Hoffmann C., Leis A., Niederweis M., Plitzko J. M. and Engelhardt H. (2008) Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci USA* **105(10):**3963-3967.

Hottenrott S., Schumann T., Plückthun A., Fischer G. and Rahfeld J. U. (1997) The *Escherichia coli* SlyD is a metal ion-regulated peptidyl-prolyl cis/trans-isomerase. *J Biol Chem* **272(25):**15697-15701.

Husson R. N., James B. E. and Young R. A. (1990) Gene replacement and expression of foreign DNA in mycobacteria. *J Bacteriol* **172:**519–524.

Jacob F. and Fuerst C. R. (1958) The mechanism of lysis by phage studied with defective lysogenic bacteria. *J Gen Microbiol* **18(2):**518-526.

Jado I., López R., García E., Fenoll A., Casal J. and García P. (2003) Phage lytic enzymes as therapy for antibiotic-resistant *Streptococcus pneumoniae* infection in a murine sepsis model. *J Antimicrob Chemother* **52(6):**967-973.

Johnson-Boaz R., Chang C. Y. and Young R. (1994) A dominant mutation in the bacteriophage lambda S gene causes premature lysis and an absolute defective plating phenotype. *Mol Microbiol* **13(3):**495-504.

Kakikawa M., Yokoi K., Kimoto H., Nakano M., Kawasaki K., Taketo A. and Kodaira K. (2002) Molecular analysis of the lysis protein Lys encoded by *Lactobacillus plantarum* phage ϕg1e. *Gene* **299:**227-234.

Kanamaru S., Ishiwata Y., Suzuki T., Rossmann M. G. and Arisaka F. (2005) Control of bacteriophage T4 tail lysozyme activity during the infection process. *J Mol Biol* **346(4):**1013-1020.

Karnik S. and Billeter M. (1983) The lysis function of RNA bacteriophage Qβ is mediated by the maturation (A2) protein. *EMBO J* **2(9):**1521-1526.

Kedzierska S., Wawrzynów A. and Taylor A. (1996) The *Rz1* gene product of bacteriophage lambda is a lipoprotein localized in the outer membrane of *Escherichia coli*. *Gene* **168(1):**1-8.

van Kessel J. C. and Hatfull G. F. (2007) Recombineering in *Mycobacterium tuberculosis*. *Nat Methods* **4(2):**147-152.

van Kessel J. C. and Hatfull G. F. (2008) Efficient point mutagenesis in mycobacteria using single stranded DNA recombineering: characterization of antimycobacterial drug targets. *Mol Microbiol* **67(5):**1094-1107.

van Kessel J. C., Marinelli L. J. and Hatfull G. F. (2008) Recombineering mycobacteria and their phages. *Nat Rev Microbiol* **6(11):**851-857.

Krogh S., Jørgensen S. T. and Devine K. M. (1998) Lysis genes of the *Bacillus subtilis* defective prophage PBSX. *J Bacteriol* **180(8):**2110-2117.

Krupovič M., Cvirkaite-Krupovič V. and Bamford D. H. (2008) Identification and functional analysis of the *Rz*/*Rz1*-like accessory lysis genes in the membrane-containing bacteriophage PRD1. *Mol Microbiol* **68(2):**492-503.

Krupovič M. and Bamford D. H. (2008) Holin of bacteriophage lambda: structural insights into a membrane lesion. *Mol Microbiol* **69(4):**781-783.

Liu J., Dehbi M., Moeck G., Arhin F., Bauda P., Bergeron D., Callejo M, Ferretti V., Ha T., Kwan T., McCarty J., Srikumar R., Williams D., Wu J. J., Gros P., Pelletier J. and DuBow M. (2004) Antimicrobial drug discovery through bacteriophage genomics. *Nature Biotech* **22(2):**185-191.

Liu J., Glazko G. and Mushegian A. (2006) Protein repertoire of double-stranded DNA bacteriophages. *Virus Res* **117(1):**68-80.

Łobocka M. B., Rose D. J., Plunkett G. 3rd, Rusin M., Samojedny A., Lehnherr H., Yarmolinsky M. B. and Blattner F. R. (2004) Genome of bacteriophage P1. *J Bacteriol* **186(21):**7032-7068.

Loeffler, J. M., Nelson, D. and V. A. Fischetti (2001) Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. *Science* **294:**2170-2172.

Loeffler J. M. and Fischetti V. A. (2003) Synergistic lethal effect of a combination of phage lytic enzymes with different activities on penicillin-sensitive and -resistant *Streptococcus pneumoniae* strains. *Antimicrob Agents Chemother* **47(1):**375-377.

Loeffler J. M., Djurkovic S. and Fischetti V. A. (2003) Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia. *Infect Immun* **71(11):**6199-6204.

Loessner M. J., Maier S. K., Daubek-Puza H., Wenlinger G. and Scherer S. (1997) Three *Bacillus cereus* bacteriophage endolysins are unrelated but reveal high homology to cell wall hydrolases from different bacilli. *J Bacteriol* **179:**2845-2851.

Loessner M. J. (2005) Bacteriophage endolysins: current state of research and applications. *Curr Opin Microbiol* **8:**480-487.

Longchamp P. F., Mauël C. and Karamata D. (1994) Lytic enzymes associated with defective prophages of *Bacillus subtilis*: sequencing and characterization of the

region comprising the *N*-acetylmuramoyl-L-alanine amidase gene of prophage PBSX. *Microbiology* **140(Pt 8):**1855-1867.

Marinelli L. J., Piuri M., Swigonová Z., Balachandran A., Oldfield L. M., van Kessel J. C. and Hatfull G. F. (2008) BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. *PLoS One* **3(12):**e3957.

Markov D., Christie G. E., Sauer B., Calendar R., Park T., Young R. and Severinov K. (2004) P2 growth restriction on an *rpoC* mutant is suppressed by alleles of the *Rz1* homolog *lysC*. *J Bacteriol* **186(14):**4628-4637.

Matsuzaki S., Rashel M., Uchiyama J., Sakurai S., Ujihara T., Kuroda M,. Ikeuchi M., Tani T., Fujieda M., Wakiguchi H. and Imai S. (2005) Bacteriophage therapy: a revitalized therapy against bacterial infectious diseases. *J Infect Chemother* **11:**211-219.

McNerney R. (1999) TB: the return of the phage. A review of fifty years of mycobacteriophage research. *Int J Tuberc Lung Dis* **3(3):**179-184.

McNerney R. and Traoré H. (2005) Mycobacteriophage and their application to disease control. *J Appl Microbiol* **99(2):**223-233.

Méderlé I., Bourguin I., Ensergueix D., Badell E., Moniz-Pereira J., Gicquel B. and Winter N. (2002) Plasmidic versus insertional cloning of heterologous genes in *Mycobacterium bovis* BCG: impact on *in vivo* antigen persistence and immune responses. *Infect Immun* **70(1):**303-314.

Mendel S., Holbourn J. M., Schouten J. A. and Bugg T. D. (2006) Interaction of the transmembrane domain of lysis protein E from bacteriophage ϕX174 with bacterial translocase MraY and peptidyl-prolyl isomerase SlyD. *Microbiology* **152(Pt 10):**2959-2967.

Moffatt B. A. and Studier F. W. (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49(2):**221-227.

Nachega J. B. and Chaisson R. E. (2003) Tuberculosis drug resistance: a global threat. *Clin Infect Dis* **36(Suppl 1):**S24-30.

Nascimento J. G., Guerreiro-Pereira M. C., Costa S. F., São-José C. and Santos M. A. (2008) Nisin-triggered activity of Lys44, the secreted endolysin from *Oenococcus oeni* phage fOg44. *J Bacteriol* **190(1):**457-461.

Navarre W. W., Ton-That H., Faull K. F. and Schneewind O. (1999) Multiple enzymatic activities of the murein hydrolase from staphylococcal phage ϕ11. Identification of a D-alanyl-glycine endopeptidase activity. *J Biol Chem* **274(22):**15847-15856.

Nelson D., Loomis L. and Fischetti V. A. (2001) Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proc Natl Acad Sci USA* **98(7):**4107-4112.

Nikaido H. (1994) Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science* **264(5157):**382-388.

O´Brien R. J. and Nunn P. P. (2001) The need for new drugs against tuberculosis. Obstacles, opportunities, and next steps. *Am J Respir Crit Care Med* **163(5):**1055-1058.

O´Flaherty S., Ross R. P. and Coffey A. (2009) Bacteriophage and their lysins for elimination of infectious bacteria. *FEMS Microbiol Rev* **33:**801-819.

Paddison P., Abedon S. T., Dressman H. K., Gailbreath K., Tracy J., Mosser E., Neitzel J., Guttman B. and Kutter E. (1998) The roles of the bacteriophage T4 *r* genes in lysis inhibition and fine-structure genetics: a new perspective. *Genetics* **1148(4):**1539-1550.

Pang T., Savva C. G., Fleming K. G., Struck D. K. and Young R. (2009) Structure of the lethal phage pinhole. *Proc Natl Acad Sci USA* **106(45):**18966-18971.

Pang T., Park T. and Young R. (2010) Mutational analysis of the S^{21} pinholin. *Mol Microbiol* **76(1):**68-77.

Park T., Struck D. K., Deaton J. F. and Young R. (2006) Topological dynamics of holins in programmed bacterial lysis. *Proc Natl Acad Sci USA* **103(52):**19713-19718.

Park T., Struck D. K., Dankenbring C. A. and Young R. (2007) The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. *J Bacteriol* **189(24):**9135-9139.

Parreira R., São-José C., Isidro A., Domingues S., Vieira G. and Santos M. A. (1999) Gene organization in a central DNA fragment of *Oenococcus oeni* bacteriophage fOg44 encoding lytic, integrative and non-essential functions. *Gene* **1226(1):**83-93.

Pashley C.A., Parish T., McAdam R.A., Duncan K. and Stoker N.G. (2003) Gene replacement in mycobacteria by using incompatible plasmids. *Appl Environ Microbiol* **69:**517–523.

Payne K., Sun Q., Sacchettini J. and Hatfull G. F. (2009) Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. *Mol Microbiol* **73(3):**367-381.

Pedulla M. L., Ford M. E., Houtz J. M., Karthikeyan T., Wadsworth C., Lewis J. A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N. R., Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S., Kriakov J., Lawrence J. G., Jacobs W. R., Hendrix R. W. and Hatfull G. F. (2003) Origins of highly mosaic mycobacteriophage genomes. *Cell* **113:**171-182.

Pelicic V., Reyrat J. M. and Gicquel B. (1996) Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Mol Microbiol* **20:**919–925.

Peng L., Chen B. W., Luo Y. A. and Wang G. Z. (2006) Effect of mycobacteriophage to intracellular mycobacteria *in vitro*. *Chin Med J (Engl)* **119(8):**692-695.

Piechota C., Gil F., Mahapatra S., Crick D. C., Yufang M., Huang H., Scherman M. S., Jones V., Pavelka M. S., Moniz-Pereira J., Pimentel M. and McNeil M. R. (2005) Mycobacteria phage Ms6 amidase reveals highly amidated mycobacterial peptidoglycan. Unpublished results.

Pimentel M. (1999) Biologia molecular da resistência à superinfecção por bacteriófagos: estudo do gene *pin* do micobacteriófago Ms6. PhD Thesis. Univ. of Lisbon.

Piuri M., Jacobs W. R. Jr. and Hatfull G. F. (2009) Fluoromycobacteriophages for rapid, specific, and sensitive antibiotic susceptibility testing of *Mycobacterium tuberculosis*. *PLoS One* **4(3):**e4870.

Portugal, I., Anes, E. and Moniz-Pereira J. (1989) Temperate mycobacteriophage from *M. smegmatis*. *Acta Leprol* **7:**243-244.

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

Raab R., Neal G., Garrett J., Grimaila R., Fusselman R. and Young R. (1986) Mutational analysis of bacteriophage lambda lysis gene S. *J Bacteriol* **167(3):**1035-1042.

Raab R., Neal G., Sohaskey C., Smith J. and Young R. (1988) Dominance in lambda S mutations and evidence for translational control. *J Mol Biol* **199(1):**95-105.

Ramanculov E. and Young R. (2001a) Genetic analysis of the T4 holin: timing and topology. *Gene* **265(1-2):**25-36.

Ramanculov E. and Young R. (2001b) Functional analysis of the phage T4 holin in a lambda context. *Mol Genet Genomics* **265(2):**345-353.

Ramanculov E. and Young R. (2001c) An ancient player unmasked: T4 *rI* encodes a t-specific antiholin. *Mol Microbiol* **41(3):**575-583.

Rashel M., Uchiyama J., Ujihara T., Uehara Y., Kuramoto S., Sugihara S., Yagyu K., Muraoka A., Sugai M., Hiramatsu K., Honke K. and Matsuzaki S. (2007) Efficient elimination of multidrug-resistant *Staphylococcus aureus* by cloned lysin derived from bacteriophage ϕMR11. *J Infect Dis* **196(8):**1237-1247.

Rietsch A., Fraisl P., Graschopf A. and Bläsi U. (1997) The hydrophilic C-terminal part of the lambda S holin is non-essential for intermolecular interactions. *FEMS Microbiol Lett* **153(2):**393-398.

Roof W. D., Horne S. M., Young K. D. and Young R. (1994) *slyD*, a host gene required for ϕX174 lysis, is related to the FK506-binding protein family of peptidyl-prolyl cis-trans-isomerases. *J Biol Chem* **269(4):**2902-2910.

Roof W. D., Fang H. Q., Young K. D., Sun J. and Young R. (1997) Mutational analysis of *slyD*, an *Escherichia coli* gene encoding a protein of the FKBP immunophilin family. *Mol Microbiol* **25(6):**1031-1046.

Roucourt B. and Lavigne R. (2009) The role of interactions between phage and bacterial proteins within the infected cell: a diverse and puzzling interactome. *Environ Microbiol* **11(11):**2789-805.

Russel M. (1995) Moving through the membrane with filamentous phages. *Trends Microbiol* **3(6):**223-228.

São-José C., Parreira R., Vieira G. and Santos M. 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. *J Bacteriol* **178:**5823-5831.

São-José C. (2002) Genome analysis and gene expression in oenophage fOg44: evidence for a new strategy of bacteriophage-induced host lysis. PhD Thesis. Univ. Lisbon.

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

São-José C., Santos S., Nascimento J., Brito-Madurro A. G., Parreira R. and Santos M. A. (2004) Diversity in the lysis-integration region of oenophage genomes and evidence for multiple tRNA loci, as targets for prophage integration in *Oenococcus oeni*. *Virology* **325(1):**82-95.

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

Savva C. G., Dewey J. S., Deaton J., White R. L., Struck D. K., Holzenburg A. and Young R. (2008) The holin of bacteriophage lambda forms rings with large diameter. *Mol Microbiol* **69(4):**784-793.

Scheilfer K. H. and Kandler O. (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* **36:**407-477.

Schmidt C., Velleman M. and Arber W. (1996) Three functions of bacteriophage P1 involved in cell lysis. *J Bacteriol* **178(4):**1099-1104.

Schuch, R., Nelson, D. and Fischetti V. A. (2002) A bacteriolytic agent that detects and kills *Bacillus anthracis*. *Nature* **418:**884-889.

Smith D. L., Young R. (1998) Oligohistidine tag mutagenesis of the lambda holin gene. *J Bacteriol* **180(16):**4199-211.

Smith D. L., Struck D. K., Scholtz J. M. and Young R. (1998) Purification and biochemical characterization of the lambda holin. *J Bacteriol* **180(9):**2531-2540.

Steiner M. and Bläsi U. (1993) Charged amino-terminal amino acids affect the lethal capacity of lambda lysis proteins S_{107} and S_{105}. *Mol Microbiol* **8(3):**525-533.

Sularkvelidze A., Alavidze Z. and Morris J. G. Jr. (2001) Bacteriophage therapy. *Antimicrob Agents Chemother* **45(3):**649-659.

Summer E. J., Berry J., Tran T. A., Niu L., Struck D. K. and Young R. (2007) *Rz/Rz1* lysis gene equivalents in phages of Gram-negative hosts. *J Mol Biol* **373(5):**1098-1112.

Summers W. C. (2001) Bacteriophage therapy. *Annu Rev Microbiol* **55:**437-451.

Sun Q., Kuty G. F., Arockiasamy A., Xu M., Young R. and Sacchettini J. C. (2009) Regulation of a muralytic enzyme by dynamic membrane topology. *Nat Struct Mol Biol* **16(11):**1192-1194.

Taylor A. (1971) Endopeptidase activity of phage lamba-endolysin. *Nat New Biol* **234(48):**144-145.

Tran T. A., Struck D. K. and Young R. (2005) Periplasmic domains define holin-antiholin interactions in T4 lysis inhibition. *J Bacteriol* **187(19):**6631-40.

Tran T. A., Struck D. K. and Young R. (2007) The T4 RI antiholin has an N-terminal signal anchor release domain that targets it for degradation by DegP. *J Bacteriol* **89(21):**7618-7625.

Vukov N., Moll I., Bläsi U., Scherer S. and Loessner M. J. (2003) Functional regulation of the *Listeria monocytogenes* bacteriophage A118 holin by an

intragenic inhibitor lacking the first transmembrane domain. *Mol Microbiol* **48(1):**173-186.

dos Vultos T. D., Méderlé I., Abadie V., Pimentel M., Moniz-Pereira J., Gicquel B., Reyrat J. M. and Winter N. (2006) Modification of the mycobacteriophage Ms6 *attP* core allows the integration of multiple vectors into different $tRNA^{Ala}$ T-loops in slow- and fast-growing mycobacteria. *BMC Mol Biol* **7:**47.

Wagner P. L. and Waldor M. K. (2002) Bacteriophage control of bacterial virulence. *Infec Immun* **70(8):**3985-3993.

Waldor M. K. (1998) Bacteriophage biology and bacterial virulence. *Trends Microbiol* **6(8):**295-297.

Walker J. T. and Walker D. H. Jr. (1980) Mutations in coliphage P1 affecting host cell lysis. *J Virol* **35(2):**519-530.

Wang I.-N., Smith D. L. and Young R. (2000) Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* **54:**799-825.

Wang I.-N., Deaton J. and Young R. (2003) Sizing the holin lesion with an endolysin-beta-galactosidase fusion. *J Bacteriol* **185(3):**779-787.

Wang I.-N. (2006) Lysis timing and bacteriophage fitness. *Genetics* **172(1):**17-26.

Weinbauer M. G. (2004) Ecology of prokaryotic viruses. *FEMS Microbiol Rev* **28:**127-181.

White R., Tran T. A., Dankenbring C. A., Deaton J. and Young R. (2010) The N-terminal transmembrane domain of lambda S is required for holin but not antiholin function. *J Bacteriol* **192(3):**725-733.

Wood H. E., Dawson M. T., Devine K. M. and McConnell D. J. (1990) Characterization of PBSX, a defective prophage of *Bacillus subtilis*. *J Bacteriol* **172(5):**2667-2674.

Xu M., Struck D. K., Deaton J., Wang I. and Young R. (2004) A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. *Proc Natl Acad Sci USA* **101:**6415-6420.

Xu M., Arulandu A., Struck D. K., Swanson S., Sacchettini J. C. and Young R. (2005) Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. *Science* **307(5706):**113-117.

Yin Y. and Fisher D. (2008) Identification and investigation of ORFans in the viral world. *BMC Genomics* **9:**24.

Yokoi K. J., Kawahigashi N., Uchida M., Sugahara K., Shinohara M., Kawasaki K., Nakamura S., Taketo A. and Kodaira K. (2005) The two-component cell lysis genes *holWMY* and *lysWMY* of the *Staphylococcus warneri* M phage ϕWMY: cloning, sequencing, expression, and mutational analysis in *Escherichia coli*. *Gene* **351:**97-108.

Young R., Way J., Way S., Yin J. and Syvanen M. (1979) Transposition mutagenesis of bacteriophage lambda: a new gene affecting cell lysis. *J Mol Biol* **132(3):**307-322.

Young R. (1992) Bacteriophage lysis: mechanism and regulation. *Microbiol Reviews* **56:**430-481.

Young R., Wang I.-N. and Roof W. D. (2000) Phages will out: strategies of host cell lysis. *Trends Microbiol* **8:**120-127.

Young R. (2002) Bacteriophage holins: deadly diversity. *J Mol Microbiol Biotechnol* **4:**21-36.

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

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

Zagotta M. T. and Wilson D. B. (1990) Oligomerization of the bacteriophage lambda S protein in the inner membrane of *Escherichia coli*. *J Bacteriol* **172(2):**912-921.

Zhang N. and Young R. (1999) Complementation and characterization of the nested *Rz* and *Rz1* reading frames in the genome of bacteriophage lambda. *Mol Gen Genet* **262(4-5):**659-667.

Zhang Y. and Telenti A. (2000) Genetics of drug resistance in *Mycobacterium tuberculosis*. *In*, Molecular genetics of mycobacteria. Hatfull G. F. and Jacobs W. R. Jr. (Eds.) ASM Press, Washington D. C., p. 235-254.

Zheng Y., Struck D. K., Bernhardt T. G. and Young R. (2008) Genetic analysis of MraY inhibition by the ϕX174 protein E. *Genetics* **180(3):**1459-1466.

Ziermann R., Bartlett B., Calendar R. and Christie G. E. (1994) Functions involved in bacteriophage P2-induced host cell lysis and identification of a new tail gene. *J Bacteriol* **176(16):**4974-4984.

Zuber B., Chami M., Houssin C., Dubochet J., Griffiths G. and Daffé M. (2008) Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J Bacteriol* **190(16):**5672-5680.

Objectives

The research presented in this thesis was motivated by three major objectives. First, our main goal was to characterize the role of the accessory lysis protein Gp1 during *Mycobacterium smegmatis* infection by the mycobacteriophage Ms6. By using a variety of experimental approaches, including both biochemical and genetic methods, we confirmed that Gp1 is involved in endolysin translocation across the cytoplasmic membrane functioning as a potential substrate-specific chaperone for the *sec*-system and is essential for an efficient lysis of mycobacteria and to a productive burst size. Second, we investigated the potential holin function of Gp4 and Gp5. As a result of their inefficacy to support an efficient lysis of *E. coli* in the presence of the Ms6 endolysin, the study of their role in lysis was further expanded to mycobacteria. Finally, given the requirement of Gp1 for lysis we intended to characterize the role of Gp1 and the endolysin interaction during *M. smegmatis* infection by Ms6.

The specific questions addressed in this thesis are:

1. Is Gp1 required for mycobacteria lysis and how is it involved in endolysin translocation across the cytoplasmic membrane?

2. What is the role of the mycobacteriophage Ms6 holin-like proteins, in view of the fact that they are not required for *M. smegmatis* lysis?

3. Does the interaction between Gp1 and the Ms6 endolysin play a role during mycobacteria infection?

Ultimately, our overarching goal is to contribute to the progress being made in understanding the molecular pathways involved in mycobacteriophage lysis: identification and characterization of the mycobacteriophage lysis effectors and the exact pathways by which they are positioned next to their targets is of interest not only for understanding the mechanisms and timing of lysis but also because of the potential therapeutic use of mycobacteriophage lysins.

2

The Mycobacteriophage Ms6 Encodes a Chaperone-Like Protein Involved in the Endolysin Delivery to the Peptidoglycan

M. J. Catalão, F. Gil, J. Moniz Pereira, and M. Pimentel

Centro de Patogénese Molecular, Unidade dos Retrovírus e Infecções Associadas, Faculty of Pharmacy, University of Lisbon, Portugal

Molecular Microbiology (In final revision)
2010

Summary

Like most double-stranded (ds) DNA phages, mycobacteriophage Ms6 uses the holin-endolysin system to achieve lysis of its host. In addition to endolysin (*lysA*) and holin (*hol*) genes, Ms6 genome encodes three accessory lysis proteins. In this study we investigated the lysis function of Gp1, which is encoded by the *gp1* gene that lies immediately upstream of *lysA*. *Escherichia coli* lysis was observed after co-expression of Gp1 and LysA in the absence of Ms6 holin. Gp1 does not belong to the holin class of proteins, and we provide evidence that it shares several characteristics with molecular chaperones. We show that Gp1 interacts with LysA, and that this interaction is necessary for LysA delivery to its target. In addition, *phoA* gene fusions showed that, in *Mycobacterium smegmatis,* LysA is exported to the extracytoplasmic environment in the presence of Gp1. Gp1 is also necessary for efficient *M. smegmatis* lysis, as Ms6 *gp1* deletion results in host lysis defects. We propose that delivery of Ms6 endolysin to the murein layer is assisted by Gp1, a chaperone-like protein, in a holin-independent manner.

Introduction

The majority of phages described to date is endowed with a tail (96%) and presents a double-stranded DNA (dsDNA) genome (Ackermann, 2003). At the end of their vegetative cycle, these phages achieve lysis of their hosts using a "holin-endolysin" strategy (Young *et al*., 2000; Young and Wang, 2006). Endolysins are proteins that target the integrity of the cell wall, designed to attack one or more of the three types of peptidoglycan covalent bonds. Holins are small membrane proteins that control the activation of the endolysin or its access to the murein (Xu *et al*., 2005; Young, 2005; Young and Wang, 2006). The lambda model of lysis was long thought to be universal (Xu *et al*., 2004). In this model, holin accumulates in the inner membrane during late gene expression without disturbing its integrity, while enzymatically active endolysin accumulates in the cytoplasm. Then, at a genetically defined time, holin triggers to disrupt the membrane allowing endolysin to cross it and attack the peptidoglycan, resulting in cell lysis (Wang *et al*., 2000; Gründing *et al*., 2001).

Recent results have shown that endolysins may be transported across the cytoplasmic membrane in a holin-independent manner. The first example was described for *Oenococcus oeni* phage fOg44 endolysin (Lys44), which possesses a typical cleavable signal peptide (São-José *et al*., 2000). Lys44 does not accumulate in the cytoplasm as lambda endolysin, but is continuously exported during phage assembly to the extracytoplasmic environment by the Sec translocon. It remains ineffective until the holin lesion acts as a signal for its activation through dissipation of the proton-motive force (p.m.f.) (São-José *et al.,* 2000; Nascimento *et al.,* 2008).

More recently, Xu *et al.* (2004) reported the existence of an atypical signal sequence named SAR (signal-arrest-release) in the N-terminal domain of phage P1 endolysin. P1 Lyz export does not require holin action, but is mediated by the N-

terminal transmembrane domain and, like fOg44 endolysin, requires host Sec function. The SAR motif is not proteolytically cleaved and operates, in a first step, as a signal-arrest domain, directing the endolysin to the periplasm in a membrane-tethered form where it remains enzymatically inactive. In a second step, membrane depolarization triggers the SAR domain to allow P1 Lyz to be released as a soluble active enzyme in the periplasm (Xu *et al.*, 2004). Sequence comparison has identified additional endolysins with N-terminal hydrophobic sequences in phages infecting Gram-negative hosts. Analysis of these sequences suggests that they could function as a signal anchor and, like P1 Lyz, could engage the Sec-system (Young and Wang, 2006).

With the lambda paradigm no longer universal new mechanisms for endolysins export may come to light. Our group has studied the lysis module of mycobacteriophage Ms6, a phage that infects *Mycobacterium smegmatis* (Portugal *et al*., 1989). The lytic operon of Ms6 is organized into five genes. The second gene (*lysA*) encodes the endolysin, while the holin is encoded by *gp4* (*hol*) (Garcia *et al*., 2002). Between these two genes lies *lysB*, a gene recently identified as coding for a protein with lipolytic activity that disrupts the outer membrane of the mycobacterium complex envelope (Gil *et al.*, 2008; Gil *et al*., 2010). The *gp1* and *gp5* genes code for proteins of unknown functions, although similar putative proteins are encoded in the lysis module of other mycobacteriophages (Hatfull *et al*., 2006; Hatfull *et al*., 2010). Despite classification as Gram-positive bacteria, mycobacteria have a complex cell wall, suggesting that phages infecting these hosts have evolved an adequate lysis module.

We recently observed that co-expression of *gp1* and *lysA* in *E. coli* results in cell lysis. No such effect occurs when only one of the two genes is independently expressed (Garcia *et al*., 2002; our unpublished observations). These data, together with the previous reported absence of lysis when *lysA* and *hol* were co-expressed in *E. coli* (Garcia *et al*., 2002), led us to investigate the role of the *gp1* gene product in

lysis. In this report we demonstrate that, similar to the above examples, Ms6 endolysin export does not require holin function. We present experimental evidence that Ms6 Gp1 is involved in lysis, assisting the export of endolysin LysA to the extracytoplasmic environment. The Ms6 lysis system is novel and may be employed by other mycobacteriophages.

Experimental Procedures

Bacterial strains, phages, plasmids and culture conditions

Bacterial strains, phages and plasmids used in this study are listed in Table 1. *E. coli* strains were grown at 37 ºC, in Luria-Bertani (LB) broth or agar supplemented with 100 µg ml^{-1} ampicillin or 30 µg ml^{-1} kanamycin, when appropriate. *M. smegmatis* recombinant strains were grown at 37 ºC in 7H9 medium (Difco) supplemented with 0.05% Tween 80, with shaking, or in Middlebrook 7H10 (Difco) containing 15 µg ml^{-1} kanamycin. For induced conditions, 2% succinate and 0.2% or 2% acetamide were also added to media.

Table 1. Strains, bacteriophages and plasmids used in this study.

Strain, bacteriophage, or plasmid	Description	Reference or source
Bacteria		
Escherichia coli		
JM109	*recA1 endA1 gyr96 thi hsdR17 supE44 relA1* Δ(*lac-proAB*) [F′ *traD36 proAB lacI^{q}Z*ΔM15]	Stratagene
BL21 (DE3)	F$^-$ *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dmc* (DE3)	Novagen
MC4100	*araD139* Δ(*argF-lac*)*U169 rpsL150 relA1 flbB5301deoC1 ptsF25 rbsR*	Silhavy *et al.*, 1984
MM52	MC4100 *secA51* (Ts)	Oliver and Beckwith, 1981
K802	F$^-$ e14$^-$ (McrA$^-$) *lacY1* or *Δ(lac)6 supE44 galK2 galT22 rfbD1 metB1 mcrB1 hsdR2*($r_k^-m_k^+$)	Sambrook and Russell, 2001
Y1088	e14 (*mcrA)* Δ(*lac*)*U169 supE supF hsdR metB trpR tonA21 proC::Tn5* [pMC9]	Stratagene
Mycobacterium smegmatis		
mc^{2}155	High-transformation-efficiency mutant of *M. smegmatis* ATCC 607	Snapper *et al.*, 1990
mc^{2}2522	mc^{2}155 Δ*secA2*	Braunstein *et al.*,2001

Bacteriophages		
$Ms6_{wt}$	Temperate bacteriophage from *M. smegmatis*	Portugal *et al*., 1989
$Ms6_{\Delta gp1}$	213 bp in-frame deletion of Ms6 *gp1* gene	This study
Ms6-LysAHis$_6$	His$_6$tag insertion at the 3' end of Ms6 *lysA* gene	This study
	cIts857 Sam100	
λgt11		Stratagene
Plasmids		
pQE30	Expression vector; T5 promoter; Amp^r	QIAGEN
pQE-80L	Expression vector; T5 promoter; Amp^r, $lacI^q$	QIAGEN
pET29b(+)	Expression vector, T7 promoter; Kan^r	Novagen
pVVAP	Mycobacteria shuttle vector carrying the acetamidase promoter; Kan^r	Unpublished
pJV53	Derivative of pLAM12 with Che9c *60* and *61* under control of the acetamidase promoter; Kan^r	van Kessel and Hatfull, 2007
pSMT3-[19-phoA]	Mycobacteria plasmid containing the structural gene for *E. coli* PhoA	Herrmann *et al*., 1996
pMP201	Mycobacteria integrative vector for gene expression under the control of Ms6 P*lys* promoter	Unpublished
pMG231A	*lysA* cloned in pQE30	Garcia *et al*., 2002
pFG1	*gp1* cloned in pQE30	This study
pMP320	*gp1* and *lysA* cloned in pQE30	This study
pMJC1	*gp1* cloned in pET29b	This study
pMJC2	*gp1* with stop codon included cloned in pET29b	This study
pMJC3	*gp1* and *lysA* cloned in pET29b	This study
pMJC4	*lysA* cloned in pET29b	This study
pMJC5	*gp1* and *lysA* cloned in pQE-80L	This study
pMJC6	*gp1* cloned in pVVAP	This study
pMJC7	*gp1* and *lysA* cloned in pVVAP	This study
pMJC8	*lysA* cloned in pVVAP	This study
pMJC9	*gp1* and *λR* cloned in pQE30	This study
pMJC10	*gp1* and *λR* cloned in pET29b	This study
pMJC11	*gp1*, *lysA*, *lysB*, *hol* and *gp5* cloned in pVVAP	This study
pMJC12	*lysA-phoA'* fusion in pMP201	This study
pMJC13	*gp1* and *lysA-phoA*' fusion in pMP201	This study
pMJC14	*gp1* and 180 bp N-terminal *lysA* deletion cloned in pQE30	This study
pMJC15	*gp1* and 180 bp N-terminal *lysA* deletion cloned in pET29b	This study
pMJC17	*gp1* and the first 180 bp of *lysA* cloned in pET29b	This study

Ms6 lysis genes Accession No. AF319619.

Plasmid construction

Unless otherwise indicated, DNA fragments were amplified using Ms6 genomic DNA as template. DNA amplification, plasmid isolation and electrophoresis were carried out using standard techniques (Sambrook and Russell, 2001). *E. coli* and *M. smegmatis* mc^2155 cells were transformed as described previously (Snapper *et al*., 1990; Sambrook and Russell, 2001). Restriction

enzymes and T4 DNA ligase (New England Biolabs) were used according to supplier's recommendations. All oligonucleotides were from Thermo Scientific and are listed in Table 2.

Table 2. Oligonucleotides used in this study.

Primer Name	Sequence 5´-3´[a]	Comment
Pr*gp1*afwd	CGGGATCCATGGACCGCTTAGGATCGTCC	Includes BamHI site to clone in pQE30 and pQE80-L
Pr*gp1*arv	ATCTTTCGTGGAAGCTTCACCGTTTGC	Includes HindIII site to clone in pQE30
Pr*lysA*brv	ATGCGAAGCTTCAGTGGCCCAACAGTTC	Includes HindIII site to clone in pQE30 and pQE80-L
Pr*gp1*bfwd	CGGGATCCCGACCGCTTAGGCATC	Includes BamHI site to clone in pET29b
Pr*gp1*brv	CTTTCGTGGAAGCTTCCGTTTGCTCC	Includes HindIII site to clone in pET29b and PVVAP
Pr*lysA*crv	GCGAAGCTTGTG GCCCAACAGTTC	Includes HindIII site to clone in pET29b and pVVAP
Pr*lysA*cfwd	CGCGGATCCCACCACGAAAGATCAAG	Includes BamHI site to clone in pET29b
Pr*gp1*crv	ATCTTTCGTGGAAGCTTCACCGTTTGC	Includes HindIII site to clone in pET29b
Pr*gp1*dfwd	CGCGGATCCGGTTTTTCTATGCC	Includes BamHI site to clone in pVVAP
Pr*lysA*dfwd	GGACATATGCCTCGGGAGCAAACGGTGAC	Includes NdeI site to clone in pVVAP
Pr*gp1*drv	GCGAGCTCCACTCATTTCTCCCGAGG	Includes SacI site to clone in pQE30 and pET29b
Pr*gp1*erv	CTGAGCTCTCACCGTTTGCTCCCGAG	Includes SacI site to clone in pQE30 and pET29b
P*gp5*rv	CGCAAGCTTCCGGTCGCGGCGTTTG	Includes HindIII site to clone in pVVAP
Prλ*R*fwd	CGAGCTCTCTGACCGGAGTAGAAGATGG	Includes the RBS and the start codon of λ*R* and a SacI site to clone in pQE30 and pET29b
Prλ*R*rv	GGAAGCTTTACATCAATCTCTCTGACCG	Includes HindIII site to clone in pQE30 and pET29b
PrΔ180bp*lysA*fwd	GCGAGCTCCGCTTCGACGGTGCC	Includes SacI site to clone in PQE30 and pET29b
Pr*lysA*180bprv	GCAAGCTTGTGTGGGTAGGAGCCGTCC	Includes HindIII site to clone in pET29b; Ms6 *gp1*Δ^{FP}
Pr*phoA*'BamHIfwd	GCCGGATCCCCTGTTCTGGAAAACC	Includes BamHI site to ligate to the 3' end of a PCR generated fragment of *lysA* containing a SGGGS 3´ linker
Pr*phoA*'HindIIIrv	GCGGCGCGGTTAAGCTTCAGGG	Includes HindIII site to clone in pMP201
Pr*gp1*pMP201fwd	CGCAAGCTTGGTTTTTCTATGCC	Includes the RBS and the start codon of *gp1* and a HindIII site to clone in pMP201
Pr*lysA*pMP201fwd	GGAAAGCTTCCTCGGGAGCAAACGGTGAC	Includes the RBS and the start

		codon of *lysA* and a HindIII site to clone in pMP201
Pr*lysA*-SGGGSrv	GTGGGATCCGCTGCCGCCGCCGCTGTGGCC CAACAGTTC	Includes a SGGGS 3′ linker and a BamHI site to ligate to the 5' end of a PCR generated fragment of *phoA*'
PrP1fwd	CGGTACTAGTCGGCCTCGGCCTGC	Ms6 *gp1*Δ [FP]
PrΔ*gp1*	CGCCACCACCTGAGTGCGCGGGGTTTTTCTAT GCCCGAAAGGACCCCGACATCCTCGGGAGCA AACGGTGACCACGAAAGATCAAGTCGCCCAA ATCACC	Ms6 *gp1* 213bpΔ
PrExtΔ*gp1*fwd	CTCATCGACGACCGACCCGACTTCTGACCCAC TGACTCCATCGACCCCGCCACCACCTGAGTGC GCGGGGTT	Extend PrΔ*gp1*
PrExtΔ*gp1*rv	GCCAGGCATTCGCTGCGGGTGTAGCCGCGCGC CTTGGCTTCGGCGATGGTGATTTGGGCGACTT GATCTTTC	Extend PrΔ*gp1*
Pr*lysA*His$_6$tag	AACCCTCGTGGACGCGGTAGCAGAACTGTTGG GCCACCACCACCACCACCACTGATGCGCATCG ACGGCCAATACGTCGGCCTCGGACC	Ms6 *lysA*His$_6$tag insertion
PrExt*lysA*His$_6$tagfwd	CGAGATCCTGCGGCAACTGCGCGGATACAACC TCACTGGCTGGCCGCAGCTCGGCGGAAAAACC CTCGTGGACGCGGTAGCAGAACTGTT	Extend Pr*lysA*His$_6$tag
PrExt*lysA*His$_6$tagrv	TAGGAGAACTTGCGCCGCATGAACGCTTTGAT CTTGCGGATCTCGTCGGATCTGTCCCCTGGTCC GAGGCCGACGTATTGGCCGTCGATG	Extend Pr*lysA*His$_6$tag
Pr*lysA*His$_6$tagfwd	GAACTGTTGGGCCACCACCACCACCAC	Ms6 *lysA*His$_6$tag [TAG]
Pr*lysB*rv	GATACCCCATGACGTACTTGGTTTCGGCG	Ms6 *lysA*His$_6$tag [FP]

[a]Underlined bases were added to provide additional restriction sites
[FP]Flanking primer.
[TAG]Primer specific to the tag sequence.

To construct plasmids pFG1 and pMP320, DNA fragments containing *gp1* or *gp1* and *lysA* genes were obtained by PCR with primers Pr*gp1*afwd/Pr*gp1*arv or Pr*gp1*afwd/Pr*lysA*brv, respectively. Primers were designed to generate restriction sites, and fragments were inserted in the corresponding sites of vector pQE30 (QIAGEN) creating a hexahistidine-tag fusion at the N-terminus. To obtain plasmids pMJC1, pMJC3 and pMJC4, DNA fragments containing *gp1*, *gp1* and *lysA* or *lysA* were amplified with primers Pr*gp1*bfwd/Pr*gp1*brv, Pr*gp1*bfwd/Pr*lysA*crv or Pr*lysA*cfwd/Pr*lysA*crv, respectively, and cloned into pET29b (Novagen) generating the following recombinant proteins: Gp1 with an N-terminal S-tag and C-terminal His$_6$-tag; Gp1 with an N-terminal S-tag and LysA with a C-terminal His$_6$-tag; LysA with an N-terminal S-tag and C-terminal His$_6$-

tag. pMJC2 was constructed by amplifying *gp1* with primers Pr*gp1*bfwd/Pr*gp1*crv, and inserting into pET29b to generate a S-Gp1 recombinant protein not fused to a C-terminal His_6-tag. pMJC5 was constructed by amplifying *gp1* and *lysA* with primers Pr*gp1*afwd/Pr*lysA*brv and cloning into pQE-80L (QIAGEN), allowing expression of Gp1 with a His_6-tag at the N-terminus. pMJC9 was constructed in two steps: *gp1* was amplified with primers Pr*gp1*afwd/Pr*gp1*drv and cloned into the BamHI/SacI sites of pQE30, followed by cloning lambda *R* (*λR)* amplified from phage λgt11 with primers Prλ*R*fwd/Prλ*R*rv, and cloning into the SacI/HindIII sites of pQE30:*gp1*. pMJC10 was constructed the same way, except that *gp1* was amplified with primers Pr*gp1*bfwd/Pr*gp1*drv and the cloning vector used was pET29b. To obtain pMJC14 and pMJC15, *lysA* deleted for the first 180 bp (Δ60aaLysA) was amplified with primers PrΔ180bp*lysA*fwd/Pr*lysA*brv, or PrΔ180bp*lysA*fwd/PrlysAcrv, and cloned into the SacI/HindIII sites of pQE30 and pET29b, respectively. In a second step, *gp1* was amplified with primers Pr*gp1*afwd/Pr*gp1*erv or Pr*gp1*bfwd/Pr*gp1*erv and cloned into the BamHI/SacI sites of pQE30:Δ180bp*lysA* or pET29b:Δ180bp*lysA*, respectively. *gp1* and the first 180 bp of *lysA* were amplified with primers Pr*gp1*bfwd/Pr*lysA*180bprv and cloned into the BamHI/HindIII sites of pET29b for pMJC17.

To express Ms6 lysis genes in *M. smegmatis*, genes were cloned downstream of the acetamidase promoter (Parish *et al.*, 1997) in vector pVVAP, which contains a pAL5000 replication origin, a kanamycin resistance cassette, and allows expression of C-terminally His_6-tagged fusion proteins. For optimal expression in mycobacteria, the consensus ribosomal-binding site (RBS) and ATG/GTG start codons of the first coding sequence were provided by the forward primers. Plasmids pMJC6 and pMJC7 were obtained by cloning *gp1* or *gp1* and *lysA* fragments amplified with primers Pr*gp1*dfwd/Pr*gp1*brv or Pr*gp1*dfwd/ Pr*lysA*crv into pVVAP. To construct pMJC8, *lysA* was amplified with primers Pr*lysA*dfwd/Pr*lysA*crv and directionally inserted into pVVAP. Plasmid pMJC11

was obtained by amplifying *gp1*, *lysA*, *lysB*, *hol* and *gp5* genes with primers P*gp1*dfw/P*gp5*rv and cloning in pVVAP, allowing the expression of the entire Ms6 lysis module in *M. smegmatis*. To construct a LysA-PhoA′ hybrid protein, pSMT3-[19-phoA] (Herrmann *et al.*, 1996) was used as a template to amplify the *phoA* gene omitting the PhoA signal peptide, using primers Pr*phoA*'BamHIfwd/ Pr*phoA*'HindIIIrv and generating the PCR fragment *phoA′*. The genomic DNA of Ms6 corresponding to *gp1* and *lysA*, or *lysA* only were PCR amplified using the forward primer Pr*gp1*pMP201fwd which includes the RBS and the ATG start codon of *gp1*, or the forward primer Pr*lysA*pMP201fwd, which includes the RBS and GTG start codon of *lysA*, both carrying a HindIII restriction site, and a reverse primer Pr*lysA*-SGGGSrv carrying the SGGGS 3′ linker and a BamHI restriction site. The PCR products were digested with BamHI, and after ligation, *lysA:phoA′* and *gp1lysA:phoA′* were digested with HindIII and cloned into the same site of pMP201, generating pMJC12 and pMJC13, respectively. Plasmid pMP201 (Pimentel, unpublished data) is a derivative of the integrative plasmid pRM16 (Freitas-Vieira *et al.*, 1998) that allows gene expression under control of the Ms6 *Plys* promoter. All constructs were verified for the insert nucleotide sequence.

Protein interaction

Crosslinking

BS^3 (Sigma) crosslinker stock solution was prepared immediately before use to decrease the extent of hydrolysis. BS^3 was dissolved in 20 mM Na-HEPES, 200 mM NaCl pH 7.0 at 10 mM final concentration. For *in vitro* crosslinking, *E. coli* BL21 (DE3) carrying pMJC1, pMJC3, pMJC4, pMJC10, pMJC15 or pMJC17 were induced at logarithmic growth phase with 1 mM IPTG, and 10 ml samples pelleted after 1 h. Cells were resuspended in phosphate-buffered saline (PBS), sonicated and centrifuged at 4 ºC. Supernatant was treated with BS^3 to a final

concentration of 1-5 mM at room temperature for 30 min. In the control samples crosslinker was omitted. Cells were centrifuged and resuspended in SDS-PAGE sample buffer to quench the reaction, and subjected to SDS-PAGE. Gp1, LysA or λR proteins were detected by Western-blotting using HRP-conjugated anti-His monoclonal antibody (Roche). To examine the interaction of Gp1 and LysA during Ms6 infection, an exponential culture of *M. smegmatis* was infected with Ms6 phage at a multiplicity of infection (m.o.i.) of 10. Two hours after infection, cells were centrifuged and concentrated 100-fold in PBS with 20 mg of lysozyme ml^{-1}. After an incubation period at 37 ºC for 30 min, SDS-PAGE sample buffer was added, followed by incubation at 100 ºC for 5 min. Crosslinking was performed as described previously for *E. coli* samples. For crosslinking purified Gp1, soluble S-Gp1-His_6 was purified by passage through a Ni-NTA column (QIAGEN), according to the manufacturer's instructions. Eluted fractions were dialyzed against PBS for 24 h at 4 ºC and 10 μg of S-Gp1-His_6 treated with 1-5 mM BS^3 for 30 min at room temperature, before SDS-PAGE and Western-blotting.

Pulldown

To test the interaction between Gp1 and LysA, we used an *in vitro* pull-down technique with Dynal Dynabeads Talon (Invitrogen). S-Gp1 and LysA-His_6 were co-expressed in *E. coli* BL21 (DE3) carrying pMJC3, by growing in LB broth supplemented with kanamycin to mid-logarithmic phase and inducing with 1 mM IPTG for 1 h. Cells were centrifuged, resuspended in 50 mM Tris-HCl, pH 8.0 and sonicated. Lysates were centrifuged at 8000 x *g* for 20 min. Recovered supernatant was incubated with 50 μl (2 mg) His-affinity Dynabeads at 4 ºC for two hours. Beads were washed four times with binding and washing buffer (PBS, pH 8.0, 300 mM NaCl and 0.01% Tween 20). To eliminate non-specific interactions, a negative control was performed in the same conditions using a His_6 untagged recombinant protein (S-Gp1) expressed in *E. coli* BL21 (DE3):pMJC2. All samples were

resuspended in SDS sample buffer, and boiled to elute proteins off the beads. Proteins were resolved on SDS-PAGE gels, and Gp1 and LysA were detected by immunoblot analysis as described.

Alkaline phosphatase activity assay

Cells from 1 ml exponential *M. smegmatis* cultures were centrifuged and resuspended in 900 µl of 1 M Tris-HCl, pH 8.0 and 100 µl of the substrate solution consisting of 0.2 M *p*-nitrophenol phosphate (Sigma) were added. Reactions proceeded at 37 ºC until a yellow colour developed, and stopped with 100 µl of 1 M K_2HPO_4. Cells were centrifuged at 12,000 x *g* for 2 min, and the supernatant read at 405 nm. Enzyme activity was expressed in arbitrary units of OD_{405} ml^{-1} of culture min^{-1}. For complementation in *trans*, *M. smegmatis*:pMJC12 was electroporated with pMJC6. Presence of *gp1* and *lysAphoA´* in kanamycin-resistant transformants was confirmed by PCR with specific primers. Exponential cultures were induced with 0.2% acetamide for 6 hours and collected by centrifugation. Alkaline phosphatase activity was determined as above.

Construction of Ms6 mutant phages

Construction of Ms6 mutant phages was performed using the Bacteriophage Recombineering of Electroporated DNA (BRED) technique in *M. smegmatis* as described previously (Marinelli *et al*., 2008). Briefly, for Ms6 *gp1* gene deletion, a 100 bp oligonucleotide (PrΔ*gp1*) with 50 bp of homology upstream and downstream of the deleted region was extended by PCR using two 75-bp extender primers, PrExtΔ*gp1*fwd/PrExtΔ*gp1*rv which have 25 bp of homology to the ends of the 100-mer, and add an additional 50 bp of homology on either end. For the His_6-tag insertion, an 89-bp oligonucleotide (Pr*lysA*His_6tag) was extended with primers PrExt*lysA*His_6tagfwd/PrExt*lysA*His_6tagrv. The final 200-bp dsDNA

products were purified using a MinElute PCR Purification Kit (QIAGEN) and co-electroporated with Ms6 genomic DNA into electrocompetent recombineering *M. smegmatis* mc^2155:pJV53. Cells were resuspended in 7H9 with 0.5% glucose and 1 mM $CaCl_2$, incubated at 37 ºC for 2 hours and plated as top agar lawns with *M. smegmatis* mc^2155. Phage plaques were picked into 100 µl phage buffer (10 mM Tris-HCl, pH 7.5; 10 mM $MgSO_4$; 68.5 mM NaCl; 1 mM $CaCl_2$), eluted for two hours at room temperature and analysed by PCR with primers PrP1fwd/ Pr*lysA*180bprv flanking the *gp1* deletion, or with Pr*lysA*His_6tagfwd /Pr*lysB*rv to detect the His_6tag insertion. Mixed primary plaques containing both the deletion mutant and wild-type DNA were eluted as described above, and serial dilutions were plated with *M. smegmatis*. Individual secondary plaques or lysates were screened by PCR for the presence of pure *gp1* deletion- or *lysA*His_6tag insertion-mutant phages.

One-step growth curves and burst size determination

One-step growth curves and burst-size determination assays were adapted from Adams (1959). One-step assays were carried out in exponential cells using an m.o.i. of 1. *M. smegmatis* cells were pelleted and resuspended in 1 ml phage suspension ($Ms6_{wt}$ or $Ms6_{\Delta gp1}$) supplemented with 1 mM $CaCl_2$. The mixture was incubated 50 min at 37 ºC for phage adsorption, and then 100 µl of 0.4% H_2SO_4 was added for five min to inactivate non-adsorbed phages. The suspension was neutralized with 100 µl of 0.4% NaOH and diluted 1:100 in 7H9 with 0.5% glucose and 1 mM $CaCl_2$. One-ml samples were taken every 30 min for 240 min and 100 µl of serial dilutions of each sample were plated with 200 µl of *M. smegmatis* cells on 7H10, as top agar lawns. Phage titer for each sample was determined after 24 h incubation at 37 ºC. The same procedure was used for burst size determination except that 10 µl of infected cells were diluted in 7H9 to obtain ≤ one infected cell

ml^{-1}. Samples of 1 ml infected culture were distributed in 50 tubes and incubated for 120 min at 37 ºC. 200 µl of *M. smegmatis* cells and top agar were added to each tube before plating on 7H10. After 24 h at 37 ºC, phage plaques were counted, and *Poisson* [P(*n*)] distribution applied to determine burst-size (BS): P (n) = (e^{-c} . c^n)/ n! (e < 1), where P (n) is the probability of samples with n infected cells, and ***c*** is the average number of infected cells per tube. (BS) = (total plaque count in the 50 plates)/(total number of infected cells).

Results

Gp1 expression induces* E. coli *lysis in the presence of Ms6 endolysin

To understand how each Ms6 lytic gene contributes to lysis, we have previously constructed a series of recombinant strains expressing different combinations of the five genes. As expected, expression of the entire Ms6 lysis module in *E. coli*, under the control of the T5 promoter, resulted in cell lysis (data not shown). Surprisingly, expression of the endolysin in the presence of *gp1* gene was sufficient to induce *E. coli* lysis, beginning 20 minutes after induction, even in the absence of the holin gene (Fig. 1). This phenotype was not observed with the endolysin expression alone, or in conjunction with the holin, unless the plasma membrane was permeabilized with chloroform (Garcia *et al.*, 2002). When *gp1* was expressed alone, growth inhibition, but not cell lysis was observed (Fig. 1), suggesting that Gp1 overexpression is toxic to cells. Toxic effects are observed when membrane proteins are overexpressed, as it happens with holins (Smith *et al.*, 1998). To confirm that the observed lysis phenotype was not a result of cytoplasmic membrane damage that allowed the Ms6 endolysin to reach the peptidoglycan, we constructed a recombinant plasmid expressing Gp1, with λ endolysin (λR) replacing the Ms6 endolysin. As shown in Fig. 1, after IPTG

induction, no lysis was observed, indicating that lysis was not a consequence of nonspecific damage of the cytoplasmic membrane by Gp1.

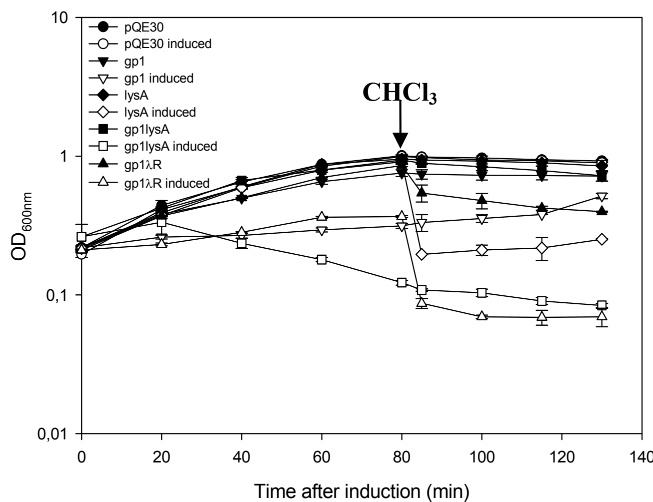


Figure 1. Effect of Ms6 lysis genes on *E. coli* growth. *E. coli* JM109 carrying plasmid pQE30 with no insert or cloned genes was grown in LB broth at 37 ºC to OD_{600nm} 0.2. At time zero, transcription of cloned lysis genes, was induced with 1 mM IPTG. Culture turbidity was monitored at 600 nm. At the time indicated by the arrow, 2% $CHCl_3$ was added to cultures. Plotted data are the means of three independent experiments, and error bars show standard deviation.

gp1 is not a holin-like protein

gp1, the first gene of the Ms6 lytic cassette is located immediately upstream of, and overlapped with the endolysin gene (Fig. 2). It encodes a small, 77-amino acid protein with a predicted molecular mass of 8.3 kDa and an acidic isoelectric point of 4.6. No Gp1 homologues were identified by BLASTp search, except for corresponding putative proteins with unknown function encoded by other mycobacteriophages. The lysis phenotype resulting from co-expression of Ms6 Gp1 and endolysin suggested that *gp1* might encode an additional Ms6 holin.

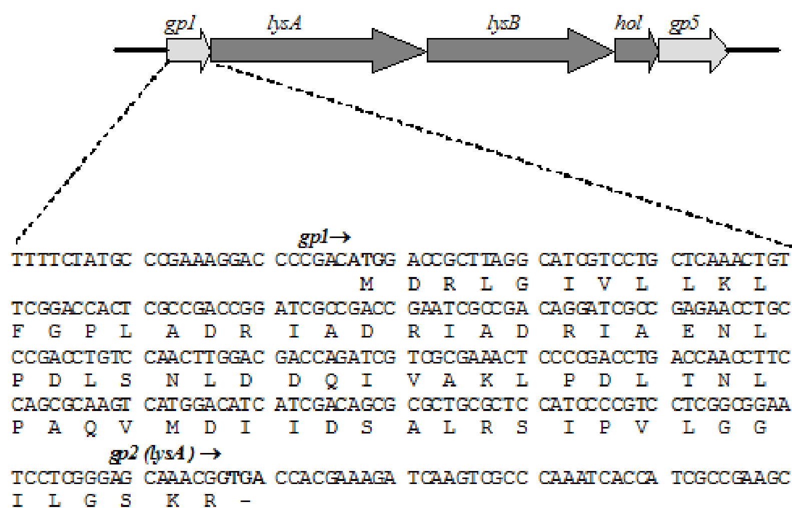


Figure 2. Organization of the mycobacteriophage Ms6 lysis module. Previously characterized genes (Garcia *et al.*, 2002) are shown by dark arrows. In the nucleotide sequence, overlapped start/stop codons of *lysA* and *gp1* are highlighted. Ms6 lysis genes: accession No. AF319619.

Although holins do not exhibit significant similarity at the amino acid level, they share well-defined features: a hydrophilic, highly charged C-terminal sequence, and at least one N-terminal transmembrane domain (Young *et al.*, 2000). We did not find these characteristics in Gp1 amino acid sequence; the N- and C-termini are hydrophobic, with a hydrophilic central region. No transmembrane regions were found using the prediction programs HMMTOP 2.0 or Top-Pred2 from the Expasy Proteomics Server of the Swiss Institute of Bioinformatics (http://www.expasy.ch). To further exclude the holin-like activity of Gp1, the nonsuppressing *E. coli* strain K802, harbouring a pQE30 empty vector or a derived plasmid containing *gp1* (pFG1) was infected with λgt11, a lambda mutant with an amber mutation in gene *S* (S100) that consequently can not induce lysis of the infected host, except in suppressing *E. coli* strains. The number of plaques forming

units (p.f.u.) produced by *E. coli* K802:pFG1 (9×10^5 p.f.u. ml^{-1}) was not higher than *E. coli* K802:pQE30 (2×10^5 p.f.u. ml^{-1}), but was significantly lower than the p.f.u. produced by the suppressing strain *E. coli* Y1088 (1.3×10^{10} p.f.u. ml^{-1}). Therefore Ms6 Gp1 did not complement the defective S100 allele of phage λgt11, and, unlike Ms6 holin which efficiently complements a λS defect (Garcia *et al.*, 2002), did not allow the access of the lambda transglycosylase (λR) to the murein. These data, together with the bioinformatic analysis, indicate that the Gp1 protein does not belong to the holin class of proteins.

Gp1 has several features of a chaperone-like protein

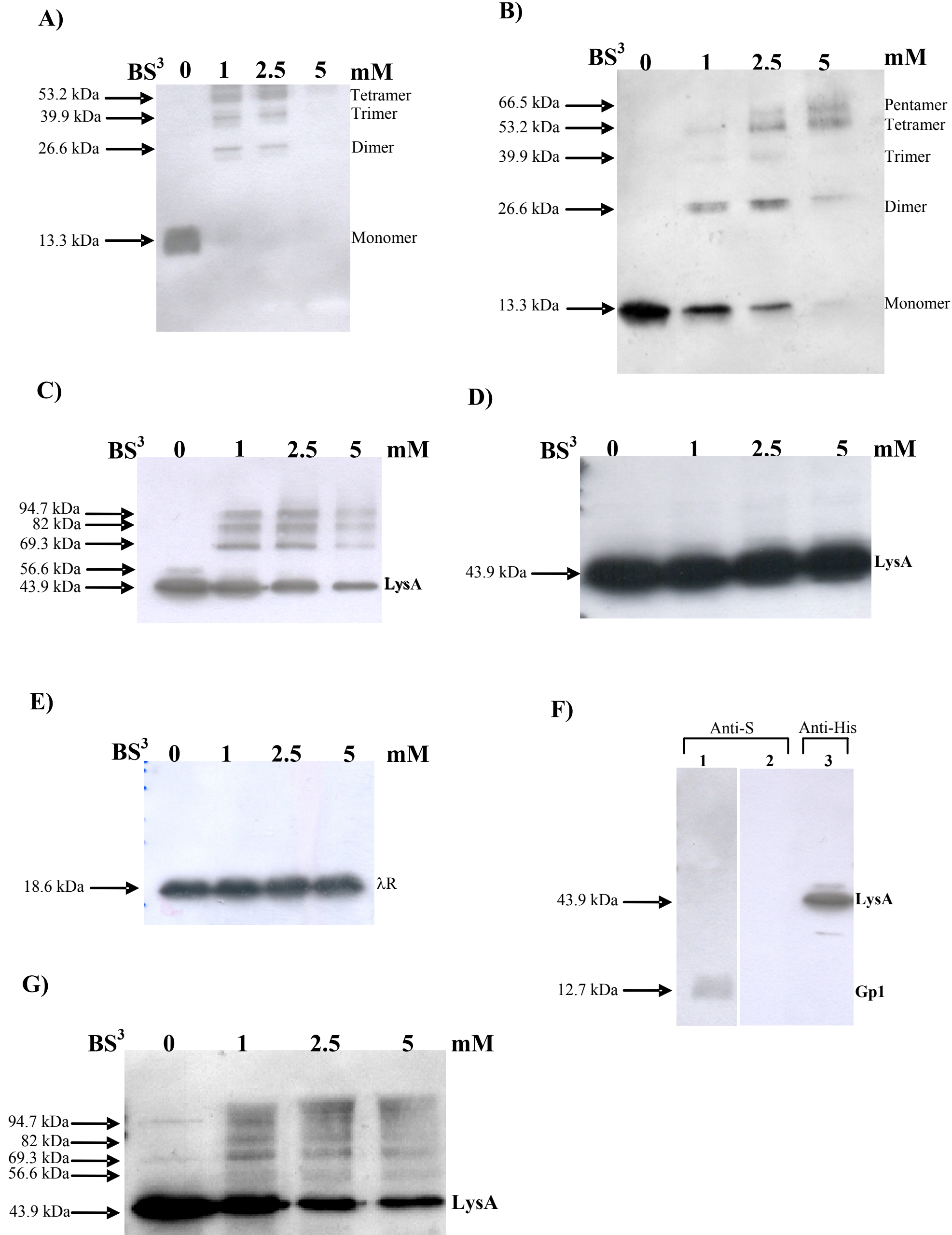
Several physical and predicted structural characteristics of Gp1 are consistent with chaperones, particularly type III secretion (TTS) system chaperones. Although they do not exhibit amino acid sequence similarities, they have common physical characteristics, including a low molecular weight (<15 kDa), an acidic pI (<5), and a predicted amphipathic helix near the C-terminus. TTS chaperone-genes are generally adjacent to the genes encoding their cognate effectors (Page and Parsot, 2002; Feldman and Cornelis, 2003). One of the main features of the molecular secretion chaperones described so far is their ability to dimerize in order to exert their biological functions, as shown for the most well-studied secretion chaperone, SecB (Zhou and Xu, 2005) and for the TTS chaperones (Feldman and Cornelis, 2003). We previously observed SDS-resistant Gp1 dimers (Catalão, unpublished data). The resistance of Gp1 dimers to boiling in SDS and β-mercaptoethanol treatment suggests strong hydrophobic interactions. We performed a crosslinking experiment with BS^3, a water-soluble, membrane-impermeant, homobifunctional sulfo-*N*-hydroxy-succinimide ester, to determine whether Gp1 can dimerize or oligomerize. BS^3 forms covalent amide bonds with accessible α-amine groups on the N-termini of proteins and ε-amines on lysine residues, and is used to identify transient interactions within the cells. Ideally, the

concentration of the crosslinker reagent is determined by the amount of each protein to be linked, but as we performed crosslinking with cell extracts, the optimal reagent concentration was empirically determined. The experiment was performed with concentrations of BS^3 ranging from 1 mM to 5 mM, added to the supernatant of sonicated *E. coli* cells expressing a Gp1 recombinant protein with an S-tag at the N-terminus, and a hexahistidine tag at the C-terminus (S-Gp1-His_6). Fig. 3A shows that, in the absence of crosslinker, only S-Gp1-His_6 monomers were detected, but as the concentration of BS^3 increased, a decrease in monomeric Gp1 was observed with a concomitant increase in higher molecular-mass oligomers. By Western-blotting, the molecular masses of crosslinked species detected with anti-His antibody were consistent with the sizes predicted for recombinant S-Gp1-His_6 monomers (13.3 kDa), dimers (26.6 kDa), trimers (39.9 kDa), and tetramers (53.2 kDa). A similar interaction pattern was obtained when crosslinking was performed with purified S-Gp1-His_6 (Fig. 3B), so we concluded that all crosslinked forms resulted from the Gp1 protein interacting with itself, and not with other cellular proteins, or as a result of overexpression.

Gp1 interacts with Ms6 endolysin

Considering the above results and the overlapping gene arrangement of *gp1* and *lysA* in the Ms6 genome, we tested the possibility that Gp1 interacts with LysA by expressing both proteins in *E. coli* under the control of the T7 promoter, with Gp1 S-tagged at the N-terminus (S-Gp1), and LysA His_6-tagged at the C-terminus. We detected a similar crosslinking pattern at all BS^3 concentrations used. Three bands with molecular masses higher than LysA-His_6 (43.9 kDa) were detected, corresponding to a 69.3 kDa S-Gp1 dimer and LysA-His_6 monomer, a 82 kDa S-Gp1 trimer and LysA-His_6 monomer, and a 94.7 kDa S-Gp1 tetramer and LysA-His_6 monomer (Fig. 3C). Interestingly, even in the absence of crosslinker, we observed a 56.6 kDa band corresponding to an S-Gp1 monomer and a LysA-His_6

monomer. Given the denaturing SDS-PAGE conditions, this indicates a strong hydrophobic interaction between the two proteins. LysA-His$_6$ did not interact with itself since no oligomers were detected with an anti-His antibody when the protein was expressed alone (Fig. 3D). To exclude the possibility that the observed interaction resulted from overexpression, the S-Gp1 protein was co-expressed with the phage lambda endolysin (λR) carrying a C-terminal His$_6$-tag. No interaction was detected, suggesting that the interaction between Gp1 and LysA was specific, and not a consequence of protein overexpression (Fig. 3E). To confirm the Gp1-LysA interaction seen in the crosslinking experiments, we performed an *in vitro* pulldown assay to determine physical interaction between the proteins. S-Gp1 and LysA-His$_6$ were co-expressed in *E. coli,* and cell lysate supernatants applied to His-affinity Dynabeads. Gp1 interacted with LysA, as detected with anti-S-protein (Fig. 3F, lane 1). No protein was detected when S-Gp1 was expressed alone from pMJC2 excluding the possibility of non-specific interaction of Gp1 with the beads (Fig. 3F, lane 2). The successful bead capture of LysA-His$_6$ was confirmed with an anti-His antiboby (Fig. 3F, lane 3). These data confirmed the interaction between Gp1 and LysA. To investigate *in vivo* interactions during Ms6 phage infection of *M. smegmatis,* we used the bacteriophage recombineering of electroporated DNA (BRED) system developed by Marinelli *et al*. (2008) to construct a mutant Ms6-LysAHis$_6$ phage by fusing the 3' end of *lysA* to a hexahistidine tag. *M. smegmatis* was infected with Ms6-LysAHis$_6$ phage at a multiplicity of infection (m.o.i.) of 10. At two hours after infection, and before lysis occur, cells were treated with lysozyme and the supernatant extracted for crosslinking with BS^3 at the concentrations used above. As shown in Fig. 3G, crosslinking patterns were similar to those observed in *E. coli*, indicating specific interaction during Ms6 infection of mycobacteria.



A)
BS3 0 1 2.5 5 mM
53.2 kDa
39.9 kDa
26.6 kDa
13.3 kDa
Tetramer
Trimer
Dimer
Monomer
B)
BS3 0 1 2.5 5 mM
66.5 kDa
53.2 kDa
39.9 kDa
26.6 kDa
13.3 kDa
Pentamer
Tetramer
Trimer
Dimer
Monomer
C)
BS3 0 1 2.5 5 mM
94.7 kDa
82 kDa
69.3 kDa
56.6 kDa
43.9 kDa
LysA
D)
BS3 0 1 2.5 5 mM
43.9 kDa
LysA
E)
BS3 0 1 2.5 5 mM
18.6 kDa
λR
F)
Anti-S
Anti-His
1 2 3
43.9 kDa
12.7 kDa
LysA
Gp1
G)
BS3 0 1 2.5 5 mM
94.7 kDa
82 kDa
69.3 kDa
56.6 kDa
43.9 kDa
LysA

Figure 3. Expression of different constructs of Gp1, LysA and analysis of their interactions. Crosslinking with different BS^3 concentrations. *E. coli* lysates with S-Gp1-His_6 alone **(A)**; purified S-Gp1-His_6 **(B)**; S-Gp1 and LysA-His_6 **(C)**; LysA-His_6 alone **(D)**; S-Gp1 and λR-His_6 **(E)**. Proteins were analysed by Western-blotting with anti-His antibody. **F)** Interaction of S-Gp1 with LysA-His_6 by *in vitro* pulldown; lane 1: immobilized S-Gp1 detected with anti S-protein; lane 2: no proteins detected with anti-S from control beads treated with S-Gp1 only; lane 3: immobilized LysA-His_6 detected with anti-His. **G)** Lysate of *M. smegmatis* cells infected with Ms6-LysAHis_6 phage, crosslinked with BS^3. Proteins were analysed by Western-blotting with anti-His antibody.

Gp1 binding requires the N-terminal 60 amino acids of LysA

To determine the region of LysA that binds to the Gp1 chaperone-like protein, progressive truncations of the LysA N-terminus were constructed, and interaction patterns examined by crosslinking. Loss of the first 60 amino acids of LysA (LysAΔ_{1-60}) was sufficient to abrogate interaction between Gp1 and endolysin (Fig. 4A). To confirm this result, we performed a crosslinking experiment using S-Gp1 and the first 60 amino acids of LysA (LysAΔ_{61-384}) tagged with a C-terminal His_6. The interaction pattern was consistent with the crosslinking between S-Gp1 and full-length LysA-His_6. The three bands with molecular masses greater than LysAΔ_{61-384} (7.7 kDa) corresponded to interaction between a LysAΔ_{61-384} monomer and a Gp1 dimer (33.1 kDa), trimer (45.8 kDa), and tetramer (58.5 kDa) (Fig. 4B). These data demonstrated that the first 60 amino acids of Ms6 LysA were necessary and sufficient for interaction with Gp1. Once more, in the absence of crosslinker, we observed a band of 20.4 kDa corresponding to the interaction of an S-Gp1 monomer and a LysAΔ_{61-384}-His_6 monomer, indicating strong hydrophobic interaction. Collectively, these results indicate that LysAΔ_{61-384}, but not LysAΔ_{1-60} interacts with Gp1. When both Gp1 and LysAΔ_{1-60} were co-expressed in *E. coli*, no lysis phenotype was observed, unless the plasma membrane

was permeabilized with chloroform (Fig. 4C), indicating that LysAΔ_{1-60} retained the ability to exert its enzymatic activity, but could not access the peptidoglycan even with Gp1 co-expression.

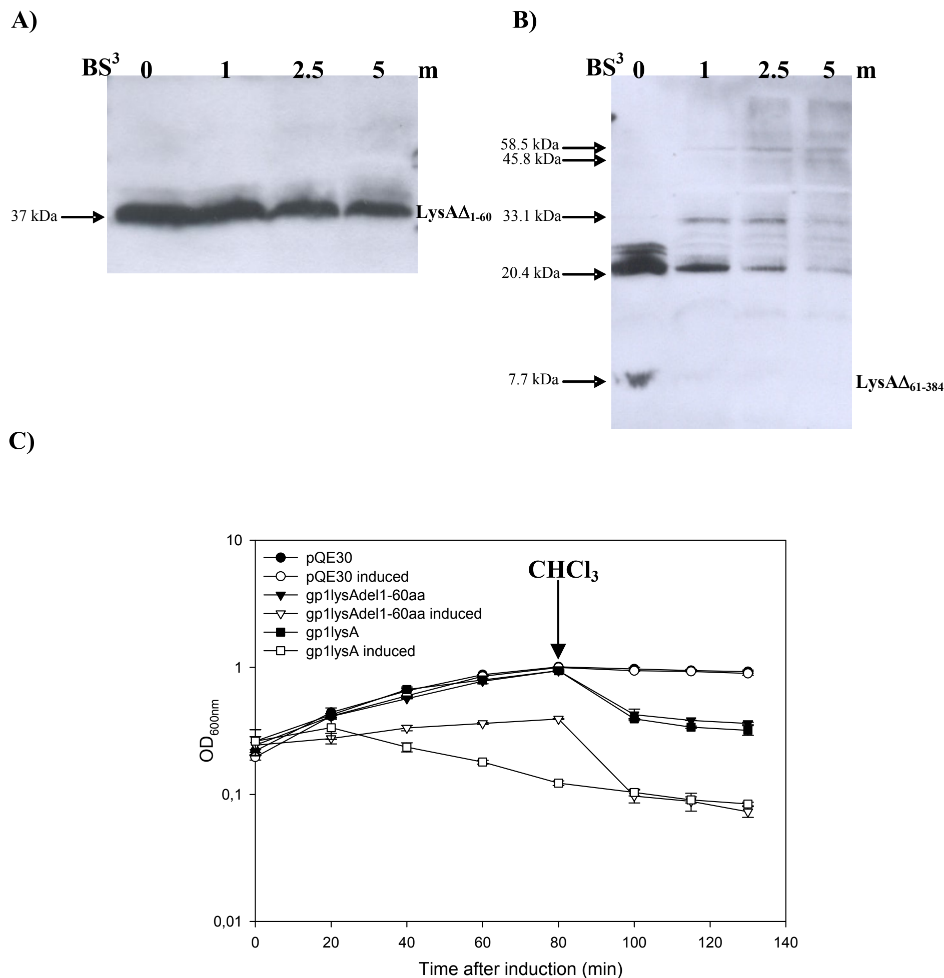


Figure 4. Interaction of different regions of His-tagged LysA with Gp1. Crosslinking with BS^3 of *E. coli* lysates expressing S-Gp1 and **A)** LysAΔ_{1-60} or **B)** LysAΔ_{61-384}. Proteins were detected with anti-His antibody. **C)** Effect on *E. coli* growth. Growth and induction conditions as described for Fig. 1.

Gp1 allows translocation of a LysA-PhoA fusion across the* M. smegmatis *membrane

Based on the above results, we tested whether Gp1 is involved in endolysin transport across the inner membrane. *E. coli* alkaline phosphatase (PhoA) is widely used as a reporter protein for studying protein export in bacteria, as it is enzymatically active only after export into the periplasmic space. If a gene is fused to *E. coli phoA* lacking its own signal peptide sequence (*phoA*'), only replacement sequences with protein export signals promote transfer across the cytoplasmic membrane, resulting in PhoA activity (Manoil and Beckwith, 1986). This approach has also been used to identify protein export sequences in *M. smegmatis* (Timm *et al.,* 1994). Using this strategy, we investigated whether Gp1 protein, which appears to be necessary for localizing the endolysin LysA to the extracytoplasmic compartment, would be sufficient for transport of a LysA-PhoA' fusion. We constructed two integrative plasmids derived from pMP201 (see Experimental Procedures) carrying a *lysAphoA' fusion*. pMJC12 expressed the hybrid protein LysA-PhoA', and pMJC13 expressed Gp1 and the LysA-PhoA'. In *M. smegmatis,* pMJC13 expression yielded blue colonies in media supplemented with the chromogenic phosphatase indicator 5-bromo-4-chloro-3-indoxyl phosphate (BCIP), while cells carrying pMP201:*lysAphoA'* (pMJC12) were colourless, suggesting that PhoA' was active only when fused to LysA and in the presence of Gp1 (data not shown). Moreover, *M. smegmatis* whole cells with pMJC13 showed higher alkaline phosphatase activity (71 U) compared to whole cells with pMJC12 (16.7 U), or empty vector pMP201 (9.7U). In addition, expression of Ms6 Gp1 in *trans* increased LysA-PhoA' fusion export*,* since *M. smegmatis*:pMJC12 expressing Gp1 protein from pMJC6 showed 68 U of alkaline phosphatase activity. These data indicate that Gp1 is necessary for translocation of LysA across the *M. smegmatis* cell membrane.

The Sec-system of E.* coli *is involved in endolysin translocation

The Sec-dependent export pathway is highly conserved among bacteria and is responsible for the export of many newly synthesized outer membrane and periplasmic proteins. Gp1 export of LysA using the host Sec-system was tested in *E. coli* by examining holin-independent lysis in the presence of sodium azide, a potent inhibitor of SecA, which is essential for protein export across the *E. coli* plasma membrane (Oliver *et al*., 1990). As shown in Fig. 5A, no lysis was detected in *E. coli* carrying the wild-type (*wt*) *sec*A locus expressing Gp1 and LysA from pMP320, in the presence of 1 mM sodium azide. The same result was obtained in a *secA51* temperature-sensitive (ts) mutant strain that displays pleiotropic defects in protein export (Oliver and Beckwith, 1981). A lysis phenotype was observed at the permissive temperature (Fig. 5B) but not at the non-permissive temperature (Fig. 5C). These results indicate that LysA transport across the *E. coli* cytoplasmic membrane is SecA-dependent.

Nisin triggers* M. smegmatis *lysis in Gp1-dependent manner

Recent research on the lambdoid phage 21 suggests that some holins, named pinholins, make holes too small to allow the passage of phage endolysins from the cytoplasm to the periplasm. Pinholins do not play a decisive role in endolysin export, and their function seems to be confined to membrane depolarization (Park *et al*., 2007) which serves as a signal for activation of secreted endolysin (Nascimento *et al*., 2008). Considering the need for the auxiliary lysis protein Gp1 to achieve lysis, we investigated if membrane-depolarizing agents triggered LysA-mediated lysis in the Ms6 mycobacterial host *M. smegmatis*, using nisin, a bacteriocin produced by *Lactococcus lactis* (Bierbaum and Sahl, 1987). Nisin depletes ATP and p.m.f. in mycobacteria by acting on the cytoplasmic membrane (Montville *et al*., 1999; Chung *et al*., 2000), forming small pores

(Wiedemann et *al.*, 2004; Breukink, 2006), and leading to efflux of low-molecular weight compounds (Ruhr and Sahl, 1985).

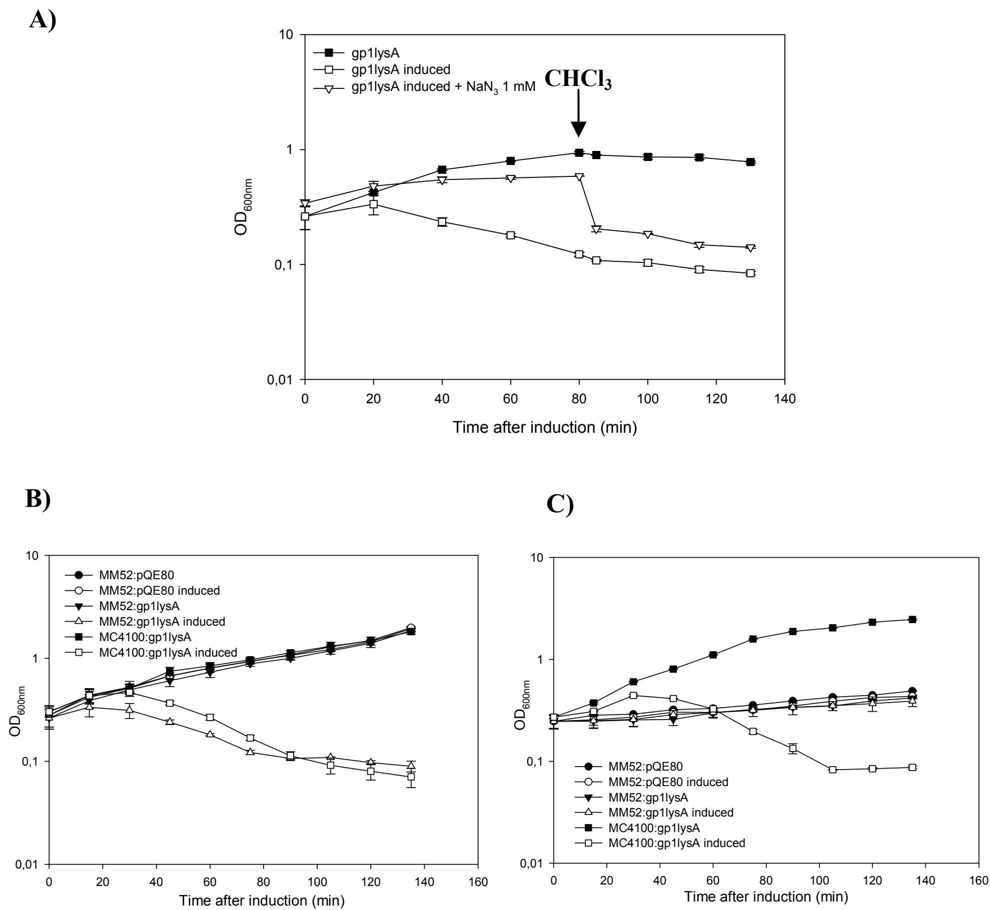


Figure 5. Gp1-LysA-mediated lysis requires the *E. coli* Sec-system. *E. coli* was grown to OD_{600nm} 0.2, and expression induced with 1 mM IPTG. **A)** *E. coli* JM109:pMP320; at the time indicated by the arrow, 2% $CHCl_3$ was added. To one culture (▽), 1 mM sodium azide was added 10 min before induction. **B)** Recombinant strains were grown at 30 ºC. At 90 min before induction, strains in **C)** were shifted from 30 ºC to 42 ºC. Data are the means of three independent experiments, and error bars show standard deviations.

As proposed by Nascimento *et al*. (2008), nisin mimics the holin disruption of the cytoplasmic membrane electrical and chemical gradients, triggering events that lead to endolysin activity on the peptidoglycan. In contrast to the complete bacterial lysis observed in *E. coli*, co-expression of Gp1 and endolysin in *M. smegmatis* caused only partial lysis. However, addition of 25 ng ml^{-1} nisin to *M. smegmatis* expressing Gp1 and LysA, at two hours after induction, resulted in complete lysis (Fig. 6). No lysis was observed when nisin was added to *M. smegmatis* expressing the endolysin or Gp1 alone (Fig. 6). This strongly suggests that the endolysin was already positioned next to its target, the peptidoglycan, at lysis onset, as the pore diameter produced by nisin (2 nm) should not allow the passage of a protein as large as Ms6 endolysin (Ruhr and Sahl, 1985). Accordingly, we conjecture that Ms6 holin needs only to depolarize the mycobacterial membrane to accomplish its role in controlling the timing of lysis.

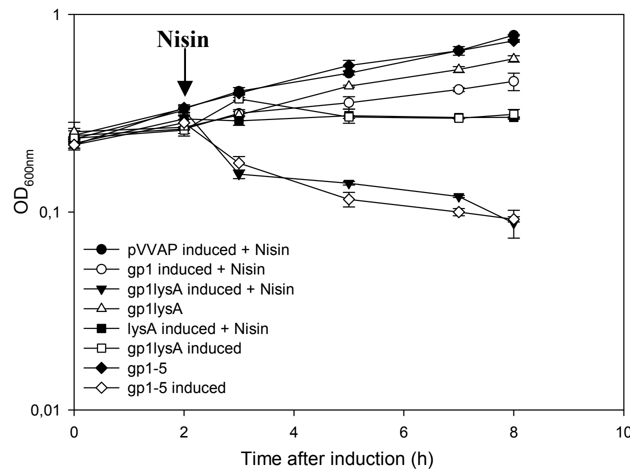


Figure 6. Gp1-LysA lysis is triggered by nisin membrane depolarization. *M. smegmatis* mc^2155 cells carrying plasmid pVVAP with no insert, cloned *gp1*, *gp1* and *lysA*, *lysA* or the five lytic genes were grown at 37 ºC. Two hours after acetamide induction, 25 ng ml^{-1} nisin was added to the indicated cultures (●, ○, ▼ and ■). Error bars show standard deviations of means from three independent experiments.

Gp1 is essential for efficient lysis of* M. smegmatis, *but not for plaque formation

To determine if Gp1 is essential for plaque development by Ms6, we tested whether it is possible to delete *gp1* from the Ms6 genome and obtain a viable phage. Using the BRED system (Marinelli *et al.,* 2008), we constructed a 213 bp internal in-frame deletion of Ms6 *gp1* by allelic gene replacement. Co-electroporation of a 200 bp substrate containing 100 bp flanking *gp1*, and Ms6 genomic DNA into a *M. smegmatis* recombineering strain yielded 55 primary plaques, 25 of which were screened by PCR with primers flanking the deletion. Two contained detectable levels of mutant phage mixed with wild-type phage (Fig. 7A). Serial dilutions of the mixed plaques were re-plated, and both the secondary lysate and individual secondary plaques were screened by PCR with the same primers. Since the mutant allele was still present in the lysate (Fig. 7B), the mutant was viable. To recover *gp1* deletion mutants, 25 secondary plaques from the re-plating were picked and PCR-screened with primers flanking the deletion. Four homogenously pure mutants were recovered at high efficiency (16%), in the absence of selection (Fig. 7B). These results demonstrated that $Ms6_{\Delta gp1}$ is viable, and that Gp1 is not required for plaque formation. However, the mutant phage produced smaller plaques lacking the size variation produced by $Ms6_{wt}$ phage (Fig. 8A, panel 1 and 2). Defective $Ms6_{\Delta gp1}$ phage was complemented by infecting a *M. smegmatis* strain expressing wild-type Gp1 under the control of the acetamidase promoter (*M. smegmatis*:pMJC6). Ms6 Gp1 provided in *trans* restored the wild-type phenotype in induced conditions (Fig. 8A, panel 3), but in uninduced conditions (Fig. 8A, panel 4), the plaques remained very small, and showed no size variation. One step growth curves and determination of phage growth parameters (latent period, rise period and burst size) were carried out to compare the phages multiplication cycles. The one step experiment (Fig. 8B) showed that both phages have a 110-minute latent period corresponding to the 60 minutes represented in Fig. 8B, plus the initial 50 min of adsorption, and a rise period of 30 minutes.

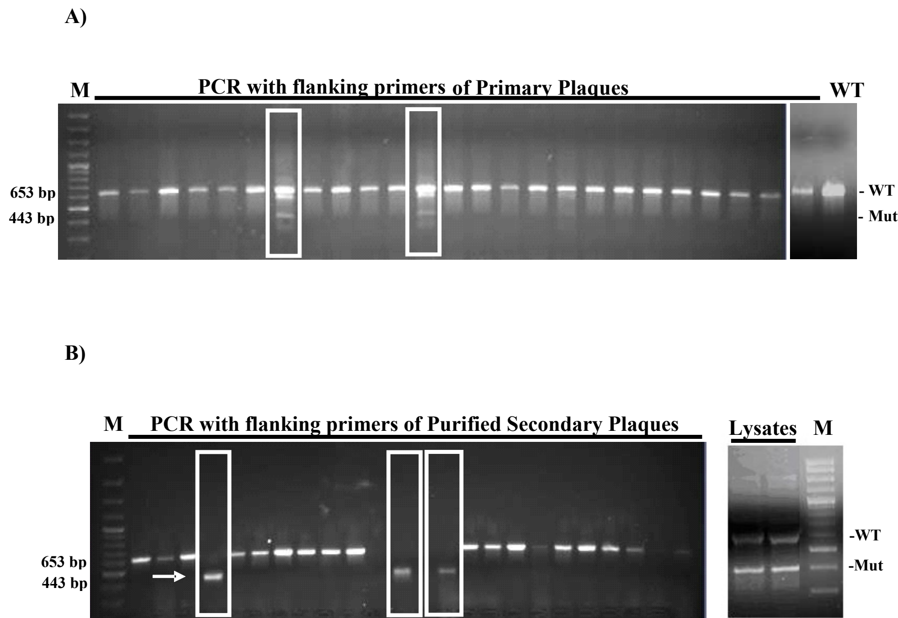
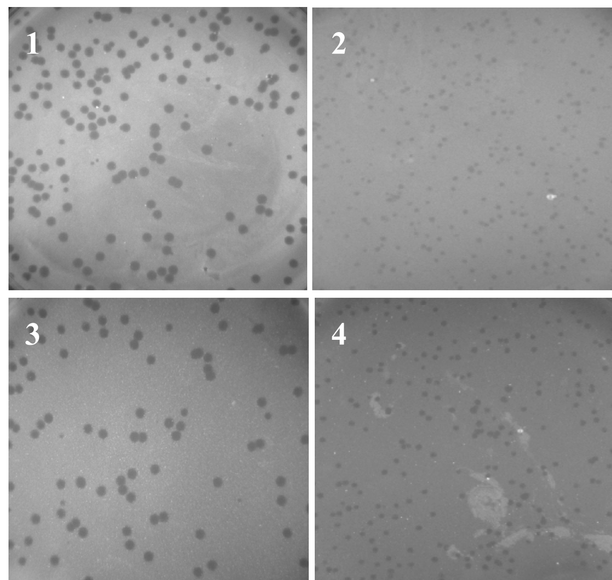


Figure 7. Construction of an Ms6 *gp1* deletion mutant. A) Following co-electroporation of 200 bp substrate and Ms6 genomic DNA, primary plaques were recovered and screened by PCR using primers flanking the deleted region. Mixed plaques containing wild-type and mutant DNA are boxed. On the right is the control with $Ms6_{wt}$ as the template. **B)** Mixed primary plaques were diluted and plated, and both lysates (on the right) and isolated secondary plaques screened with the same primers. Pure *gp1* deletion mutants are boxed. An arrow indicates the obtained PCR product.

This means that there is not a delay in the detection of phage released from $Ms6_{\Delta gp1}$ *M. smegmatis* infected cells when compared to cells infected with $Ms6_{wt}$. Thus, lack of *gp1* in the infecting virion had no evident effect on infection cycle duration. However, the number of infective particles released after $Ms6_{\Delta gp1}$ phage infection was lower than in an $Ms6_{wt}$ infection. Single-burst experiments were done to compare the viable progeny released from single cells infected with $Ms6_{wt}$ or $Ms6_{\Delta gp1}$. In our experimental conditions, $Ms6_{wt}$ infection released an average of 149

viable phages per bacterium, while $Ms6_{\Delta gp1}$ yielded a reduced burst size of approximately 45 viable phages per infected cell. These results show that Gp1 is necessary to obtain a normal burst of infective phages, although it has no influence on the latent period duration.

A)



B)

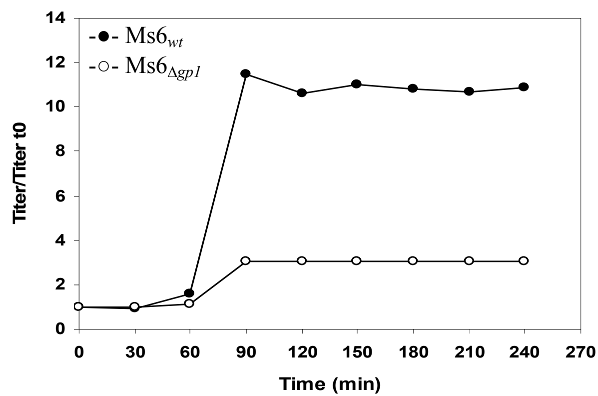


Figure 8. Gp1 is essential for efficient mycobacterial cell lysis. A) Ms6$_{\Delta gp1}$ phage forms smaller plaques than Ms6$_{wt}$. **1.** *M. smegmatis* infected with Ms6$_{wt}$ phage; **2.** *M. smegmatis* infected with Ms6$_{\Delta gp1}$; **3.** *M. smegmatis*:pMJC6 infected with Ms6$_{\Delta gp1}$ in induced conditions. Larger plaques are restored by complementation with Gp1; **4.** *M. smegmatis*:pMJC6 infected with Ms6$_{\Delta gp1}$ in uninduced conditions. **B**) One-step growth curves of phages Ms6$_{wt}$ (-●-) and Ms6$_{\Delta gp1}$ (-○-) infecting *M. smegmatis* mc^{2}155.

Discussion

In the present work we provide evidence that *gp1*, the first gene in the lysis module of mycobacteriophage Ms6, encodes a secretion chaperone-like protein that binds the endolysin, assisting its export to the extracytoplasmic environment and is required for efficient *M. smegmatis* lysis. Previous work demonstrated that co-expression of the Ms6 endolysin and holin is not sufficient to support *E. coli* lysis (Garcia *et al.*, 2002). These data, along with the observation that in *E. coli,* lysis is achieved by co-expression of Gp1 and LysA, but not with endolysin alone, suggest that LysA must reach the peptidoglycan in a holin-independent manner. Analysis of the Gp1 secondary structure predicted no transmembrane domains or hydrophilic C-terminus, hallmarks of holins secondary structure (Young *et al.*, 2000; Young, 2002). Together with the inability of Gp1 to complement the λgt11 holin defect, these results indicate that Gp1 is not a holin. It has already been described for some bacteriophages of both Gram-negative and Gram-positive hosts that endolysins are not always dependent on holins for export (São-José *et al.*, 2000; Xu *et al.*, 2004). Analysis of the Ms6 LysA amino acid sequence did not predict an amino-terminal signal sequence, as observed for bacteriophage fOg44 endolysin (São-José *et al.*, 2000), or an N-terminal SAR domain seen in phages P1 and P21 endolysins (Xu *et al.*, 2004; Xu *et al.*, 2005). This suggests that Ms6 endolysin must be exported in a different way.

Previous observation of SDS-resistant dimers of the Gp1 protein (data not shown) led us to consider a chaperone-like activity for Gp1. This hypothesis was strengthened by the physical properties of Gp1 that fit the characteristics of the well-characterized TTS system chaperones, such as a low molecular weight (8.3 kDa) and a low pI (4.6). In addition, the genes encoding these chaperones are often adjacent to the genes encoding their substrates (Page and Parsot, 2002) as observed for *gp1*. The *gp1* gene is immediately upstream of the endolysin (*lysA*) gene, in a different reading frame, with overlapping stop and start codons. This gene arrangement is not uncommon in phage genomes, especially among genes encoding lysis proteins. Embedded and overlapped classes of *Rz/Rz1* gene pairs have been described for phages infecting Gram-negative hosts. This physical association suggests a strong selective pressure for linkage of these genes. Moreover, having genes overlapped in the same direction in different reading frames can result in much tighter linkage than simple adjacency of the genes (Summer *et al*., 2007). Crosslinking experiments showed that Gp1 oligomerizes in dimers, trimers, tetramers and pentamers. We believe that this characteristic is important for its role in lysis as Gp1 homo-oligomers interacted with Ms6 endolysin during phage infection of *M. smegmatis.* This is an interesting finding as no other protein has been shown to interact with a lysin in any system, except for the transcriptional regulatory interaction between RNA polymerase and lysozyme of bacteriophage T7 (Moffatt and Studier, 1987).

In this study we have determined that the N-terminal 60 amino acids of Ms6 LysA are necessary and sufficient for Gp1 binding, and are essential for LysA export. The absence of a lysis phenotype when active LysAΔ_{1-60} was co-expressed with Gp1 in *E. coli* suggested that an interaction between Gp1 and LysA is required for LysA reach its target, the peptidoglycan. These data were further supported by alkaline phosphatase fusion experiments in *M. smegmatis*. PhoA activity of the LysA-PhoA’ hybrid protein increased in the presence of Gp1 in a single-copy

integrative system. This indicates that Gp1 promotes the translocation of LysA-PhoA' across *M. smegmatis* cytoplasmic membrane. Considering these data and the secreted properties of endolysins, we conjectured whether Ms6 LysA could be targeted to the extracytoplasmic environment as it is being synthesized, as suggested for P1 endolysin (Xu *et al.*, 2004), rather than at the end of phage maturation, as observed in λ-like systems (Young, 2005).

Despite the common properties of Gp1 and TTS chaperones, the fact that the Sec-dependent export pathway is highly conserved among different bacteria led us to investigate the involvement of the Sec-system on LysA export. The reason of our experiments relied on the fact that, holin-independent export of endolysins requires the Sec-system of their hosts (São José *et al*., 2000; Xu *et al.*, 2004). We observed that the SecA inhibitor, sodium azide, prevented lysis achieved by Gp1 and LysA co-expression in an *E. coli* host carrying the wild-type *secA* locus. Similarly, no lysis was observed at the non-permissive temperature in the thermosensitive *secA51 E. coli* strain. Mycobacteria have two homologues of *secA*, *secA1* and *secA2*; *secA1* is essential in *M. smegmatis*, while *secA2* is non-essential for viability. Failure to obtain *secA1* deletion mutants strongly suggests that deletion of *secA1* in *M. smegmatis* is a lethal event (Braunstein *et al*., 2001). Nevertheless, we tried to evaluate the dependence on the *M. smegmatis* SecA1 to achieve an optimal lysis after Ms6 infection. Using an antisense RNA strategy, we found *secA1* depletion to be highly lethal to *M. smegmatis*, even though antisense expression was regulated by the inducible acetamidase promoter (data not shown). For this reason, direct involvement of *M. smegmatis* SecA1 in LysA translocation during Ms6 infection of the host cell remains to be demonstrated. Although SecA2 is necessary for the export of some proteins (Rigel and Braunstein, 2008), Ms6 infection of a *M. smegmatis* Δ*secA*2 strain did not alter the lysis phenotype (data not shown). Of note is the fact that LysA lacks a signal sequence or an N-terminal transmembrane domain, and Sec-exported proteins are synthesized as precursors

with an N-terminal signal sequence (Rigel and Braunstein, 2008). Since Gp1 interacts with LysA, it may be a potential candidate for engaging the Sec system. Future investigation is needed to understand how the complex Gp1-LysA interacts with the Sec system for transport. Although we currently do not know how the endolysin is activated once it is localized next to the peptidoglycan, we hypothesize that Ms6 holin functions as described for the lambdoid phage 21 holin ($S^{21}68$) (Park *et al*., 2007). When $S^{21}68$ triggers, it eliminates the p.m.f., causing endolysin activation, but does not form holes in the membrane large enough to allow passage of the cytoplasmic endolysin. Proton-motive force depletion in *M. smegmatis* expressing Gp1 and LysA after nisin addition resulted in complete lysis. The absence of lysis observed in nisin-treated cells expressing the endolysin alone suggests a depolarizing role for the Ms6 holin.

Construction of $Ms6_{\Delta gp1}$, a viable mutant phage deleted for *gp1* confirmed that Gp1 is not essential for plaque formation. Analysis of the phage infection cycle by one-step growth curves showed that the latent period was similar in both wild-type and *gp1*-deleted phages. $Ms6_{\Delta gp1}$ adsorption efficiency was not lower than in wild-type phage, and no discrepancy in lysis timing was observed. However, the mutant phage had a burst-size reduced almost four-fold compared to wild-type. One interpretation could be that the absence of Gp1 during the infective viral cycle led to incomplete lysis, with virions trapped inside the cells, due to a deficient access of the endolysin to the peptidoglycan and thus a deficient cell lysis. However, attempts to determine if the burst size of the mutant phage, using chloroform to disrupt incomplete lysed cells, equals the wild-type burst size were unsuccessful.

In conclusion, our results indicate that Gp1 is required for efficient lysis of *M. smegmatis*, mainly because Ms6 holin alone does not provide a way for endolysin to reach the peptidoglycan. A proposed model for Ms6 Gp1 role is presented in Fig. 9. This model does not seem to be conserved among all

mycobacteriophages. A recent analysis of 60 mycobacteriophage genomes (Hatfull *et al*., 2010) revealed that Gp1 homologues are present in the lytic cassette of only 18 of 60 genomes, suggesting that endolysins from the other 42 phages are delivered by another mechanism. Ms6 Gp1 is highly similar to homologous proteins from mycobacteriophages that are grouped in the same cluster, F1 (Hatfull *et al*., 2010), thus it is expected that these phages use a similar model for endolysin export. It should be noted that an alignment of endolysins from phages belonging to this cluster does not show the same degree of similarity as observed for Ms6 Gp1 and its homologues, especially at the N-terminal region. However, such variation does not mean that an interaction between Gp1 homologues and their endolysins does not occur. We must stress that for the TTS chaperones and also for SecB no apparent consensus in the amino acid sequence is observed among their effectors. Molecular chaperones bind to hydrophobic residues and/or unstructured backbone regions in a non-native state (Bukau and Horwich, 1998; Page and Parsot, 2002; Zhou and Xu, 2005). This is, to our knowledge, the first phage secretion chaperone protein identified within a bacteriophage lysis module and opens new avenues for studying mycobacterial secretion of mycobacteriophage lysis effectors.

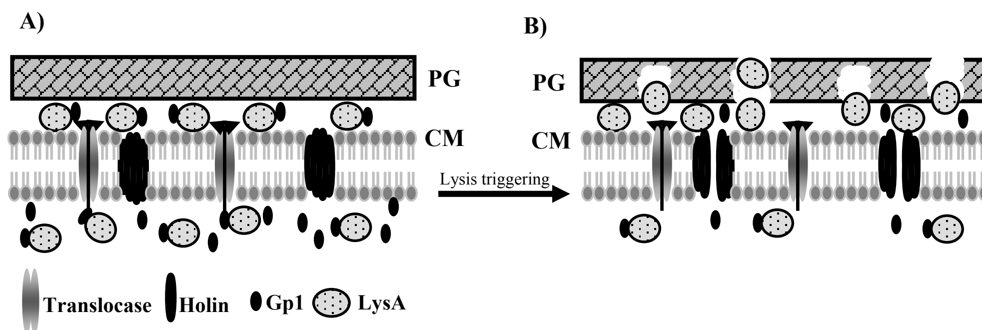


Figure 9. A model for Ms6 LysA export mediated by the Gp1 chaperone-like protein. **A**) LysA binds to Gp1 and is continuously exported to the extracytoplasmic environment. **B**) When lysis is triggered, the already positioned endolysin cleaves the peptidoglycan

bonds. Membrane lesions formed by holin are not large enough to allow LysA to access the PG. **PG-**Peptidoglycan, **CM-**Cytoplasmic Membrane.

Acknowledgments

We thank Dr. J. L. Herrmann (University of Versailles-Saint-Quentin-en-Yvelines, France) for plasmid pSMT3-[19-phoA], Dr. Varalakshmi Vissa and Dr. Michael McNeil (Colorado State University, USA) for pVVAP vector, Dr. Miriam Braunstein (University of North Carolina, USA) for the *M. smegmatis* Δ*secA2* mutant strain, Dr. Paula Leandro (University of Lisbon, Portugal) for assistance with the crosslinking experiments and Dr. Graham Hatfull, Dr. Julia van Kessel and Dr. Laura Marinelli (University of Pittsburgh, USA) for supplying plasmid pJV53 and for technical assistance with the recombineering experiments. We also thank Dr. Ry Young (Texas A&M University, USA) and Dr. Carlos São-José (Institute of Molecular Medicine, Portugal) for helpful suggestions. This work was supported by funds provided by FCT (Fundação para a Ciência e Tecnologia)-PTDC/SAU-FCF/73017/2006. Maria João Catalão and Filipa Gil are recipients of Ph.D. fellowships from FCT (SFRH/BD/24452/2005) and (SFRH/BD/29167/2006).

References

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

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

Bierbaum G. and Sahl H. G. (1987) Autolytic system of *Staphylococcus simulans* 22: influence of cationic peptides on activity of *N*-acetylmuramoyl-L-alanine amidase. *J Bacteriol* **169:**5452-5458.

Braunstein M., Brown A. M., Kurtz S. and Jacobs W. R. Jr. (2001) Two nonredundant SecA homologues function in Mycobacteria. *J Bacteriol* **183:**6979-6990.

Breukink E. (2006) A lesson in efficient killing from two-component lantibiotics. *Mol Microbiol* **61:**271-273.

Bukau B. and Horwich A. L. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell* **92:**351-366.

Chung H. J., Montville T. J. and Chikindas M. L. (2000) Nisin depletes ATP and proton-motive force in Mycobacteria. *Lett Appl Microbiol* **31:**416-420.

Feldman M. F. and Cornelis G. R. (2003) The multitalented type III chaperones: all you can do with 15 kDa. *FEMS Microbiol Lett* **219:**151-158.

Freitas-Vieira A., Anes E. and Moniz-Pereira J. (1998) The site-specific recombination locus of mycobacteriophage Ms6 determines DNA integration at the tRNA (Ala) gene of *Mycobacterium* spp. *Microbiology* **144:**3397-3406.

Garcia M., Pimentel M. and Moniz-Pereira J. (2002) Expression of mycobacteriophage Ms6 lysis genes is driven by two sigma (70)-like promoters and is dependent on a transcription termination signal present in the leader RNA. *J Bacteriol* **184:**3034-3043.

Gil F., Catalão M. J., Moniz-Pereira J., McNeil M., Leandro P. and Pimentel M. (2008**)** The lysis cassette of mycobacteriophage Ms6 encodes an enzyme with lipolytic activity. *Microbiol*ogy **154:**1364-1371.

Gil F., Grzegorzewicz A. E., Catalão M. J., Vital J., McNeil M. and Pimentel M. (2010) The mycobacteriophage Ms6 LysB specifically targets the outer membrane of *Mycobacterium smegmatis*. *Microbiology* **156:**1497-1504.

Gründling A., Manson M. D. and Young R. (2001) Holins kill without warning. *Proc Natl Acad Sci USA* **98:**9348-9352.

Hatfull G. F., Pedulla M. L., Jacobs-Sera D., Cichon P. M., Foley A., Ford M. E., Gonda R. M., Houtz J. M., Hryckowian A. J., Kelchner V. A., Namburi S. Pajcini K. V., Popovich M. G., Schleicher D. T., Simanek B. Z., Smith A. L., Zdanowicz G. M., Kumar V., Peebles C. L., Jacobs W. R. Jr., Lawrence J. G. and Hendrix R. W. (2006) Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genetics* **2:**0835-0847.

Hatfull G. F., Jacobs-Sera D., Lawrence J. G., Pope W. H., Russell D. A., Ko C. C., Weber R. J., Patel M. C., Germane K. L., Edgar R. H., Hoyte N. N., Bowman C.

A., Tantoco A. T., Paladin E. C., Myers M. S., Smith A. L., Grace M. S., Pham T. T., O'Brien M. B., Vogelsberger A. M., Hryckowian A. J., Wynalek J. L., Donis-Keller H., Bogel M. W., Peebles C. L., Cresawn S. G. and Hendrix R. W. (2010) Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. *J Mol Biol* **397(1):**119-143.

Herrmann J. L., O'Gaora P., Gallagher A., Thole J. E. R. and Young D. B. (1996) Bacterial glycoproteins: a link between glycosilation and proteolytic cleavage of a 19 kDa antigen from *Mycobacterium tuberculosis*. *EMBO J* **15:**3547-3554.

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

Manoil C. and Beckwith J. (1986) A genetic approach to analyzing membrane protein topology. *Science* **233:**1403-1408.

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

Moffatt B.A. and Studier F. W. (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49:**221–227.

Montville T. J., Chung H. J., Chikindas M. L. and Chen Y. (1999) Nisin A depletes intracellular ATP and acts in bactericidal manner against *Mycobacterium smegmatis*. *Lett Appl Microbiol* **28:**189-193.

Nascimento J. G., Guerreiro-Pereira M. C., Costa S. F., São-José C. and Santos M. A. (2008) Nisin-triggered activity of Lys44, the secreted endolysin from *Oenococcus oeni* phage fOg44. *J Bacteriol* **190:**457-461.

Oliver D. B. and Beckwith J**.** (1981) *E. coli* mutant pleiotropically defective in the export of secreted proteins. *Cell* **25:**765-772.

Oliver D. B., Cabelli R. J., Dolan K. M. and Jarosik G. P. (1990) Azide-resistant mutants of *Escherichia coli* alter the SecA protein, an azide-sensitive component of the protein export machinery. *Proc Natl Acad Sci USA* **87:**8227-8231.

Page A. L. and Parsot C. **(**2002) Chaperones of the type III secretion pathway: jacks of all trades. *Mol Microbiol* **46:**1-11.

Parish T., Mahenthiralingam E., Draper P., Davis E. O. and Colston M. J. **(**1997) Regulation of the inducible acetamidase gene of *Mycobacterium smegmatis*. *Microbiology* **143:**2267-2276.

Park T., Struck D. K., Dankenbring C. A. and Young R. (2007) The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. *J Bacteriol* **189:**9135-9139.

Portugal I., Anes E. and Moniz-Pereira J. (1989) Temperate mycobacteriophage from *M. smegmatis*. *Acta Leprol* **7:**243-244.

Rigel N. W. and Braunstein M. (2008) A new twist on an old pathway - acessory Sec systems. *Mol Microbiol* **69**:291-302.

Ruhr E. and Sahl H. G. (1985) Mode of action of the peptide antibiotic nisin and influence on the membrane potencial of whole cells and on cytoplasmic and artificial membrane vesicles. *Antimicrob Agents Chemother* **27:**841-845.

Sambrook J. and Russell D. W. (2001). Molecular Cloning: a Laboratory Manual, 3rd Ed. Cold Spring Harbor, NY: Cold Spring Harbor Press.

São-José C., Parreira R., Vieira G. and Santos M. 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. *J Bacteriol* **182:**5823-5831.

Silhavy T. J., Berman M. L. and Enquist L. W. (1984) Experiments with gene fusions. Cold Spring Harbor Laboratory Press, N.Y., pp 111-112.

Smith D. L., Chang C. Y. and Young R. (1998) The lambda holin accumulates beyond the lethal triggering concentration under hyperexpression conditions. *Gene Expr* **7:**39-52.

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

Summer E. J., Berry J., Tran T. A. T., Niu L., Struck D. K. and Young R. (2007) *Rz/Rz1* lysis genes equivalent in phages of gram-negative hosts. *J Mol Biol* **373:**1098-1112.

Timm J., Perilli M. G., Duez C., Trias J., Orefici G., Fattorini L. *et al.* (1994) Transcription and expression analysis, using *lacZ* and *phoA* gene fusions, of *Mycobacterium fortuitum* beta-lactamase genes cloned from a natural isolate and a high-level beta-lactamase producer. *Mol Microbiol* **12:**491-504.

Wang I.-N., Smith D. and Young R. (2000). Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* **54:**799-825.

Wiedemann I., Benz R. and Sahl H. G. (2004) Lipid II-mediated pore formation by the peptide antibiotic nisin: a black lipid membrane study. *J Bacteriol* **186:**3259-3261.

Xu M., Struck D. K., Deaton J., Wang I.-N. and Young R. (2004) A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. *Proc Natl Acad Sci USA* **101:**6415-6420.

Xu M., Arulandu A., Struck D. K., Swanson S., Sacchettini J. C. and Young R. (2005) Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. *Science* **307:**113-117.

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

Young R. (2002) Bacteriophage holins: deadly diversity. *J Mol Microbiol Biotechnol* **4:**21-36.

Young R. (2005) Phage lysis. *In*, Phages: Their Role in Bacterial Pathogenesis and Biotechnology. Waldor M. K., Friedman D. I., and Adhya S. L. (Eds.). Washington D.C., American Society for Microbiology, pp. 92-127.

Young R. and Wang I.-N. (2006) Phage lysis. *In*, The Bacteriophages. Calendar R. (Ed.). New York Oxford University Press, pp. 104-125.

Zhou J. and Xu Z. (2005) The structural view of bacterial translocation-specific chaperone SecB: implications for function. *Mol Microbiol* **58:**349-357.

3

Functional Analysis of the Holin-Like Proteins of Mycobacteriophage Ms6

M. J. Catalão, F. Gil, J. Moniz Pereira, and M. Pimentel

Centro de Patogénese Molecular, Unidade dos Retrovírus e Infecções Associadas, Faculty of Pharmacy, University of Lisbon, Portugal

Manuscript in preparation
2010

Summary

The mycobacteriophage Ms6 is a temperate double-stranded DNA (dsDNA) bacteriophage which, in addition to the endolysin (*lysA*)-holin (*hol*) lysis system, encodes three additional proteins within its lysis module: *gp1*, *lysB* and *gp5*. Ms6 *gp4* (*hol*) was previously described as a putative class II holin-like protein with high similarity to the bacteriophage r1t holin, based on its structural features. By analysis of the amino acid sequence of Hol, an N-terminal signal-arrest-release (SAR) domain was identified, followed by a typical transmembrane domain (TMD) characteristic of pinholins. In addition, a second putative holin gene (*gp5*) with a predicted single TMD at the N-terminal region was identified at the end of the Ms6 lytic operon. Neither the putative class II holin nor the single TMD polypeptide could trigger lysis in pairwise combinations with the endolysin LysA in *E. coli*. Furthermore, crosslinking experiments showed that Hol and Gp5 oligomerize and that both proteins interact. One step growth curves and single burst size experiments of different Ms6 derivatives deleted in different regions of the lysis operon demonstrated that the gene products of *hol* and *gp5*, although nonessential for phage viability, appear to play a role in controlling the timing of lysis: an $Ms6_{\Delta hol}$ deletion mutant caused slightly accelerated lysis whereas an $Ms6_{\Delta gp5}$ deletion mutant delayed lysis, which is consistent with holin function. Deletion of the chaperone-like protein *gp1*, in both deletion mutants, was catastrophic for lysis, with more than five-fold reduction of the burst size. These results suggest that *hol* and *gp5* encode holin-like proteins whose combined action could play the role of a holin and that expression of both proteins is necessary to effect host cell lysis at the correct and programmed timing. Moreover, the Gp1 accessory lysis protein of mycobacteriophage Ms6 was proven to be essential for an efficient lysis of mycobacteria and to a productive burst size.

Introduction

The majority of dsDNA bacteriophages described so far terminates each infection cycle through the programmed and regulated activity of two phage-encoded proteins, the endolysin and the holin, a small membrane protein that controls the endolysin function and the access to the peptidoglycan (Young, 1992; Young *et al.*, 2000). Endolysins are characterized by their ability to directly target covalent bonds in the peptidoglycan (PG) layer of the bacterial cell wall; the result of this activity is disruption of the rigid murein layer and release of newly synthesized virions (Loessner, 2005). During phage assembly, holin molecules accumulate in the cytoplasmic membrane without detectable effect on the host (Wang *et al.*, 2000; Gründling *et al.*, 2001). Then, at an allele-specific time programmed into their primary structure, holins trigger to disrupt the cytoplasmic membrane (Gründling *et al.*, 2000a; Gründling *et al.*, 2001). Holins are extremely diverse, found in many unrelated sequence families with at least three membrane topologies suggesting that they may have evolved from multiple distinct origins to allow precisely scheduled efficient lysis and rapid adjustment of the lysis time, either on the basis of genetic selection, or in some cases, in real time in response to environmental changes (Wang *et al.*, 2000; Young, 2002). The canonical holins, such as those of phages λ and T4, form very large holes that allow fully-folded and active endolysins accumulated in the cytosol to pass through the cytoplasmic membrane and attack the peptidoglycan. These holes are nonspecific and allow the passage of unrelated endolysins (Young, 2002) and proteins larger than 480 kDa (Wang *et al.*, 2003). In addition, hole formation is absolutely required for lysis. Many phages also encode an antiholin which contributes to control the timing of host lysis by inhibiting the holin. In some cases, the antiholin is encoded by the holin gene, with an additional N-terminal extension of several amino acids -dual-start motif- (Bläsi and Young, 1996), or an alternative intragenic start codon

(Vukov *et al*., 2003). In other cases, the antiholin is encoded by an independent gene (Walker and Walker Jr., 1980; Ziermann *et al*., 1994; Ramanculov and Young, 2001).

Recently, an alternative and remarkably different class of holin-endolysin systems has emerged (Park *et al*., 2006; Park *et al*., 2007). This class, represented by the lambdoid bacteriophage 21, utilizes endolysins having N-terminal secretory SAR signals and pinholins. For phages encoding SAR endolysins, the holin protein needs only to produce lesions large enough to allow the passage of ions and depolarize the cytoplasmic membrane in order to fulfil its role in controlling the timing of lysis. Indeed, unlike lesions formed by the λ holin, lesions formed by the phage 21 holin do not allow the passage of λ endolysin (Park *et al*., 2007). The term ''pinholin'' has been proposed to differentiate the small-hole (pinhole) forming character of the phage 21 holin from the canonical holins that form large, nonspecific holes (Park *et al*., 2007).

The genetic organization of the mycobacteriophage Ms6 lysis functions was previously described (Garcia *et al*., 2002). In addition to the endolysin (*lysA*) and a holin (*hol*) gene (*gp4*), three accessory lysis genes were also identified: *gp1*, *lysB* and *gp5*. The *gp1* gene was recently identified as encoding a chaperone-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). The product of *lysB* gene was also characterized and encodes an enzyme with lipolytic activity that hydrolyzes the mycolic acids from the mycolyl-arabinogalactan-peptidoglycan complex (Gil *et al*., 2008; Gil *et al*., 2010).

The Ms6 holin-like gene (*hol*) shares some structural characteristics with class II holins, which are usually hydrophobic in nature and small in size, with a hydrophilic carboxy-terminal domain and two potential transmembrane domains. Hol function was also supported by its ability to complement a λ phage S mutant (suggesting that the Ms6 holin allows the non-specific release of the λR endolysin

to the periplasm) and the observed lethal phenotype when overexpressed in *E. coli*, explained by the introduction of non-specific lesions in the cytoplasmic membrane. However, unlike some holins as the lambda S holin it lacks a dual-start motif (Garcia *et al*., 2002). Analysis of the amino acid sequence of Hol identified an N-terminal signal-arrest-release (SAR) sequence followed by a typical transmembrane domain, characteristic of the pinholin class of proteins. In addition, a second putative holin gene (*gp5*) encoding a protein with a predicted single TMD at the N-terminal region was identified at the end of the Ms6 lytic operon. It starts at an ATG codon that overlaps the *hol* TGA stop codon in a different reading frame. A BLASTp search using the Ms6 Gp5 deduced amino acid sequence identified a number of putative proteins with unknown function restricted to the mycobacteriophages, with a high degree of sequence identity (Hatfull *et al*., 2010). In this study we characterized the function of *hol* and *gp5* gene products and demonstrated that, although nonessential for the infective cycle, both appear to play a role in controlling the timing of lysis. We also present evidence suggesting that Hol has structural features of a pinholin and that Gp5 encodes a holin-like protein, whose combined action is essential to effect host cell at the correct programmed lysis timing.

Experimental Procedures

Bacterial strains, phages, plasmids and culture conditions

Bacterial strains, phages and plasmids used throughout this study are listed in Table 1. *E. coli* strains were grown at 37 ºC, in Luria-Bertani (LB) broth or agar supplemented with 100 µg ml^{-1} ampicillin or 30 µg ml^{-1} kanamycin, when appropriate. *M. smegmatis* recombinant strains were grown at 37 ºC in 7H9 medium (Difco) supplemented with 0.05% Tween 80 and 0.5% glucose, with

shaking or Middlebrook 7H10 (Difco), containing 15 µg ml^{-1} kanamycin. For induced conditions 0.2% succinate and 0.2% acetamide were also added to media.

Plasmid construction

Unless indicated otherwise, the DNA fragments were obtained by PCR using Ms6 genomic DNA as a template. DNA amplification, plasmid isolation and electrophoresis were carried out using standard techniques (Sambrook and Russell, 2001). *E. coli* and *M. smegmatis* mc^{2}155 cells were transformed as described previously (Snapper *et al.*, 1990; Sambrook and Russell, 2001). Restriction enzymes and T4 DNA ligase (New England Biolabs) were used according to the supplier's recommendations. All oligonucleotides were from Thermo Scientific and are listed in Table 2.

In order to construct plasmids pMJC21 and pMJC22, DNA fragments containing genes *gp5* or *hol* and *gp5* were obtained by PCR amplification with primers Porf5a/Porf5-c2 or Porf4-1/Porf5-c2, respectively. Primers were designed in order to generate restriction sites, and the DNA fragments were inserted in the same sites of vector pQE30 (QIAGEN) allowing a fusion to a hexahistidine tag at the N-terminus. To obtain plasmid pMJC23, the DNA fragment containing gene *gp11*, was amplified by PCR using D29 genomic DNA as a template with primers P*hol*D29fwd/P*hol*D29rv and cloned into SacI/HindIII sites of pQE30. pMJC24 and pMJC25 were constructed in two steps: the λ*R* gene was amplified using the genomic DNA of bacteriophage λgt11 as a template with primers Pλ*R*fwd/Pλ*R*rv and cloned into BamHI/SacI sites of pQE30. *hol* or *gp5* genes were amplified by PCR using the Ms6 genomic DNA as template with primers Porf4-1/Porf4-c1 or *gp5*RBSfwd/Porf5-c2 and cloned into SacI/HindIII sites of pQE30:λ*R*, generating plasmids pMJC24 and pMJC25, respectively. pMJC27, pMJC28 and pMJC29 were

Table 1. Strains, bacteriophages and plasmids used in this study.

Strain, bacteriophage, or plasmid	Description	Reference or Source
Bacteria		
Escherichia coli		
JM109	*recA1 endA1 gyr96 thi hsdR17 supE44 relA1* Δ(*lac-proAB*) [F´ *traD36 proAB lacI*^{q}ZΔM15]	Stratagene
BL21	F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dmc*	Novagen
BL21 (DE3)	F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dmc* (DE3)	Novagen
Mycobacterium smegmatis mc^{2}155	High-transformation-efficiency mutant of *M. smegmatis* ATCC 607	Snapper *et al*., 1990
Bacteriophages		
D29	Lytic phage that infects both fast and slow-growing mycobacterial species	Institute Pasteur collection
λgt11	cIts857 Sam100	Stratagene
Ms6$_{wt}$	Temperate bacteriophage from *M. smegmatis*	Portugal *et al*., 1989
Ms6$_{\Delta gp1}$	213 bp in-frame deletion of the Ms6 *gp1* gene	This study
Ms6$_{\Delta hol}$	210 bp in-frame deletion of the Ms6 *hol* gene	This study
Ms6$_{\Delta gp5}$	366 bp in-frame deletion of the Ms6 *gp5* gene	This study
Ms6$_{\Delta gp1hol}$	213 bp and 210 bp in-frame deletions of the Ms6 *gp1* and *hol* genes, respectively	This study
Ms6$_{\Delta gp1gp5}$	213 bp and 366 bp in-frame deletions of the Ms6 *gp1* and *gp5* genes, respectively	This study
Plasmids		
pQE30	Expression vector; T5 promoter; Amp^r, lacI^q	QIAGEN
pET29a(+)	Expression vector, T7 promoter; Kan^r	Novagen
pJV53	Derivative of pLAM12 with Che9c *60* and *61* under control of the acetamidase promoter; Kan^r	van Kessel and Hatfull, 2007
pMG231A	*lysA* cloned into pQE30	Garcia *et al*., 2002
pMP300	*lysA* and *hol* Ms6 cloned in pQE30	Garcia *et al*., 2002
pMP310	*hol* Ms6 cloned in pQE30	Garcia *et al*., 2002
pMJC21	*gp5* Ms6 cloned in pQE30	This study
pMJC22	*hol* and *gp5* Ms6 cloned in pQE30	This study
pMJC23	*gp11* D29 cloned in pQE30	This study
pMJC24	λ*R* and *hol* Ms6 cloned in pQE30	This study
pMJC25	λ*R* and *gp5* Ms6 cloned in pQE30	This study
pMJC27	*gp5* Ms6 cloned in pMG231A	This study
pMJC28	*hol* and *gp5* Ms6 cloned in pMG231A	This study
pMJC29	*gp11* D29 cloned in pMG231A	This study
pMJC30	*hol* Ms6 cloned in pET29a(+)	This study
pMJC31	*gp5* Ms6 cloned in pET29a(+)	This study
pMJC32	*hol* and *gp5* Ms6 cloned in pET29a(+)	This study

Ms6 lysis genes Accession No. AF319619

obtained by amplifying *gp5* or *hol* and *gp5* with primers *gp5*RBSfwd/Porf5-c2 or Porf4-1/Porf5-c2, or *gp11* with primers P*hol*D29fwd/P*hol*D29rv using the genomic DNA of mycobacteriophage D29 as a template. The DNA fragments were introduced into the SacI/HindIII sites of pMG231A. To obtain plasmids pMJC30, pMJC31 and pMJC32, the *hol*, *gp5* or *hol* and *gp5* genes of Ms6 were amplified

with primers Pr*hol*Ms6fwd/*hol*Ms6rv, *gp5*Ms6fwd/*gp5*Ms6rv or Pr*hol*Ms6fwd/ *gp5*Ms6rv, respectively, and cloned into BamHI/HindIII sites of pET29a(+). All constructs were validated by verifying the insert nucleotide sequence.

Table 2. Oligonucleotides used in this study

Oligonucleotides	Sequence[a] 5´-3´	Reference
Porf4-1	CTACGCGAGCTCACATAGGAGGCAC	Garcia *et al.* 2002
Porf4-c1	GCACCGAAGCTTCCCAGATCATGCGG	Garcia *et al.* 2002
*gp5*Ms6fwd	CGGGATCCATCTGGGAATCGGTG	This study
*gp5*Ms6rv	CGCAAGCTTCCGGTCGCGGCGTTTG	This sudy
*hol*Ms6rv	CGCAAGCTTTGCGGTGGTGACCGC	This study
Porf-5a	GAGCGGATCCCACCGCATGATCTGG	This sudy
Porf5-c2	GTGACAAGCTTCGTGCGCCCTTC	This sudy
P*hol*D29fwd	GCGAGCTCCGCCAGGAATGGAGCC	This sudy
P*hol*D29rv	GCAAGCTTTCATCGGTTCCAGGGC	This sudy
Pλ*R*fwd	CGGGATCCGTAGAAATCAATCAACG	This sudy
Pλ*R*rv	CCGTCGAGCTCTGCTCATACATC	This sudy
Pr*hol*Ms6fwd	GCGGATCCCTGACACGTTCATTCTGG	This sudy
*gp5*RBSfwd	CTGAGCTCTAGAGCGGTCACCACCGC	This study
PrΔ*hol*	GACTACAACCCCCAGCCCGCCATCGACTACCTACGCAC ATAGGAGGCACCCATGAGAGCGGTCACCGCATGATCTG GGAATCGGTGCGCGAAGCGGTGAAC	This sudy
PrExtenderΔ*hol*fwd	GATCCTCGACGCCGGCATGTTCTTCGCGAAACGCACCG GCCCGCACGTGGACTACAACCCCCAGCCCGCCATCGA	This sudy
PrExtenderΔ*hol*rv	GATGATGAGCAGTCCTATCAGGTCGATACCGTCGTCGG GCTGGTACGCCGCGTTCACCGCTTCGCGCACCGATTC	This sudy
PrΔ*gp5*	GGAAACCCCGGAACGGCGTCGGCCACTAGAGCGGTCA CCACCGCATGATCTAACCACCAGGAAAGAAGGGCGCA CGAATGTCACTACTGGCCGATCTCGC	This sudy
PrExtenderΔ*gp5*fwd	GCGGCGCGGCCGTCGTATCACTGCTGATGTCGATCGGC GCCGAACGCCGCGGAAACCCCGGAACGGCGTCGGCCA	This sudy
PrExtenderΔ*gp5*rv	ACCGGGCGCCCGCCCAGTCGCATGCGGGGCATGTG CGGGGTTGCAGGCCCGCGAGATCGGCCAGTAGTGA CATTC	This sudy
PrΔ*holgp5*	GACTACAACCCCCAGCCCGCCATCGACTACCTACGC ACATAGGAGGCACCCTAACCACCAGGAAAGAAGGG CGCACGAATGTCACTACTGGCCGATCTCGC	This study
PrΔ*gp1*	CGCCACCACCTGAGTGCGCGGGGTTTTTCTATGCCCG AAAGGACCCCGACATCCTCGGGAGCAAACGGTGACC ACGAAAGATCAAGTCGCCCAAATCACC	Catalão *et al.*,2010
PrExtenderΔ*gp1*fwd	CTCATCGACGACCGACCCGACTTCTGACCCACTGACT CCATCGACCCCGCCACCACCTGAGTGCGCGGGGTT	Catalão *et al.*,2010
PrExtenderΔ*gp1*rv	GCCAGGCATTCGCTGCGGGTGTAGCCGCGCGCCTTGG CTTCGGCGATGGTGATTTGGGCGACTTGATCTTTC	Catalão *et al.*,2010
*lysB*fwd[FP]	CGACTAGTCGCGACCTCGACCTGTTC	This sudy
Ms6rv[FP]	CTCTAGACGCACAACGCTCATGCGC	This sudy
PrP1fwd[FP]	CGGTACTAGTCGGCCTCGGCCTGC	Catalão *et al.*,2010
Pr*lysA*180bprv[FP]	GCAAGCTTGTGTGGGTAGGAGCCGTCC	Catalão *et al.*,2010

[a] underlined bases were added to provide additional restriction sites
[FP]Flanking primer

Protein interaction experiments

Crosslinking

BS^3 (Sigma) crosslinker stock solution at 10 mM final concentration, was prepared immediately before use to decrease the extent of hydrolysis in 20 mM Na-HEPES, 200 mM NaCl pH 7.0. For *in vitro* crosslinking experiment, *E. coli* BL21 (DE3) carrying plasmid pMJC30, pMJC31 or pMJC32, was induced at the logarithmic growth phase with 1 mM IPTG and ten-mililiter samples were withdrawn and pelleted after 1 h. Cells were resuspended in PBS, broken by sonication and centrifuged at 4º C. The proteins of the membrane fraction were extracted with 1% Triton X-100 for 2 hours at 37 ºC. The detergent fraction was treated with BS^3 solution to a final concentration between 1 and 5 mM at room temperature for 30 min. In the control samples the crosslinker was omitted. After incubation at room temperature, samples were resuspended in SDS-PAGE sample buffer that quenches the reaction. Aliquots were subjected to SDS-PAGE and Hol or Gp5 proteins were detected by Western-blotting using HRP-conjugated anti-His_6 monoclonal antibody (Roche).

Beta-galactosidase activity assay

β-galactosidase activity (Miller, 1972) was measured in the supernatants of induced *E. coli* BL21 cells carrying plasmids pQE30, pMP310, pMJC21, pMJC22 and pMJC23. Following 1 hour induction with 1 mM IPTG, 1 ml aliquots of exponential growing cultures were centrifuged and 30 µl of supernatants were added to 66 µl of *o*-nitrophenyl-β-D-galactopyranoside-ONPG solution (Sigma, US) (4mg/ml in sodium phosphate buffer 0.1 M, pH 7.5), 3 µl of 4.5 M β-mercaptoethanol, 0.1 M $MgCl_2$ solution and 200 µl sodium phosphate buffer 0.1 M, pH 7.5. The reaction was performed at 37 ºC for 30 minutes and then stopped by the addition of 500 µl Na_2CO_3. The amount of *o*-nitrophenol released was

measured at 405 nm. Enzyme activity was expressed in arbitrary units of OD_{405nm} ml^{-1} of culture min^{-1}.

Construction of Ms6 mutant phages

Construction of Ms6 mutant phages was performed using Bacteriophage Recombineering of Electroporated DNA (BRED) in *M. smegmatis*. Recombineering substrates and BRED strategy were done as described previously (Marinelli *et al*., 2008; Catalão *et al*., 2010). Briefly, for deletion of the Ms6 *hol, gp5* or *hol and gp5* genes, 100 bp oligonucleotides, PrΔ*hol*, PrΔ*gp5* or PrΔ*holgp5*, that have 50 bp of homology upstream and downstream of the region to be deleted were extended by PCR using two 75 bp extender primers, PrExtΔ*hol*fwd/ PrExtΔ*hol*rv, PrExtΔ*gp5*fwd/PrExtΔ*gp5*rv or PrExtΔ*hol*fwd/PrExtΔ*gp5*rv, respectively, which have 25 bp of homology to the ends of the 100-mer and add an additional 50 bp of homology on either end. For deletion of the Ms6 *gp1* gene a 100 bp oligonucleotide (PrΔ*gp1*) was extended with primers PrExtΔ*gp1*fwd/ PrExtΔ*gp1*rv. The final 200 bp dsDNA products were purified using MinElute PCR Purification Kit (QIAGEN) and co-electroporated with Ms6 genomic DNA (for *hol*, *gp5* or *hol* and *gp5* deletion), or with $Ms6_{\Delta hol}$ or $Ms6_{\Delta gp5}$ genomic DNA (for *gp1* deletion) into electrocompetent recombineering cells of *M. smegmatis* mc^2155:pJV53. Cells were resuspended in 7H9 supplemented with 0.5% glucose and 1 mM $CaCl_2$, incubated at 37 ºC for 2 hours and plated as top agar lawns with *M. smegmatis* mc^2155. Phage plaques were picked into 100 µl phage buffer (10 mM Tris-HCl pH 7.5; 10 mM $MgSO_4$; 68.5 mM NaCl; 1 mM $CaCl_2$), eluted for two hours at room temperature and analysed by PCR with primers PrP1Fwd/Pr*lysA*180bprv flanking the *gp1* deletion or with primers *lysB*fwd/Ms6rv to detect the *hol*, *gp5* or *hol* and *gp5* deletions. Mixed primary plaques containing both the deletion mutant and the wild-type DNA were eluted as described above, and serial dilutions were plated with *M. smegmatis.* Individual secondary plaques

or lysates were screened by PCR with primers flanking the deletions for the presence of pure mutant phages.

One-step growth curves and burst size determination

One-step growth curve and burst-size determination were described previously (Catalão *et al*., 2010). The one step assays were carried out in cells in exponential growth using a multiplicity of infection (m.o.i.) of 1. *M. smegmatis* cells were pelleted and resuspended in 1 ml of a phage suspension ($Ms6_{wt}$, $Ms6_{\Delta gp1}$, $Ms6_{\Delta hol}$, $Ms6_{\Delta gp5}$, $Ms6_{\Delta gp1hol}$ or $Ms6_{\Delta gp1gp5}$) supplemented with 1 mM $CaCl_2$. The mixture was incubated 50 min at 37 ºC to allow adsorption of the phages. 100 µl of 0.4% H_2SO_4 was added to inactivate the non-adsorbed phages and the incubation continued for five min. The suspension was neutralized with 100 µl of 0.4% NaOH and diluted 1:100 in 7H9 supplemented with 0.5% glucose and 1 mM $CaCl_2$. 1 ml samples were withdrawn every 30 min until reaching 300 min. 100 µl of serial dilutions of each sample were plated with 200 µl of *M. smegmatis* cells, on 7H10 as top agar lawns and the phage titer for each sample was determined after 24h incubation at 37 ºC. The same experimental procedure was used for burst size determination except that 10 µl of infected cells were diluted in order to obtain ≤ one infected cells ml^{-1} in 7H9. Samples of 1 ml of infected cultures were distributed in 50 tubes and incubated for 180 min at 37 ºC. 200 µl of *M. smegmatis* cells and top agar (4 ml) were added to each tube and plated on 7H10. After 24h at 37 ºC, the phage plaques were counted and the distribution of *Poisson* [P(*n*)] was applied to determine the burst-size (BS): $P(n) = (e^{-c} \cdot c^{n})/ n!$ $(e < 1)$, where $P(n)$ is the probability of samples having *n* infected cells and ***c*** is the average number of infected cells per tube; and BS = (total plaque count in the 50 plates)/(total number of infected cells).

Results

Sequence analysis of the Ms6 holin-like genes

The mycobacteriophage Ms6 *hol* encodes a 77-amino acid polypeptide with a predicted molecular mass of 8.5 kDa that was previously described as a holin-like protein with high similarity to the *Lactococcus lactis* bacteriophage r1t holin (Garcia *et al*., 2002). This assumption was also based on the structural features of the protein which shares characteristics with the class II holins, high toxicity when overexpressed in *E. coli* leading to a lethal phenotype (Fig. 1A) and finally, the ability to complement a lambda S mutant. However, unlike some holins as the λS holin it lacks a dual start motif (Garcia *et al*., 2002). According to the sequence similarity to putative holin genes from mycobacteriophages of subcluster F1 (Fig. 2A) the Ms6 holin was included in Pham95 (Hatfull *et al*., 2010). An intriguing observation was the absence of a lysis phenotype when Hol was co-expressed with the Ms6 endolysin (Garcia *et al*., 2002) as both proteins were shown to be expressed in detectable levels (Catalão M. J., unpublished data). These results together with the statement that Gp1 is required for LysA access to the peptidoglycan (Catalão *et al*., 2010) led us to reanalyze the amino acid sequence of Hol. The Ms6 Hol possesses two TMDs, the most hydrophobic of which is TMD_2 (residues 39 to 59) (Fig. 2B) and a predicted N-out, C-out topology (Fig. 2C) according to the TMpred program (http://www.ch.embnet.org/software/TMPRED_form.html) from the Expasy server. TMD_1, not predicted by every TMD search algorithms, is present from residues 17 – 33 and has characteristics of a SAR domain, with a high percentage (11 out of 17) of weakly hydrophobic or polar uncharged residues (Fig. 2B) like Gly, Ala, Glu and Thr (Park *et al*., 2006). The presence of a SAR domain followed by a typical TMD suggests that Ms6 Hol is a pinholin, analogous to other pinholins already characterized, as the holin of phage

21, S21 68 (Park *et al*., 2007; Pang *et al*., 2009) or the holin of phage Xfas53 (Summer *et al*., 2010).

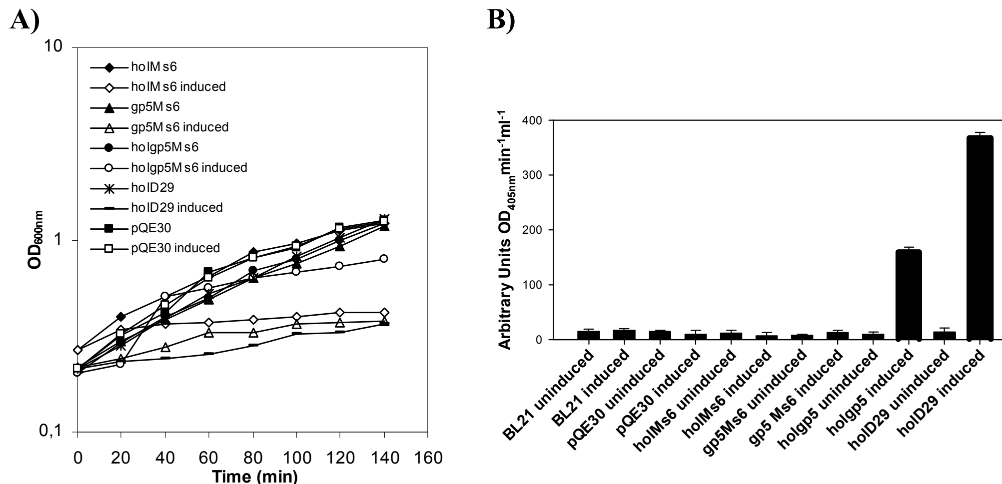


Figure 1. A) Expression of the holin-like proteins from mycobacteriophages Ms6 and D29 in *E. coli*. *E. coli* JM109 cells carrying plasmid pQE30 containing no insert or cloned genes were grown in LB broth at 37 ºC to an OD_{600nm} of 0.2. At time zero, transcription of cloned lysis genes was induced by addition of 1 mM IPTG. **B) Release of β-galactosidase from *E. coli* BL21 expressing Ms6 and D29 holin-like proteins.** Activity was determined in the supernatants of induced cultures. Results are averages of triplicate experiments.

As already mentioned, the Ms6 lysis module is organized into five genes. So far, no function has been assigned for *gp5*. It has the potential to codify for a 124-amino acid polypeptide with a predicted molecular mass of 14.1 kDa. The *gp5* gene starts at an ATG codon that overlaps the *hol* TGA stop codon in a different reading frame (Garcia *et al*., 2002). A common feature of most holins described to date is a lack of sequence similarity. A BLASTp search using the Ms6 Gp5 deduced amino acid sequence identified a number of putative proteins with unknown functions, restricted to the mycobacteriophage group of phages with a

high degree of sequence identity. Owing to their related amino acid sequences, they have been recently grouped in a mycobacteriophage gene family Pham96 (Hatfull *et al.*, 2010). Analysis of the amino acid sequence of Gp5 using TMpred showed that it possesses the structural characteristics of the class III holins, with a single membrane-spanning α-helix domain from residues 24-47 in the N-terminus, a highly charged and hydrophilic C-terminal domain (Fig. 2B) and a predicted N-out, C-in topology (Fig. 2C). There are three possible start codons in Ms6 *gp5* at positions 1, 6 and 10 that would produce 124 ($Gp5_{124}$), 118 ($Gp5_{118}$) and 114 ($Gp5_{114}$) amino acid products (with 14.1, 13.5 and 13 kDa, respectively), all preceded by potential ribosome-binding sites (5´-GAGCGG-3´ for $Gp5_{124}$, 5´-GGGAA-3´ for $Gp5_{118}$, 5´-GCGAAG-3´ for $Gp5_{114}$). When overexpressed in *E. coli*, we observed that bacterial growth was drastically inhibited suggesting that Gp5 forms lesions in the cell membrane (Fig. 1A). We tested β-galactosidase leakage from the BL21 strains expressing the Ms6 holin-like proteins Hol, Gp5 or both, toxic when expressed in *E. coli* (Fig. 1A). This assay has been used to search for proteins with canonical holin-like activity, as the damage caused to the cytoplasmatic membrane by the holin protein is sufficient to allow the leakage of cytoplasmatic contents including large proteins such as the cytoplasmatic β-galactosidase enzyme (Delisle *et al.*, 2006). β-galactosidase leakage phenomenon (measured by an increase in enzymatic activity) was not observed when Hol or Gp5 were expressed from pQE30 in *E. coli* BL21 suggesting that the lesions formed by these proteins are not large enough to allow the passage through the cytoplasmic membrane of proteins as large as β-galactosidase (Fig. 1B). However, an increase in β-galactosidase release was observed with the concomitant expression of Hol and Gp5 (Fig. 1B). In contrast to what is observed in Ms6 and other related mycobacteriophages, the mycobacteriophage D29 grouped in subcluster A2 (Hatfull *et al.*, 2010), does not possess homologues of the accessory lysis proteins Gp1 or Gp5. The lytic genes are clustered together, with the holin-like gene (*gp11*)

localized between the *lysA* and *lysB* genes (São-José *et al.*, 2003; Hatfull *et al.*, 2010). The D29 *gp11* gene has the potential to codify for a 141 amino acid polypeptide with a predicted molecular mass of 14.6 kDa and possesses structural characteristics of the class II holins, presenting two transmembrane domains from residues 11-29 (TMD_1) and 37-58 (TMD_2) in the N-terminus of the protein (Fig. 2B), a highly charged and hydrophilic C-terminal domain and N-in, C-in topology (Fig. 2C) as predicted by TMpred. D29 Gp11 (Hol) overexpression in *E. coli* inhibits cell growth as observed for holin-like proteins (Fig. 1A), and allowed the release of β-galactosidase to the culture medium with an increase of the enzymatic activity by comparison to the control (BL21 cells) (Fig. 1B). These results suggest that the D29 holin (Gp11) functions as a canonical holin, forming large lesions in the cytoplasmic membrane sufficient to allow the access of the cytoplasmic endolysin to the peptidoglycan and bring about an effective lysis of the host. These data support a potential holin function for D29 Gp11 (Hol) as previously proposed (Payne *et al.*, 2009; Hatfull *et al.*, 2010).

A)

```
                                    TMD1                     TMD2
Llij gp32         MLTRSFWIDAAERAARTFAQTAIATLGAGAVDLLATDWVSVLSVSGGAAVVSLLMSIGA 59
Pacc40 gp32       MLTRSFWIDAAERAARTFAQTAIATLGAGAVDLLATDWVSVLSVSGGAAVVSLLMSIGA 59
PMC gp32          MLTRSFWIDAAERAARTFAQTAIATLGAGAVDLLATDWVSVLSVSGGAAVVSLLMSIGA 59
Che8 gp34         MLTRSFWIDAAERAARTFAQTAIATLGAGAVDLLATDWVSVLSVSGGAAVVSLLMSIGA 59
Ms6 gp4           MLTRSFWIDAAERAIRTFAQTAIATLGAGAVDLMTTDWISVLSVSGGAAVVSLLMSIGA 59
Fruitloop gp31    MLTRSFWIDAAERAIRTFAQTAIATLGAGAVDLMTTDWISVLSVSGGAAVVSLLMSIGA 59
Tweety gp32       MLTRSFWIDAAERAIRTFAQTAIATLGAGAVDLMTTDWISVLSVSGGAAVVSLLMSIGA 59
Ramsey gp35       MLTRSFWIDAAERAIRTFAQTAIATLGAGAVDLMTTDWISVLSVSGGAAVVSLLMSIGA 59
Boomer gp34       MLTRSFWIDAAERAARTFAQTAIATLGAGAVDLMTTDWISVLSVSGGAAVVSLLMSIGA 59
                  ************** ******************::***:********************

Llij gp32         ERRGNPGTASATRAVTAA 77
Pacc40 gp32       ERRGNPGTASATRAVTAA 77
PMC gp32          ERRGNPGTASATRAVTAA 77
Che8 gp34         ERRGNPGTASATRAVTAA 77
Ms6 gp4           ERRGNPGTASATRAVTTA 77
Fruitloop gp31    ERRGNPGTASATRAVTTA 77
Tweety gp32       ERRGNPGTASATRAVTAA 77
Ramsey gp35       ERRGNPGTASATRAVTAA 77
Boomer gp34       ERRGNHGTASATRAVTAA 77
                  ***** **********.*
```

B)

Hol from Ms6 Phage

T**AGGAGG**CACCC**ATG**CTGACACGTTCATTCTGGATCGACGCCGCCGAGCGTGCCATACGCACATTCGCCCAAACC
M L T R S F W I D A A E R A I R t F a q t
\+ - - + +

GCGATCGCCACCCTCGGCGCCGGGGCAGTCGACCTGATGACCACCGACTGGATATCGGTGCTGTCCGTGTCCGGC
a I a t L g a g a V D L M T T D W I S V L S V S G
\- -

GGCGCGGCCGTCGTATCACTGCTGATGTCGATCGGCGCCGAACGCCGCGGAAACCCCGGAACGGCGTCGGCCACT
G A A V V S L L M S I G A E R R G N P G T A S A T
\- + +

AGAGCGGTCACCACCGCA**TGA**
R A V T T A *
\+

Gp5 from Ms6 Phage

GAGCGGTCACCACCGC**ATG**ATCT**GGGAA**TCG**GTG**C**GCGAAG**CG**GTG**AACGCGGCGTACCAGCCCGACGACGGT
M I W E S V/M R E A V/M N A A Y Q P D D G
\- + - - -

ATCGACCTGATAGGACTGCTCATCATCGGACTGCCCTCCACCATCGCCGCCATCGGAACAGGGATCGTCGGCGTACTC
I D L I G L L I I G L P S T I A A I G T G I V G V L
\-

ACCGTTCGGGGACAGCGCAAAGGCCGGGAGCGCGCACGCCAGATCGACGCGAAAACCGATGAGATTCACGAGCAG
T V R G Q R K G R E R A R Q I D A K T D E I H E Q
\+ + + + - + + - + - - + -

ACCGTCAACACCCACGACACCAACATGCGTGACGACCTCGACGAGATACGCGATCTGGTGCGCGACGGCTTCAAA
T V N T H D T N M R D D L D E I R D L V R D G F K
\+ - + - - - - + - + - +

CAGATCCAACGCGACATCGGAGGACTGCGGGAGGAGCTGCGAACCGAACGACTGGAACGAATCGAAGGCGACAAA
Q I Q R D I G G L R E E L R T E R L E R I E G D K
\- - + + + - - + - + - + - + +

CGCCGCGACCGG**TAA**
R R D R *
\+ + - +

Gp11 (Hol) from D29 Phage

GGCGTACATCGCC**AGGAA**T**GGAG**CCCT**ATG**AGCCCCAAGATCCGTGAAACGCTCTACTACGTCGGCACTCTCGTC
M S P K I R E T L Y Y V G T L V
\+ + -

CCCGGCATCCTGGGCATCGCCCTGATCTGGGGCGGGATCGACGCGGGCGCAGCCGCGAACATCGGCGACATCGTC
P G I L G I A L I W G G I D A G A A A N I G D I V
\- -

GCTGGCGCTCTCAACCTGGTCGGCGCAGCCGCACCGGCCACGGCCGCTGTCAAGGTCAACCAGCAGCGCAAGGAT
A G A L N L V G A A A P A T A A V K V N Q Q R K D
\+ + + -

GGCACGCTGACCACCTCCCCGGTGGATCAGGTCACCAGGGGCGTCGAGCAGGTGCTCGCGGCCAAGCAGAACGCT
G T L T T S P V D Q V T R G V E Q V L A A K Q N A
\- + - +

GAGGCTGAGGTCGAGCGCGTCAAGCAGGCTCTGGAGTCCGCTGTCAACGGCGCGGTCCCCCAGCTCGGCCCGCTG
E A E V E R V K Q A L E S A V N G A V P Q L G P L
\- - - + + -

GCCAGCCAGATCCTCAACGGCATCCAACCGGCCTACAGCCAGCCGTTCGACCCGCACACGCAGCCCTGGAACCGA
A S Q I L N G I Q P A Y S Q P F D P H T Q P W N R
\- + +

TGA
*

C)

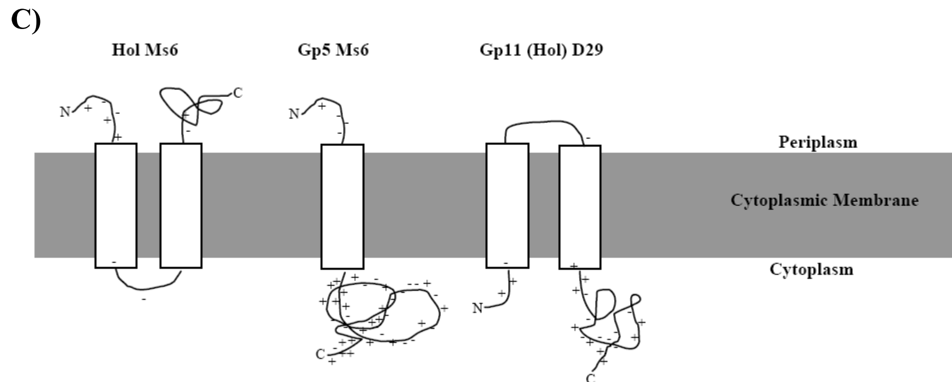


Figure 2. Holin-like proteins of mycobacteriophages Ms6 and D29. A) CLUSTALW alignment of Ms6 Hol (AAG48320) with similar sequences of Pham95 members included in subcluster F1: Llij Gp32 (ABD58248), Pacc40 Gp32 (YP002241616), PMC Gp32 (ABE67533), Che8 Gp34 (NP817372), Fruitloop Gp31 (YP002241716), Tweety Gp32 (YP001469265), Ramsey Gp35 (YP002241822) and Boomer Gp34 (YP002014250); the primary accession numbers of the UniProtKB/TrEmbl database are given in parentheses. Identical (*) and highly similar (:) amino acids are indicated. Numbers refer to the amino acids positions. The two TMDs are indicated in a grey box. **B)** Sequences of genes coding for the class II holin (*hol*) and class III holin (*gp5*) of Ms6 and class II holin of D29 (*gp11*). Charged residues are indicated by + or -. TMDs are indicated in grey. Amino acid residues in the SAR domain of Hol that are predicted to be weakly hydrophobic are shown in lowercase. Potential translation start codons and corresponding Shine-Dalgarno sequences are in bold and underlined (RBS). **C)** Topological model for Ms6 Hol (N-out, C-out), Gp5 (N-out, C-in) and D29 Gp11 (Hol) (N-in, C-in).

Expression of holin/endolysin pairwise combinations in E. coli

In contrast to what happens with phage endolysins that possess a narrow range of activity regarding the infected bacteria, holins are not species specific and do not specifically interact with the endolysins (Young, 2002; São-José *et al.*, 2003). As the canonical holins, such as λS, form very large nonspecific holes that allow fully-folded unrelated endolysins to pass through the membrane and attack

the murein, we expressed in *E. coli* different pairwise combinations of endolysins/holins in an attempt to clarify the role of Hol and Gp5 in bacterial lysis. We observed that the concomitant expression of Ms6 LysA with Hol or Gp5 was not sufficient to support a lysis phenotype in *E. coli*. Co-expression of LysA with both Gp4 and Gp5 was much more deleterious to cell viability than when LysA was co-expressed with Hol or Gp5 alone (Fig. 3A) although a marked decrease in optical density was not observed. The fact that the lack of a lysis phenotype was a consequence of the inability of Hol or Gp5 to form lesions on the cytoplasmic membrane large enough to allow the passage of LysA to the periplasm was further confirmed by concomitantly expressing the Ms6 endolysin LysA with the D29 Gp11 holin. Lysis of *E. coli* was observed, beginning 20 minutes after induction which suggests that the D29 holin is functional in *E. coli* and allows the access of LysA to the peptidoglycan (Fig. 3A). In addition, the Ms6 Hol but not Gp5 allows the access of the lambda transglycosylase to the murein as demonstrated by complementation of a λS mutant (Garcia *et al*., 2002; Catalão M. J., unpublished results), and unlike Hol, Gp5 was unable to promote the cytoplasm release of λR, the cytosolic endolysin of phage λ (Fig. 3B). We interpret this to mean that the Ms6 Hol and Gp5 alone make holes too small to allow the passage of LysA in *E. coli*.

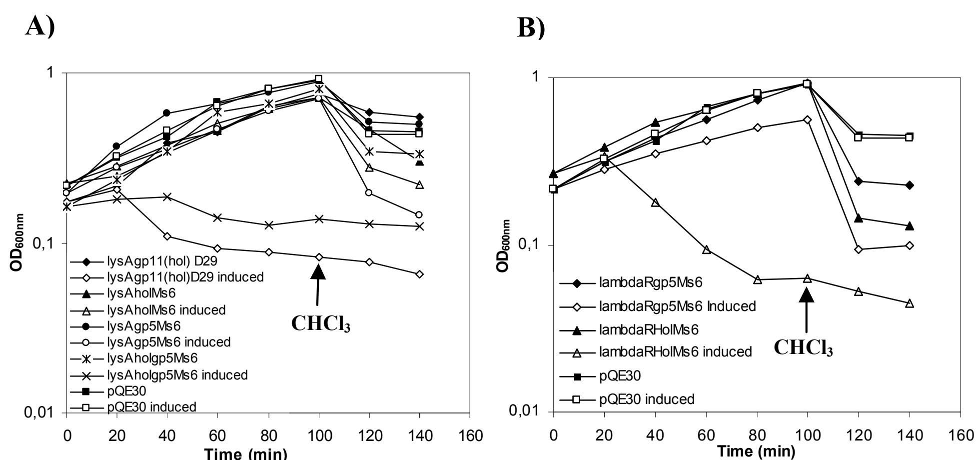


Figure 3. A) Effect of the expression of phage endolysin/holin pairwise combinations on *E. coli* growth. **B)** Effect of the expression of pairwise combinations of the λR endolysin with the Ms6 holin-like proteins Hol and Gp5 in *E. coli*. *E. coli* JM109 cells carrying plasmid pQE30 containing no insert or cloned genes were grown in LB broth at 37 ºC to an OD_{600nm} of 0.2. At time zero, transcription of cloned lysis genes was induced by addition of 1 mM IPTG. At the time indicated by the arrow, 2% $CHCl_3$ was added to the cultures. Results are averages of triplicate experiments.

Crosslinking of Hol and Gp5 in the* E. coli *cell membrane

To identify the oligomeric states of Ms6 Hol and Gp5, the membrane fractions from *E. coli* expressing Hol or Gp5 (fused to a S-tag at the N-terminus and a His_6-tag at the C-terminus) from a derivative plasmid of pET29a, were collected 60 minutes after induction, proteins were extracted from the membranes with Triton X-100 and subjected to crosslinking with the water-soluble membrane-impermeant, homobifunctional sulfo-*N*-hydroxy-succinimide ester, BS^3. This crosslinker forms covalent amide bonds with accessible α-amine groups present on the N-termini of proteins and ε-amines on lysine residues. The presence of a band with 24.8 kDa in the absence of the crosslinker shows that Hol forms SDS-resistant dimers during membrane extraction with Triton X-100 (Fig. 4A). Furthermore, Hol specific bands corresponding to molecular masses of 12.4 and 24.8, 37.2, 49.6 kDa and 62 kDa, up to pentamers could be detected by Western-blotting analysis (Fig. 4A). Gp5, which does not complement a λS defective mutant phage, formed only trimers but not higher oligomers under the same conditions (Fig. 4B). This result might help explaining its inability to support an efficient lysis of *E. coli* when co-expressed with different endolysins, as the ability of holin molecules to oligomerize is essential for the lytic step in holin function (Zagotta and Wilson, 1990; Gründling *et al*., 2000a; Gründling *et al*., 2000b).

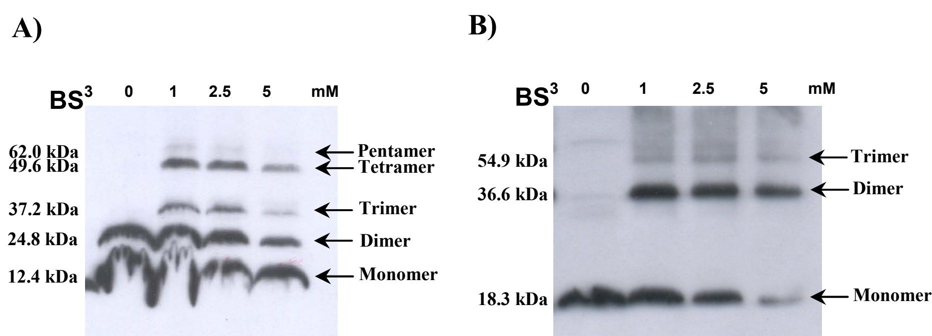


Figure 4. Ms6 Hol (A) or Gp5 (B) oligomerization. Proteins from membranes of *E. coli* BL21 (DE3) expressing Hol or Gp5 were extracted with Triton X-100 and treated with different BS^3 concentrations as described in Material and Methods. Proteins were detected by Western-blotting with an anti-His_6 antibody. Predicted molecular masses are indicated to the left of the panels. Oligomerization bands are indicated by arrows.

Evidence for an interaction between Hol and Gp5

Despite the toxicity observed after expression induction of Hol or Gp5 membrane proteins in *E. coli*, simultaneous expression of these proteins attenuates the lethal effect (Fig. 1A). For bacteriophage lambda, it has been proposed that the lysis inhibitor S_{107} inhibits lysis through intermolecular interaction with the lysis effector S_{105} (Gründling *et al.*, 2000c). Accordingly, the ability of Gp5 to inhibit Hol lethal effect suggests that it may interact with the holin. To demonstrate this, both proteins were expressed from the same vector in *E. coli* with Hol fused to an S-tag at the N-terminus and Gp5 fused to a His_6-tag at the C-terminus. 60 minutes post-induction, cell membranes were collected by centrifugation, proteins were extracted with 1% Triton X-100 and this fraction was subjected to chemical crosslinking with BS^3. In the absence of the crosslinker we detected a band with a molecular mass of 14.9 kDa corresponding to the size predicted for Gp5-His_6 monomer and a faint band with ~28 kDa corresponding to an interaction between

an S-tag Hol monomer and a Gp5-His_6 monomer (Fig. 5). Crosslinking using 1 mM of BS^3 revealed an additional band of 30 kDa corresponding to the Gp5 dimer. The increase in BS^3 concentration to 5 mM led to the appearance of a band with ~55 kDa as a result of oligomer formation between Gp5 homodimer and Gp4 homodimer. This result suggests that Gp5 interacts with the Ms6 holin and suggests that it may control Hol function during the lytic cycle.

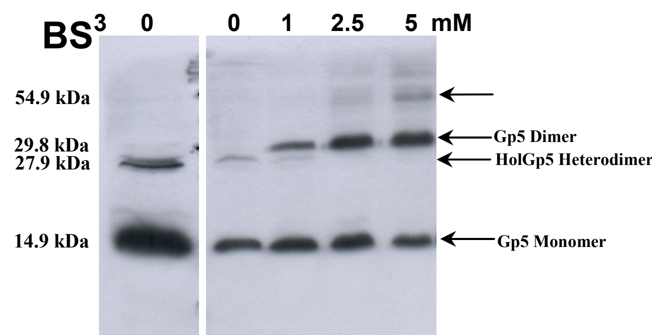


Figure 5. Interaction between Hol and Gp5 of mycobacteriophage Ms6. Crosslinking and sample preparation for Western-blotting analysis was performed as described to Fig. 4. Proteins were detected with an anti-His_6 antibody. In the absence of BS^3, Gp5 monomer and HolGp5 heterodimer are detected. Predicted molecular masses are indicated to the left of the panels. Oligomerization bands are indicated by arrows.

Construction of Ms6 mutant phages defective for lysis

Concerning the possibility that Ms6 Hol and Gp5 may behave differently in the heterologous *E. coli* host and in Ms6 natural host, *M. smegmatis*, we constructed Ms6 mutant phages defective for Hol, Gp5 or both proteins synthesis. Using the mycobacteriophage recombineering system (Marinelli *et al.,* 2008) we constructed internal in-frame deletions of the Ms6 *hol*, *gp5* or *hol* and *gp5* in the Ms6 wild-type phage and of *hol* and *gp5* genes in a previously constructed $Ms6_{\Delta gp1}$

phage (Catalão *et al*., 2010) by allelic gene replacement. Co-electroporation of a 200 bp substrate containing 100 bp flanking *hol or gp5* and $Ms6_{wt}$ genomic DNA or a 200 bp substrate containing 100 bp flanking *holgp5* and $Ms6_{\Delta gp1}$ genomic DNA into a *M. smegmatis* recombineering strain yielded several primary plaques for each transformation. 30 of the primary plaques of each transformation were screened by PCR with primers flanking the deletions and those containing detectable levels of the mutant phage mixed with the wild-type were serially diluted and re-plated. Both the secondary lysate and individual secondary plaques were screened by PCR with the same pair of primers. As the mutant allele is still present in all the lysates screened except for the mutant phage $Ms6_{\Delta gp1holgp5}$, all mutant phages with this exception are viable. To recover the deletion mutants, 30 secondary individual plaques of the re-plating were picked and screened by PCR with primers flanking the deletions. Even though we have not yet been able to recover a purified mutant derivative of phage $Ms6_{\Delta holgp5}$, probably reflecting poor viability of the mutant, pure mutants of $Ms6_{\Delta hol}$ and $Ms6_{\Delta gp5}$ were readily identified in a high efficiency and in the absence of any selection (Fig. 6). In view of the fact that the absence of Hol or Gp5 from the Ms6 virion has no apparent effect on phage viability and our recent observations that the accessory lysis protein Gp1 is required for a normal burst of infective phage particles (Catalão *et al*., 2010), we further investigate its function during the Ms6 lytic cycle. We constructed an internal in-frame deletion of the *gp1* gene in defective phages $Ms6_{\Delta hol}$ and $Ms6_{\Delta gp5}$ using the same recombineering strategy described above and pure mutants of $Ms6_{\Delta gp1hol}$ and $Ms6_{\Delta gp1gp5}$ were isolated after PCR screening of secondary individual plaques (Fig. 6).

These results demonstrate that $Ms6_{\Delta hol}$, $Ms6_{\Delta gp5}$, $Ms6_{\Delta gp1hol}$ and $Ms6_{\Delta gp1gp5}$ are viable, neither *hol* nor *gp5* are essential for plaque formation and that an Ms6 mutant phage lacking the *holin* and *gp5* lysis genes is able to infect and lyse mycobacteria.

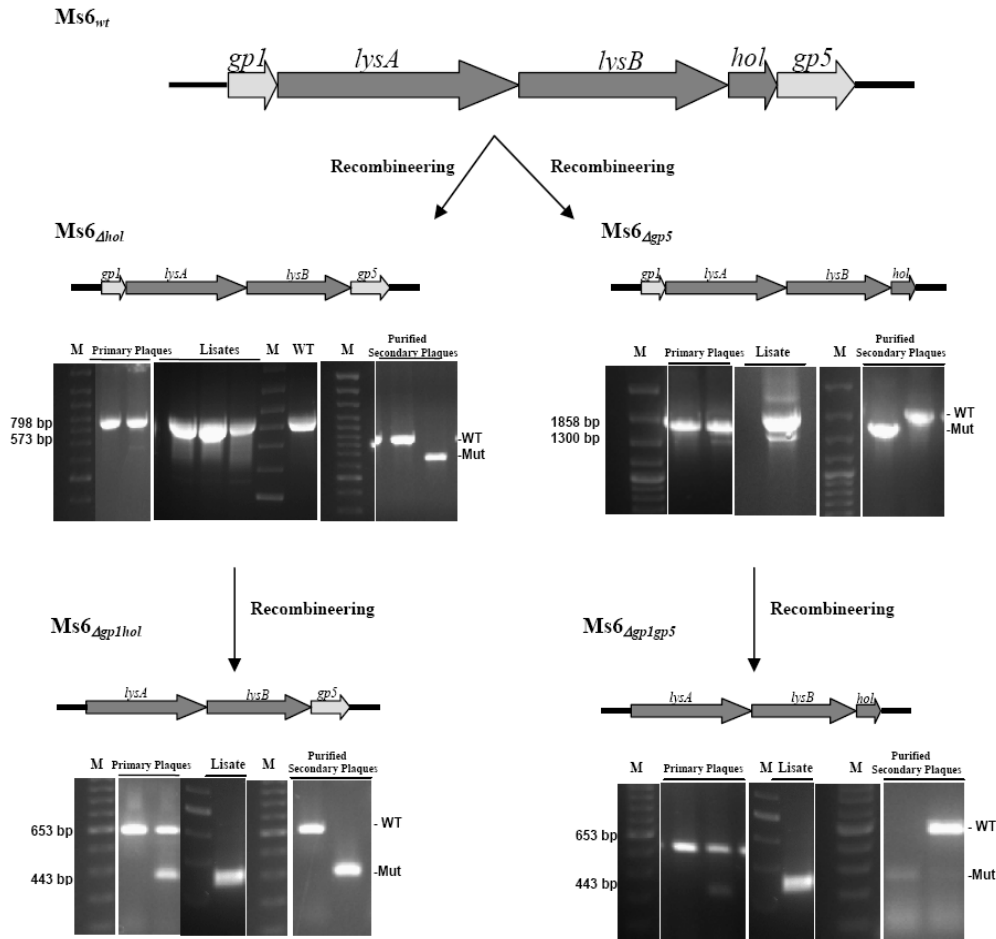


Figure 6. Strategy for construction of Ms6 lysis genes deletion mutants. A 200 bp dsDNA substrate that has 100 bp homology flanking the deletion was designed. Following co-electroporation of the 200 bp substrates and genomic DNA (of $Ms6_{wt}$ to obtain $Ms6_{\Delta hol}$/$Ms6_{\Delta gp5}$ mutants or of $Ms6_{\Delta hol}$/$Ms6_{\Delta gp5}$ to obtain $Ms6_{\Delta gp1hol}$/$Ms6_{\Delta gp1gp5}$ mutants), primary plaques were recovered to identify a mixed plaque containing wild-type and mutant phages. The mixed primary plaque was diluted and plated; the lysate was screened to check for phage viability and purified secondary plaques were screened to identify homogenous deletion mutants.

Ms6 holin-like proteins are not required for mycobacteria lysis

One step growth curves and determination of phage growth parameters (latent period, rise period and burst size) were carried out to compare the Ms6 mutant phages replication cycle. Results are summarized in Table 3. The one step growth experiment (Fig. 7A) shows that Hol and Gp5, although nonessential for lysis, have an effect on the lysis timing since an Ms6 *hol* deletion mutant caused slightly accelerated lysis (80 minutes), whereas an Ms6 *gp5* deletion mutant delayed lysis (170 min) which is consistent with holin function. These lysis times correspond to the latent time represented in Fig. 7A in addition to the initial 50 min of adsorption and were compared to the Ms6 wild-type phage (110 min) in the same experimental conditions. Thus, the absence of *hol* or *gp5* in the infecting virion has an evident effect in the timing of lysis. Single-burst experiments were done to compare the viable progeny released from single cells infected with $Ms6_{wt}$ or the mutant phages. The number of infective particles released after infection with the $Ms6_{\Delta hol}$ phage is lower than in an $Ms6_{wt}$ infection, whereas after infection with the $Ms6_{\Delta gp5}$ phage an increase in the burst size was detected. In our experimental conditions, when $Ms6_{wt}$ infects *M. smegmatis* mc^2155 there is an average of 149 viable phages released from one bacterium while infection with $Ms6_{\Delta hol}$ or $Ms6_{\Delta gp5}$ yielded an altered burst size of approximately 115 or 221 viable phages per infected cell, respectively. Deletion of the *gp1* gene from $Ms6_{\Delta hol}$ or $Ms6_{\Delta gp5}$, drastically reduced the burst size to ~28 phages/infected cell or ~77 phages/infected cell for $Ms6_{\Delta gp1hol}$ or $Ms6_{\Delta gp1gp5}$ phages, respectively, which is in agreement with our previously results (Catalão *et al*., 2010). When analysing the plating ability and the plaque size of the lysis defective phages we observed that all, except $Ms6_{\Delta gp5}$ produce smaller plaques with no size variation than those produced by $Ms6_{wt}$ phage (Fig. 7B). The size of the plaques produced by $Ms6_{\Delta gp1hol}$ was < 1 mm whereas the plaques produced by the $Ms6_{\Delta gp5}$ mutant were very large

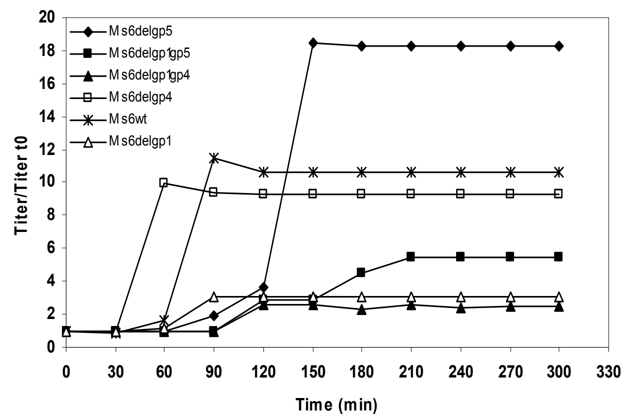
with 4-5 mm of diameter in agreement with a lower and a larger burst size, respectively, when compared to $Ms6_{wt}$ phage.

Table 3. Viability and phage growth parameters of mycobacteriophage Ms6 and lysis genes deletion derivatives.

Phage	Plating Ability	Burst Size	Latent time	Plaque Size
$Ms6_{wt}$	Positive	149±32	110 min	2-3 mm
$Ms6_{\Delta gp1}$	Positive	45±13	110 min	~1 mm
$Ms6_{\Delta hol}$	Positive	115±39	80 min	1-2 mm
$Ms6_{\Delta gp5}$	Positive	221±56	170 min	4-5 mm
$Ms6_{\Delta gp1hol}$	Positive	28±6	140 min	1-2 mm
$Ms6_{\Delta gp1gp5}$	Positive	77±27	200 min	< 1 mm
$Ms6_{\Delta holgp5}$	Positive	-	-	-
$Ms6_{\Delta gp1holgp5}$	Negative	-	-	-

Taken together these results suggest that in addition to *hol*, *gp5* encode a holin-like protein and they must act in concert to control the timing of lysis. Furthermore, as previously observed (Catalão *et al.*, 2010) the presence of Gp1 in the virion is crucial to obtain a normal burst of infective phages although it has no influence on duration of the latent time of the lytic cycle.

A)



B)

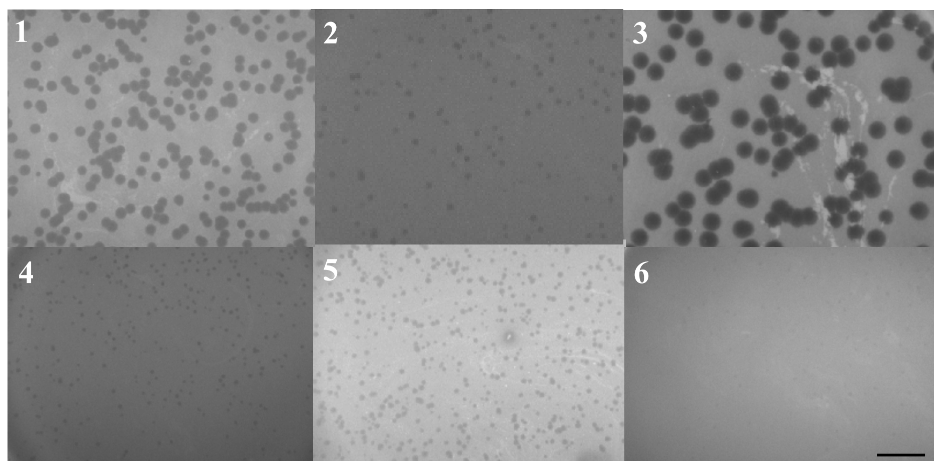


Figure 7. A) One step growth curves of mycobacteriophage Ms6 and lysis genes deletion derivatives. For each curve the titers measured were divided by the titer at t=0 for normalization (titer/titer t_0). Results are averages of three independent experiments. **B)** Plating ability of the different lysis genes mutant bacteriophages. **1.** *M. smegmatis* infected with $Ms6_{wt}$ phage; **2.** *M. smegmatis* infected with $Ms6_{\Delta hol}$; **3.** *M. smegmatis* infected with $Ms6_{\Delta gp5}$; **4.** *M. smegmatis* infected with $Ms6_{\Delta gp1}$; **5.** *M. smegmatis* infected with $Ms6_{\Delta gp1hol}$; **6.** *M. smegmatis* infected with $Ms6_{\Delta gp1gp5}$. Scale bar indicates 1 cm.

Discussion

Even though the mechanisms underlying mycobacteriophage lysis of mycobacteria are poorly understood, recent work has contributed significantly to the progress in the field (Gil *et al*., 2008; Payne *et al*., 2009; Henry *et al*., 2010; Gil *et al*., 2010; Catalão *et al*., 2010). Notwithstanding, the exact mechanism by which the lysis effectors LysA and LysB are localized to their substrates remains elusive in the majority of the mycobacteriophages. Very recently, we have identified the product of *gp1* gene as a chaperone-like protein that specifically interacts with the

endolysin and is involved in its translocation across the cytoplasmic membrane, positioning the protein in the periplasm, independently of holin function. Moreover, removal of Gp1 function in mycobacteriophage Ms6 showed that although not essential for plaque formation, the protein is required for efficient phage release (Catalão *et al.*, 2010). These data, together with the previous reported absence of lysis when LysA and Hol were co-expressed in *E. coli* (Garcia *et al.*, 2002) led us to investigate the role of two holin-like proteins *hol* and *gp5* in a heterologous host (*E. coli*) and in the mycobacteriophage Ms6 infection context.

The majority of phage lysis genes are clustered and transcribed in the order holin-endolysin. In many cases the genes are not linked (i.e. T4) (Ramanculov and Young, 2001), and in many phages of *Streptococcus thermophilus* (Sheehan *et al.*, 1999) and in phage Av-1, which infects the Gram-positive bacteria *Actinomyces naeslundii* (Delisle *et al.*, 2006), two putative holin genes precede the endolysin gene. Another gene arrangement was found in the *Bacillus subtilis* prophage PBSX; it was proposed that two open reading frames preceding the endolysin *xlyA,* designated *xhlA* and *xhlB*, encode polypeptides that associate in the membrane to form a functional holin complex that allow XlyA access to the peptidoglycan (Longchamp *et al.*, 1994; Krogh *et al.*, 1998). In *Staphylococcus aureus* phage P68 the endolysin gene *lys16* was identified within the structural genes, and embedded in the -1 register at the distal end of *lys16*, a putative holin gene, *hol15*, was further recognized. A second putative holin gene, *hol12*, was later identified at the end of the structural genes (Takáč *et al.*, 2005). In mycobacteriophage Ms6 another endolysin-holin arrangement was found (Garcia *et al.*, 2002). The endolysin *lysA* (*gp2*) is located upstream of the identified holin gene (*gp4*). A second putative holin-like gene (*gp5*) which may play a role in lysis was also identified using bioinformatic analysis. The start codon of *gp5* overlaps the stop codon of the upstream-located *hol* gene within four nucleotides (Fig. 2B) but in a different reading frame. This lysis genes arrangement was also found in phage µ1/6

(Farkašovská and Godány, 2008), and in several other mycobacteriophages included in subcluster F1 (Hatfull *et al*., 2010). Examination of the predicted amino acid sequence of Hol showed that it possesses two TMDs. The first TMD has characteristics of a SAR domain with a high percentage of hydrophobic or uncharged polar residues. The presence of a SAR domain followed by a typical TMD has been described for the characterized pinholin of the lambdoid phage 21 (Park *et al*., 2007; Summer *et al*., 2010) and suggests that Ms6 Hol may be a pinholin. Pinholins are a recently discovered class of holins that make small holes in the host membrane sufficient to depolarize it and allow exit of SAR endolysins, but are not large enough to allow escape of canonical cytoplasmic endolysins (Park *et al*., 2007; Pang *et al*., 2009). However, although Hol was unable to support the Ms6-mediated lysis of *E. coli* cells (Garcia *et al*., 2002), unlike the pinholin of phage 21 S^{21}, it was able to promote the release of R, the cytosolic endolysin from phage λ (Park *et al*., 2007). Furthermore, Hol was found to oligomerize (up to at least pentameric order in detergent) as described for the S^{21} pinholin (Pang *et al*., 2009) but the ultimate degree of oligomerization is unknown. Although the Ms6 endolysin is not endowed with an N-terminal SAR domain as should be expected due to the presence of a pinholin, it interacts with Gp1, a chaperone-like protein involved in its delivery to the peptidoglycan (Catalão *et al*., 2010). Recently, according to the sequence similarity to putative holin genes from mycobacteriophages of subcluster F1 (Hatfull *et al*., 2010) the Ms6 holin was included in Pham95. Interestingly, all phages belonging to subcluster F1 where this family of holins had been assigned possess the *gp1* gene upstream of the endolysin. Although additional holin genes have yet to be identified bioinformatically or by functional studies, as in the majority of the sequenced mycobacteriophage genomes there is no closely linked and easily recognizable holin gene, another family of holin genes has been identified (Pham1981), in phages included in clusters A2, E and I (Hatfull *et al*., 2010). Differing from what is observed in Ms6 and other

related mycobacteriophages, the mycobacteriophage D29 (subcluster A2) does not possess homologues of the accessory lysis proteins Gp1 or Gp5. The lytic genes are clustered together, with the putative holin gene localized between the LysA and LysB genes (São-José *et al*., 2003; Hatfull *et al*., 2010). Bioinformatic analysis of the amino acid sequence of Gp11 (Hol) identified two transmembrane domains and a cytoplasmatic highly charged and hydrophilic C-terminal domain. The holin of D29 facilitated LysA-mediated lysis in *E. coli*, which suggests that, unlike the Ms6 Hol, it belongs to the canonical group of holins that form large holes and allow fully folded endolysins to pass through the cytoplasmic membrane. In addition to the Ms6 Hol, another holin-like protein (Gp5) was also identified within the lytic operon. It possesses a single periplasmic TMD and a very highly charged and hydrophilic C-terminal domain. Gp5 was found to be unable to support both LysA- or λR-mediated lysis in *E. coli* and in contrast to Hol, oligomerization appeared to be blocked at the trimer stage in detergent (as for Hol, the final degree of oligomerization is not yet known). Despite the toxicity observed when independently expressing the integral membrane proteins Hol and Gp5 in *E. coli*, simultaneous expression of these proteins attenuates the lethal effect which suggested that they may interact. For bacteriophage lambda, it has been proposed that the lysis inhibitor S_{107} inhibits lysis through dimeric interactions with the lysis effector S_{105} (Gründling *et al*., 2000c). Indeed, using chemical crosslinking we were able to obtain biochemical evidence for a direct interaction between Hol and Gp5; however, the exact mechanism by which Gp5 acts to control Hol function remains elusive. Owing to the concerns that exist when expressing holin genes from strong inducible promoters, we constructed different Ms6 mutant phages, deleted in the holin-like genes. Indeed, it is well known that a membrane protein overproduced from a multicopy plasmid can insult the membrane sufficiently to cause release of cytoplasmic endolysins (Young and Wang, 2006). Although more time consuming, this strategy has two important advantages: i) it allows the

function of these proteins to be examined in their natural host and ii) the level of each gene product in the cell is the same as that produced when the wild-type prophage is induced: expression of the late operon is very tightly regulated because of the lethal consequences of prophage induction. When we analysed the holin genes-deleted mutant phages infection cycle by one step growth curves we observed that the gene products of *hol* and *gp5*, although nonessential for phage viability, appear to play a role in controlling the timing of lysis. The $Ms6_{\Delta hol}$ deletion mutant caused slightly accelerated lysis whereas an $Ms6_{\Delta gp5}$ deletion mutant delayed lysis, which is consistent with holin function. None of the holin genes have a dual-start motif, indicating that the timing of *M. smegmatis* lysis by Ms6 may not be dependent on an antiholin-holin interaction, such as occurs in the phage λ S_{107}-S_{105} system (Graschopf and Bläsi, 1999). We also considered whether Gp5 could act as an antiholin. Antiholins generally delay phage lysis in order to optimize progeny phage production. This should have been observed if Gp5 acts as an antiholin and its absence would result in earlier lysis and smaller plaques due to premature lysis. Unexpectedly, this phenotype was observed when Hol was deleted from the lytic operon. Contrary, deletion of Gp5 delayed lysis and resulted in very large plaques due to an increase in the burst size. Not all dsDNA phages utilize an antiholin to regulate lysis timing since some, apparently simply rely on delaying expression of their holin genes (Loessner *et al.*, 1998). These results suggest that *hol* and *gp5* encode holin proteins whose combined action could play the role of a holin and that expression of both proteins is necessary to effect host cell lysis at the correct and programmed timing as described for other phages (Takáč *et al.*, 2005; Delisle *et al.*, 2006). Moreover, interaction of Gp5 with Hol may contribute to very precise adjustment of the timing of hole formation and to keep the infected cell productive allowing the assembly of more virions.

Deletion of the chaperone-like protein *gp1*, in both deletion mutants $Ms6_{\Delta hol}$ *and* $Ms6_{\Delta gp5}$ was catastrophic for lysis, with more than five-fold reduction

of the burst size, even though the mutant phages are viable and could be isolated. Remarkably, although Gp5 was unable to allow endolysin-mediated lysis in *E. coli*, a mutant phage lacking both *gp1* and *hol* was able to infect *M. smegmatis* cells and undergo lysis. Although lysis is generally considered an essential phenotype, deletion of genes underlying lysis is not as likely to stop replication entirely. It is expected that there will be many possible pathways to *de novo* lysis (Heineman *et al*., 2009). We conclude that for mycobacteriophage Ms6 and related mycobacteriophages, the presence of the endolysin in addition to one of the lytic genes, *gp1*, *hol* or *gp5*, is sufficient for a lysis phenotype. However, this results in dramatic changes in the infective cycle and to lower viability of the mutant phages. The presence of the mycobacterium-specific lysis factors Gp1 and Gp5 that are restricted to mycobacteriophages (Garcia *et al*., 2002; Hatfull *et al*., 2010) may confer a selective advantage not only for fitness under different environmental conditions but also as an alternative to lysis exclusively holin-dependent: it has been shown that single missense changes within the holin proteins can have a profound effect on both the process of host lysis and its timing, unpredictably advancing or retarding the instant of triggering (Gründling *et al*., 2000a; Young and Wang, 2006). With this study we hope to contribute to a better understanding of lysis timing regulation by mycobacteriophages.

Acknowledgments

We would like to thank Dr. Graham Hatfull, Dr. Julia van Kessel and Dr. Laura Marinelli (University of Pittsburgh, USA) for supplying plasmid pJV53 and for technical assistance with the recombineering experiments. This work was supported by funds provided by FCT (Fundação para a Ciência e Tecnologia) - PTDC/SAU-FCF/73017/2006. Maria João Catalão and Filipa Gil are the recipients of FCT Ph.D. fellowships (SFRH/BD/24452/2005) and (SFRH/BD/29167/2006).

References

Bläsi U. and Young R. (1996) Two beginnings for a single purpose: the dual-start holins in the regulation of phage lysis. *Mol Microbiol* **21(4):**675-682.

Catalão M. J., Gil F., Moniz-Pereira J. and Pimentel M. (2010) A chaperone-like protein is involved in the endolysin delivery to the peptidoglycan. *Mol Microbiol* (In final revision).

Delisle A. L., Barcak G. J. and Guo M. (2006) Isolation and expression of the lysis genes of *Actinomyces naeslundii* phage Av-1. *Appl Environ Microbiol* **72(2):**1110-1117.

Farkasovská J. and Godány A. (2008) The lysis system of the *Streptomyces aureofaciens* phage μ1/6. *Curr Microbiol* **57(6):**631-637.

Garcia M., Pimentel M. and Moniz-Pereira J. (2002) Expression of mycobacteriophage Ms6 lysis genes is driven by two sigma (70)-like promoters and is dependent on a transcription termination signal present in the leader RNA. *J Bacteriol* **184:**3034-3043.

Gil F., Catalão M. J., Moniz-Pereira J., Leandro P., McNeil M. and Pimentel M. (2008) The lytic cassette of mycobacteriophage Ms6 encodes an enzyme with lipolytic activity. *Microbiology* **154(Pt 5):**1364-1371.

Gil F., Grzegorzewicz A. E., Catalão M. J., Vital J., McNeil M. and Pimentel M. (2010) The mycobacteriophage Ms6 LysB specifically targets the outer membrane of *Mycobacterium smegmatis*. *Microbiology* **156:**1497-1504.

Graschopf A. and Bläsi U. (1999) Molecular function of the dual-start motif in the lambda S holin. *Mol Microbiol* **33(3):**569-582.

Gründling A., Bläsi U. and Young R. (2000a) Biochemical and genetic evidence for three transmembrane domains in the class I holin, lambda S. *J Biol Chem* **275(2):**769-776.

Gründling A., Bläsi U. and Young R. (2000b) Genetic and biochemical analysis of dimer and oligomer interactions of the lambda S holin. *J Bacteriol* **182(21):**6082-6090.

Gründling A., Smith D. L., Bläsi U. and Young R. (2000c) Dimerization between the holin and holin inhibitor of phage lambda. *J Bacteriol* **182(21):**6075-6081.

Gründling A., Manson M. D. and Young R. (2001) Holins kill without warning. *Proc Natl Acad Sci USA* **98:**9348-9352.

Hatfull G. F., Jacobs-Sera D., Lawrence J. G., Pope W. H., Russell D. A., Ko C. C., Weber R. J., Patel M. C., Germane K. L., Edgar R. H., Hoyte N. N., Bowman C. A., Tantoco A. T., Paladin E. C., Myers M. S., Smith A. L., Grace M. S., Pham T. T., O'Brien M. B., Vogelsberger A. M., Hryckowian A. J., Wynalek J. L., Donis-Keller H., Bogel M. W., Peebles C. L., Cresawn S. G. and Hendrix R. W. (2010) Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. *J Mol Biol* **397(1):**119-143.

Heineman R. H., Bull J. J. and Molineux I. J. (2009) Layers of evolvability in a bacteriophage life history trait. *Mol Biol Evol* **26(6):**1289-1298.

Henry M., O´Sullivan O., Sleator R. D., Coffey A., Ross R. P., McAuliffe O. and O´Mahony J. M. (2010) *In silico* analysis of Ardmore, a novel mycobacteriophage isolated from soil. *Gene* **453(1-2):**9-23.

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

Krogh S., Jørgensen S. T. and Devine K. M. (1998) Lysis genes of the *Bacillus subtilis* defective prophage PBSX. *J Bacteriol* **180(8):**2110-2117.

Loessner M. J., Gaeng S., Wendlinger G., Maier S. K. and Scherer S. (1998) The two-component lysis system of *Staphylococcus aureus* bacteriophage Twort: a large TTG-start holin and an associated amidase endolysin. *FEMS Microbiol Lett* **162(2):**265-274.

Loessner M. J. (2005) Bacteriophage endolysins: current state of research and applications. *Curr Opin Microbiol* **8:**480-487.

Longchamp P. F., Mauël C. and Karamata D. (1994) Lytic enzymes associated with defective prophages of *Bacillus subtilis*: sequencing and characterization of the region comprising the *N*-acetylmuramoyl-L-alanine amidase gene of prophage PBSX. *Microbiology* **140(Pt 8):**1855-1867.

Marinelli L. J., Piuri M., Swigonová Z., Balachandran A., Oldfield L. M., van Kessel J. C. and Hatfull G. F. (2008) BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. *PLoS One* **3(12):**e3957.

Miller J. H. (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, N.Y., p. 113-129.

Pang T., Savva C. G., Fleming K. G., Struck D. K. and Young R. (2009) Structure of the lethal phage pinhole. *Proc Natl Acad Sci USA* **106(45):**18966-18971.

Park T., Struck D. K., Deaton J. F. and Young R. (2006) Topological dynamics of holins in programmed bacterial lysis. *Proc Natl Acad Sci USA* **103(52):**19713-19718.

Park T., Struck D. K., Dankenbring C. A. and Young R. (2007) The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. *J Bacteriol* **189(24):**9135-9139.

Payne K., Sun Q., Sacchettini J. and Hatfull G. F. (2009) Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. *Mol Microbiol* **73(3):**367-381.

Portugal I., Anes E. and Moniz-Pereira J. (1989) Temperate mycobacteriophage from *M. smegmatis*. *Acta Leprol* **7:**243-244.

Ramanculov E. and Young R. (2001) Genetic analysis of the T4 holin: timing and topology. *Gene* **265(1-2):**25-36.

Sambrook J. and Russell D. W. (2001). Molecular Cloning: a Laboratory Manual, 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Press.

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

Sheehan M. M., Stanley E., Fitzgerald G. F. and van Sinderen D. (1999) Identification and characterization of a lysis module present in a large proportion of bacteriophages infecting *Streptococcus thermophilus*. *Appl Environ Microbiol* **65(2):**569-577.

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

Summer E. J., Enderle C. J., Ahern S. J., Gill J. J., Torres C. P., Appel D. N., Black M. C., Young R. and Gonzalez C. F. (2010) Genomic and biological analysis of phage Xfas53 and related prophages of *Xylella fastidiosa*. *J Bacteriol* **192(1):**179-190.

Takáč M., Witte A. and Bläsi U. (2005) Functional analysis of the lysis genes of *Staphylococcus aureus* phage P68 in *Escherichia coli. Microbiology* **151(Pt 7):**2331-2342.

Vukov N., Moll I., Bläsi U., Scherer S. and Loessner M. J. (2003) Functional regulation of the *Listeria monocytogenes* bacteriophage A118 holin by an intragenic inhibitor lacking the first transmembrane domain. *Mol Microbiol* **48(1):**173-186.

Walker J. T. and Walker D. H. Jr. (1980) Mutations in coliphage P1 affecting host cell lysis. *J Virol* **35(2):**519-530.

Wang I.-N., Smith D. L. and Young R. (2000) Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* **54:**799-825.

Wang I.-N., Deaton J. and Young R. (2003) Sizing the holin lesion with an endolysin-beta-galactosidase fusion. *J Bacteriol* **185(3):**779-787.

Young R. (1992) Bacteriophage lysis: mechanism and regulation. *Microbiol Reviews* **56:**430-481.

Young R., Wang I.-N. and Roof W. D. (2000) Phages will out: strategies of host cell lysis. *Trends Microbiol* **8:**120-127.

Young R. (2002) Bacteriophage holins: deadly diversity. *J Mol Microbiol Biotechnol* **4:**21-36.

Young R. and Wang I.-N. (2006) Phage lysis. *In*, The Bacteriophages. R. Calendar (Ed.) 2nd ed. Oxford University Press, New York., p. 104-125.

Zagotta M. T. and Wilson D. B. (1990) Oligomerization of the bacteriophage lambda S protein in the inner membrane of *Escherichia coli*. *J Bacteriol* **172(2):**912-921.

Ziermann R., Bartlett B., Calendar R. and Christie G. E. (1994) Functions involved in bacteriophage P2-induced host cell lysis and identification of a new tail gene. *J Bacteriol* **176(16):**4974-4984.

4

Sequence Analysis, Characterization and Lytic Activity of Lysin A, the Cell Wall Hydrolase Encoded by the Mycobacteriophage Ms6

M. J. Catalão, F. Gil, J. Moniz Pereira, and M. Pimentel

Centro de Patogénese Molecular, Unidade dos Retrovírus e Infecções Associadas, Faculty of Pharmacy, University of Lisbon, Portugal

Manuscript in preparation

2010

Summary

The ability of mycobacteriophages to lyse and destroy mycobacteria has stimulated renewed research on their potential use as therapeutic agents against tuberculosis, one of the main causes of morbidity and mortality throughout the world. The mycobacteriophage Ms6 accomplishes lysis by producing two lytic enzymes, Lysin A (LysA) that cleaves the amide bond between the *N*-acetylmuramic acid and L-alanine residues in the oligopeptide crosslinking chains of the peptidoglycan and Lysin B (LysB) a mycolylarabinogalactan esterase that hydrolyzes the mycolic acids from the mycolyl-arabinogalactan-peptidoglycan complex. Ms6 LysA possesses a central peptidoglycan recognition protein (PGRP) super-family conserved domain with its 3 sites: amidase catalytic site, a Zn binding site and a substrate binding site. Examination of the endolysin ($lysin_{384}$) DNA sequence revealed the existence of an overlapped gene ($lysin_{241}$) encoded in the same reading frame and preceded by a consensus ribosome-binding site. Even though LysA revealed to be essential for *M. smegmatis* lysis, an Ms6 mutant lacking the longer (Lysin_{384}) or the shorter (Lysin_{241}) endolysin was shown to be viable, but defective in the normal timing, progression and completion of host cell lysis. In addition, both endolysins demonstrated broad lytic activity against various Gram-positive bacteria and mycobacteria. We propose that Lysin_{241} acts at a later stage of infection to facilitate lysis initiated by the secreted and activated Lysin_{384}, compromising the peptidoglycan linkages and allowing completion of mycobacterial host cell lysis.

Introduction

At the end of the replication cycle, bacteriophages must exit the host cell and disperse their newly formed progeny to infect new prey cells. The main barrier to host lysis is the continuous meshwork of peptidoglycan, a strong, stable structure that allows the bacterial envelope to withstand internal osmotic pressure (Young *et al*., 2000): compromising the cell wall is thus, the fundamental goal for lytic processes. With the exception of filamentous phages that as a result of their unique morphology and morphogenesis can extrude through the envelope without fatal consequences for the host, all other phages must either degrade or otherwise compromise the peptidoglycan to cause lysis (Young, 2002; São-José *et al*., 2007). Most of the tailed double-stranded DNA (dsDNA) phages achieve the proper time for lysis by the consecutive use of two lysis proteins – holin and endolysin. Holins are small hydrophobic membrane proteins which during the late phase of phage development, progressively accumulate in the cytoplasmic membrane of the host and while the proton-motive force is maintained assemble into oligomers and rafts of intrinsic stability (Gründling *et al*., 2001; Young, 2002). At a precise time programmed into its primary structure and upon a specific trigger event (critical holin effector concentration and partial depolarization of the membrane), the holin suddenly causes disruption of the membrane with non-specific hole formation and collapse of the membrane potential which sets the time of lysis by allowing the destruction of the cell wall by the released or activated phage encoded muralytic enzymes, the endolysins (Wang *et al*., 2000; Young and Wang, 2006). The term endolysin is used to describe the dsDNA bacteriophage-encoded peptidoglycan hydrolases, which are synthesized in phage-infected cells at the end of the multiplication cycle. They are characterized by their ability to directly target bonds in the peptidoglycan (PG) layer of the bacterial cell wall; the result of this activity is degradation of the rigid murein layer and release of newly assembled virions by

way of lysis (Loessner, 2005; Borysowski *et al*., 2006). Endolysins distribute between five major functional types; i) *N*-acetylmuramidases (lysozymes); ii) endo-β-*N*-acetylglucosaminidases, which all hydrolyze the β-1, 4 glycosidic bonds in the murein; iii) transglycosylases, which attack the same bonds but form a cyclic 1, 6 anhydro-*N*-acetylmuramic acid product; iv) *N*-acetyl-muramoyl-L-alanine amidases, which hydrolyze the amide bond between the *N*-acetylmuramic acid and L-alanine residues in the oligopeptide crosslinking chains; v) endopeptidases, which attack the peptide bonds in the same chains. In phages infecting Gram-positive bacteria, the most common architecture includes two clearly separated functional domains: the N-terminal domain that generally harbours the enzymatic activity, whereas the cell wall binding domain (CWBD) located at the C-terminal region directs the enzymes to their substrates and may restrain the enzyme lytic action to a particular type of cell wall (Loessner, 2005; São-José *et al*., 2007). Nonetheless, at least four bifunctional lysins have also been reported, consisting of an N-terminal and central catalytic domain with different specificity and a C-terminal substrate-binding domain; examples are the endolysins encoded by *Streptococcus agalactiae* bacteriophage B30 (muramidase and endopeptidase) (Pritchard *et al*., 2004), *Staphylococcus aureus* phage ϕ11 (endopeptidase and amidase) (Navarre *et al*., 1999), *S. agalactiae* phage NCTC 11261 (endopeptidase and muramidase) (Cheng *et al*., 2005), and *Staphylococcus warneri* M phage ϕWMY (endopeptidase and amidase) (Yokoi *et al*., 2005). Most known endolysins lack signal peptides sequences and depend entirely on the cognate holins for release to the peptidoglycan. However, several endolysins containing N-terminal secretory signals have already been described. The endolysin of *Oenococcus oeni* phage fOg44 is endowed with a bona-fide signal peptide (SP) that is processed by the leader peptidase during infection, presumptive evidence that the endolysin is exported by the *sec* machinery (São-José *et al*., 2000). A survey of orthologous endolysins from other phages of Gram-positive bacteria suggested that some, but

not all of these endolysins have N-terminal sequences resembling secretory signals, although in every case an adjacent holin gene was also present (São-José *et al*., 2000; Young and Wang, 2006). In addition, the endolysin of *Lactobacillus plantarum* phage ϕg1e had also been reported to be processed in *E. coli* by the Sec machinery (Kakikawa *et al*., 2002). Particularly remarkable cases are the endolysins of *E. coli* phages P1 and 21, which feature an N-terminal signal-arrest-release (SAR) sequence that allows the enzyme to be exported to the membrane where it is arrested, and to be released as a soluble active enzyme in the periplasm (Xu *et al*., 2004; Xu *et al*., 2005; Sun *et al.*, 2009).

The mycobacteriophage Ms6 is a temperate bacteriophage with an unusual lytic cassette: in addition to the endolysin-holin lysis system, encoded by genes *lysA* (*gp2*) and *hol* (*gp4*), respectively, the Ms6 lytic cassette comprises three additional lysis proteins encoded by genes *gp1*, *gp3* (*lysB*) and *gp5* (Garcia *et al*., 2002). The product of *lysB* gene has been previously characterized and encodes an enzyme with lipolytic activity that hydrolyzes the mycolic acids from the mycolyl-arabinogalactan-peptidoglycan (Gil *et al.*, 2008; Gil *et al*., 2010) acting at a later stage of infection to facilitate lysis by compromising the integrity of the mycobacterial outer membrane linkage to the arabinogalactan-peptidoglycan layer (Payne *et al*., 2009). The Ms6 *lysA* gene was shown to encode a 384 amino acid polypeptide (LysA) with significant similarity to lysins with an *N*-acetylmuramoyl-L-alanine amidase activity of some bacteriophages (Garcia *et al*., 2002). Several types of cell wall hydrolases seem to be produced by mycobacteriophages. While phages Ms6 and TM4 encode enzymes with an "Amid-2" type domain (São-José *et al*., 2003), others such as D29 and Bxb1 employ hydrolases with lysozyme-like activity to bring about host cell lysis. Mycobacteriophage endolysins containing SAR domains that allow secretion of the endolysin into the periplasmic space have not yet been described (Payne *et al*., 2009). Interestingly, however, our group has recently identified the product of *gp1* gene as a chaperone-like protein that

specifically interacts with the N-terminal region of the endolysin, is involved in its translocation across the cytoplasmic membrane independently of the holin function and is required for efficient phage release (Catalão *et al.*, 2010). During an attempt to purify LysA as a C-terminal histidine-tagged fusion product, we detected the synthesis of two proteins, rather than a single polypeptide. Further examination of *lysA* ($lysin_{384}$) DNA sequence revealed a second possible gene ($lysin_{241}$) in the same reading frame and preceded by a potential ribosome-binding site (RBS). Here, we report studies directed at dissecting the precise role of the *lysA*-encoded gene products during *M. smegmatis* infection by the mycobacteriophage Ms6. In addition, the lytic activity spectrum of both proteins was also examined in both Gram-positive and Gram-negative bacteria and also in mycobacteria regarding the potential therapeutic use of mycobacteriophage lysins.

Experimental Procedures

Bacterial strains, phages, plasmids and culture conditions

Bacterial strains, phages and plasmids used throughout this study are listed in Table 1. *E. coli* strains were grown at 37 ºC, in Luria-Bertani (LB) broth or agar supplemented with 100 µg ml^{-1} ampicillin or 30 µg ml^{-1} kanamycin, when appropriate. *M. smegmatis* recombinant strains were grown at 37 ºC in 7H9 medium (Difco) supplemented with 0.05 % Tween 80, with shaking or Middlebrook 7H10 (Difco), containing 15 µg ml^{-1} kanamycin. For induced conditions 0.2 % succinate and 0.2% acetamide were also added to media.

Plasmid construction

Unless otherwise indicated, DNA fragments obtained by PCR were amplified using Ms6 genomic DNA as template. DNA amplification, plasmid isolation and electrophoresis were carried out using standard techniques (Sambrook and Russell, 2001). *E. coli* and *M. smegmatis* mc^2155 cells were transformed as described previously (Snapper *et al*., 1990; Sambrook and Russell, 2001). Restriction enzymes and T4 DNA ligase (New England Biolabs) were used according to the supplier's recommendations. All oligonucleotides were purchased from Thermo Scientific and are listed in Table 1.

Table 1. Strains, bacteriophages and plasmids used in this study.

Strain, bacteriophage, or plasmid	Description	Reference or Source
Bacteria		
Escherichia coli		
JM109	*recA1 endA1 gyr96 thi hsdR17 supE44 relA1* Δ(*lac-proAB*) [F′ *traD36 proAB lacI^q^*ZΔM15]	Strategene
BL21 (DE3)	F^- *ompT* $hsdS_B$ (r_B^- m_B^-) *gal dmc* (DE3)	Novagen
Mycobacterium smegmatis mc^2155	High-transformation-efficiency mutant of *M. smegmatis* ATCC 607	Snapper *et al*., 1990
Bacteriophages		
$Ms6_{wt}$	Temperate bacteriophage from *M. smegmatis*	Portugal *et al*., 1989
Ms6-$LysAHis_6$	His_6tag insertion at the 3' end of Ms6 *lysA*	Catalão *et al*., 2010
$Ms6_{\Delta gp1}$-$LysAHis_6$	His_6tag insertion at the 3' end of Ms6 *lysA* in $Ms6_{\Delta gp1}$	This study
Ms6-$LysAHis_{6TGA}$	Stop codon introduced in the *lysA* gene of Ms6-$LysAHis_6$	This study
Ms6-$LysAHis_{6GTG\rightarrow TGG}$	GTG→TGG change in codon 143 of *lysA* in Ms6-$LysAHis_6$	This study
Plasmids		
pQE30	Expression vector; T5 promoter; Amp^r	QIAGEN
pET29b(+)	Expression vector, T7 promoter; Kan^r	Novagen
pJV53	Derivative of pLAM12 with Che9c *60* and *61* under control of the acetamidase promoter; Kan^r	van Kessel and Hatfull, 2007
pMG231A	$lysin_{384}$ Ms6 cloned into pQE30	Garcia *et al*., 2002
pMJC40	$lysin_{241}$ Ms6 cloned into pQE30	This study
pMJC41	$lysin_{384}$ Ms6 cloned into pET29b(+)	This study
pMJC42	$lysin_{241}$ Ms6 cloned into pET29b(+)	This study
pMJC43	$lysin_{384\Delta GTG}$ cloned into pET29b(+)ΔRBS	This study
Oligonucleotides	**Sequence[a] 5′-3′**	
$Lysin_{241}$fwd-1	CAGGATCCCCCGACGAACCACGCCCC	Ms6 $lysin_{241}^{CP}$ pQE30
$Lysin_{241}$fwd-2	CAGGATCCCGACGAACCACGCCCC	Ms6 $lysin_{241}^{CP}$ pET29b

gp2A	CGCGGATCCCACCACGAAAGATCAAG	Ms6 *lysin*$_{384}$[CP] pET29b
pORF2-c1	ATGCGAAGCTTCAGTGGCCCAACAGTTC	Ms6 *lysin*$_{241}$[CP]
p*lysA*ΔGTGfwd	GCTCTAGACCACGAAAGATCAAGTCGC	Ms6 *lysin*$_{384\Delta GTG}$[CP] pET29b
PrΔ*lysA*	CTGCGCTCCATCCCCGTCCTCGGCGGAATCCTCG GGAGCAAACGGTGACTGGCCGCAGCTCGGCGGA AAAACCCTCGTGGACGCGGTAGCAGAACTGTT	Ms6 *lysA* 1089bpΔ
PrExtenderΔ*lysA*fwd	CGACCTGACCAACCTTCCAGCGCAAGTCATGGAC ATCATCGACAGCGCGCTGCGCTCCATCCCCGTCC TCGGCG	Extend PrΔ*lysA*
PrExtenderΔ*lysA*rv	CCCCTGGTCCGAGGCCGACGTATTGGCCGTCGAT GCGCATCAGTGGCCCAACAGTTCTGCTACCGCGT CCACGA	Extend PrΔ*lysA*
Pr*lysA*$_{TGA}$HindIIIfwd	CCACGAAAGATCAAGTCGCCCAAATCACCATCGC CTGAAGCTTCAAGGCGCGCGGCTACACCCGCAGC GAATG	Ms6 *lysA*$_{TGA}$
Pr*lysA*$_{TGA}$HindIIIrv	CATTCGCTGCGGGTGTAGCCGCGCGCCTTGAAGC TTCAGGCGATGGTGATTTGGGCGACTTGATCTTTC GTGG	Ms6 *lysA*$_{TGA}$
Pr*lysA*$_{GTG\rightarrow TGG}$MscIfwd	ATACCTCGACAAGTACTGGCCCGCCGATGGAGGT ACCGCCTGGCCAGACGAACCACGCCCCGACTTCA ACGAGTTTCCGATCTGGT	Ms6 *lysA*$_{GTG\rightarrow TGG}$
Pr*lysA*$_{GTG\rightarrow TGG}$MscIrv	ACCAGATCGGAAACTCGTTGAAGTCGGGGCGTGG TTCGTCTGGCCAGGCGGTACCTCCATCGGCGGGCC AGTACTTGTCGAGGTAT	Ms6 *lysA*$_{GTG\rightarrow TGG}$
Pr*lysA*His$_6$tag	AACCCTCGTGGACGCGGTAGCAGAACTGTTGGG CCACCACCACCACCACCACTGATGCGCATCGAC GGCCAATACGTCGGCCTCGGACC	Ms6 *lysA*His$_6$tag insertion
PrExtender*lysA*His$_6$tagfwd	CGAGATCCTGCGGCAACTGCGCGGATACAACCT CACTGGCTGGCCGCAGCTCGGCGGAAAAACCCT CGTGGACGCGGTAGCAGAACTGTT	ExtendPr*lysA*His$_6$tag
PrExtender*lysA*His$_6$tagrv	TAGGAGAACTTGCGCCGCATGAACGCTTTGATC TTGCGGATCTCGTCGGATCTGTCCCCTGGTCCGA GGCCGACGTATTGGCCGTCGATG	ExtendPr*lysA*His$_6$tag
*lysA*His$_6$tagfwd	GAACTGTTGGGCCACCACCACCACCAC	Ms6 *lysA*His$_6$tag[TAG]
*lysB*rv	GATACCCCATGACGTACTTGGTTTCGGCG	Ms6 *lysA*His$_6$tag [FP]
PrP1fwd	CGGTACTAGTCGGCCTCGGCCTGC	Ms6 *lysA*$_{TGA}$ [FP]
Pr*lysA*180bprv	GCAAGCTTGTGTGGGTAGGAGCCGTCC	Ms6 *lysA*$_{TGA}$ [FP]
gp2B	CGATGCTGCAGTCAGTGGCCCAACAC	Ms6 *lysA*$_{GTG\rightarrow TGG}$ [FP]
Δ*lysA*DADA-PCR	GAATCCTCGGGAGCAAACGGTGACCTGG	Ms6 Δ*lysA*[DP]
*lysB*DADA-PCR	GATACCCCATGAACGTACTTGGTTTCGGCG	Ms6 Δ*lysA*[DP]

[a] underlined bases were added to provide additional restriction sites
[CP] Cloning Primer
[FP] Flanking primer.
[DP] DADA-PCR primer
[TAG] Primer specific to the tag sequence.
Ms6 lysis genes Accession No. AF319619

In order to construct plasmids pMJC40 and pMJC42, *lysin*$_{241}$ was amplified using primers Lysin$_{241}$fwd-1/pORF2-c1 or Lysin$_{241}$fwd-2/pORF2-c1 and the resulting DNA fragments were introduced into the BamHI/HindIII sites of vector pQE30, allowing a fusion to a hexahistidine tag at the N-terminus, or pET29b

allowing a C-terminal hexahistidine tag fusion, respectively. To obtain plasmid pMJC41, the DNA fragment containing *lysin*$_{384}$ (*lysA*), was amplified by PCR with primers gp2A/pORF2-c1 and cloned into BamHI/HindIII sites of pET29b fused to a C-terminal His$_6$tag. Plasmid pMJC43 was constructed by PCR amplifying *lysin*$_{384}$ (*lysA*) lacking the GTG start codon with primers p*lysA*ΔGTGfwd/pORF2-c1 and cloning the DNA fragment, after restriction with XbaI/HindIII which removed the vector translational signals (RBS and start codons), in pET29b. All constructs were validated by sequencing the insert nucleotide sequence.

Construction of Ms6 mutant phages

Construction of Ms6 mutant phages was performed using Bacteriophage Recombineering of Electroporated DNA (BRED) in *M. smegmatis*. Recombineering substrates and BRED strategy were done as described previously (Marinelli *et al*., 2008; Catalão *et al*., 2010). Briefly, for Ms6 *lysA* deletion, a 100 bp oligonucleotide, PrΔ*lysA*, that has 50 bp of homology upstream and downstream of the region to be deleted was extended by PCR using two 75 bp extender primers, PrExtΔ*lysA*fwd/ PrExtΔ*lysA*rv, which have 25 bp of homology to the ends of the 100-mer and add additional 50 bp of homology on either end. The final 200 bp dsDNA product was purified using MinElute PCR Purification Kit (QIAGEN) and co-electroporated with Ms6-LysAHis$_6$ genomic DNA into electrocompetent recombineering cells of *M. smegmatis* mc^{2}155:pJV53. To abolish synthesis of Lysin$_{384}$ and Lysin$_{241}$ we designed two complementary 73 bp oligonucleotides (Pr*lysA*$_{TGA}$HindIIIfwd/ Pr*lysA*$_{TGA}$HindIIIrv) that introduce a stop codon and a HindIII restriction site downstream of the start codon of *lysin$_{384}$*, or two complementary 86 bp oligonucleotides (Pr*lysA*$_{GTG \rightarrow TGG}$MscIfwd/ Pr*lysA*$_{GTG \rightarrow TGG}$MscIrv) that modify the *lysin$_{241}$* GTG start codon (valine) to TGG (tryptophan) and introduce an MscI restriction site, respectively. Complementary oligonucleotides were co-transformed with Ms6-LysAHis$_6$ genomic DNA into

recombineering cells of *M. smegmatis* mc^{2}155:pJV53. Cells were resuspended in 7H9 supplemented with 0.5% glucose and 1 mM of $CaCl_2$, incubated at 37 ºC for 2 hours (prior to lysis) and plated as top agar lawns with *M. smegmatis* mc^{2}155. Phage plaques were picked into 100 µl phage buffer (10 mM Tris-HCl, pH 7.5; 10 mM $MgSO_4$; 68.5 mM NaCl; 1 mM $CaCl_2$), eluted for two hours at room temperature and analysed by DADA-PCR (Marinelli *et al*., 2008) with primers Δ*lysA*DADA-PCR/*lysB*DADA-PCR to detect *lysA* deletion. To detect *lysA* TGA insertion and *lysA* GTG_{143}→TGG change, phage plaques were analysed by PCR with primers PrP1fwd/*lysA*180bprv or Gp2A/Gp2B followed by HindIII or MscI restriction, respectively. Mixed primary plaques containing both mutant and wild-type DNA were eluted as described above, and serial dilutions were plated with *M. smegmatis.* Individual secondary plaques or lysates were screened by DADA-PCR for *lysA* deletion or by PCR and restriction with the same primers referred above to identify pure mutant phages. Construction of phage $Ms6_{\Delta gp1}$-$LysAHis_6$ was done as described previously (Catalão *et al*., 2010) using $Ms6_{\Delta gp1}$ genomic DNA.

Lysin expression in M. smegmatis-infected cells

Examination of $Lysin_{384}$ and $Lysin_{241}$ synthesis in *M. smegmatis* was performed as previously described (São-José *et al*., 2000). An exponentially growing culture of *M. smegmatis* mc^{2}155 was infected with Ms6-$LysAHis_6$, $Ms6_{\Delta gp1}$-$LysAHis_6$, Ms6-$LysAHis_{6TGA}$ or Ms6-$LysAHis_{6GTG\rightarrow TGG}$ at an approximate multiplicity of infection (m.o.i) of 10 and incubated at 37 ºC for 30 minutes. Ten-mililiter samples were withdrawn at 30-min intervals, cells were pelleted by centrifugation and frozen at -20 ºC. After thawing, cells were concentrated 100-fold in phosphate-buffered saline (PBS) supplemented with 20 mg of lysozyme ml^{-1}. After an incubation period at 37 ºC for 1 hour, 25 µl of 5× SDS-PAGE sample buffer were added followed by incubation at 100 ºC for 5 minutes to complete cell

lysis. *M. smegmatis* extracts were analysed by Western-blotting and lysin immunodetection was performed using horseradish-peroxidase-conjugated anti-His monoclonal antibody (Roche).

One-step growth curves

One-step growth curve and burst-size determination were previously described (Catalão *et al*., 2010). The one step assays were carried out in *M. smegmatis* exponential growth cells using an m.o.i. of 1. Cells were pelleted and resuspended in 1 ml of a phage suspension ($Ms6_{wt}$, Ms6-$LysAHis_{6TGA}$ or Ms6-$LysAHis_{6GTG \rightarrow TGG}$) supplemented with 1 mM $CaCl_2$. The mixture was incubated 50 min at 37 ºC to allow adsorption of the phages. 100 µl of 0.4% H_2SO_4 was added to inactivate the non-adsorbed phages and the incubation continued for five min. The suspension was neutralized with 100 µl of 0.4% NaOH and diluted 1:100 in 7H9 supplemented with 0.5% glucose and 1 mM $CaCl_2$. 1 ml samples were withdrawn every 30 min until reaching 300 min. 100 µl of serial dilutions of each sample were plated with 200 µl of *M. smegmatis* cells, on 7H10 as top agar lawns and the phage titer for each sample was determined after 24h incubation at 37 ºC. Results are averages of three independent experiments.

Expression of $Lysin_{384}$ and $Lysin_{241}$ proteins in E. coli

E. coli BL21 (DE3):pMJC41 or *E. coli* BL21 (DE3):pMJC42 were grown in LB medium to an OD_{600nm} of 0.6, and expression of the recombinant $Lysin_{384}$-His_6 or $Lysin_{241}$-His_6 was induced for 4 h following the addition of IPTG to a final concentration of 1 mM. Bacterial cells were harvested by centrifugation, washed, resuspended in 50 mM Tris-HCl (pH 7.5) supplemented with a cocktail of protease inhibitors (Calbiochem), and disrupted by passage through a French pressure cell. Cell debris were removed by centrifugation, and the recombinant proteins present

in the supernatant were analysed by SDS-PAGE, followed by Coomassie blue staining and Western-blotting and detected as described above.

Assay of the antibacterial activity of Lysin$_{384}$ and Lysin$_{241}$

The antibacterial activity was screened using a sensitivity test (Davidson *et al*., 2005; Son *et al.*, 2009) with some modifications. 100 µl of an exponential growing culture of the test strain was plated on LB or 7H10 + OADC (for mycobacteria) as top agar lawns. 20 µl of induced *E. coli*:pMJC41 or *E. coli*:pMJC42 extracts containing Lysin$_{384}$ or Lysin$_{241}$ were spotted onto the bacterial lawn of the test strain and incubated overnight at 37 ºC. After overnight incubation, the presence of a clear zone was examined. *E. coli*:pET29b induced extract was used as a negative control. Activity assays were performed in triplicate. Several bacterial strains were used to test the range of antibacterial activity and were obtained from the American Type Culture Collection (ATCC) or from the Institute Pasteur Collection, Paris.

Zymogram analysis

Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis was performed as described by Laemmli (1970), and the zymogram assays were carried out as outlined by Piuri and Hatfull (2006). Briefly, 0.2% autoclaved and lyophilized *Micrococcus luteus* cells were included in 15% polyacrylamide gels for detection of bacteriolytic activity. Protein samples were heated for 3 min at 100 °C in sample buffer (62.5 mM Tris-HCl, pH 6.8; 2% SDS; 5% mercaptoethanol; 20% glycerol; 0.01% bromophenol blue), and then separated on SDS-gels containing the autoclaved cells. After electrophoresis, the zymograms were washed for 30 min with distilled water at room temperature and then transferred into renaturation buffer containing 25 mM Tris-HCl (pH 7.5) and 0.1% Triton X-100 followed by

further incubation for 16 h at 37 °C. The zymograms were rinsed with distilled water, stained with 0.1% methylene blue in 0.01% KOH for 2 h at room temperature, and then destained with distilled water. Peptidoglycan hydrolase activity was detected as a clear zone on a dark blue background of stained peptidoglycan. Gels not containing peptidoglycan were stained with Coomassie Brilliant Blue. Lysozyme and bovine serum albumin (BSA) were used as positive and negative controls, respectively. Molecular masses were determined by comparison with prestained molecular weight standards that were electrophoresed on the same gel.

Results

Identification of two overlapping gene products from lysA

The 1152 bp *lysA* gene of mycobacteriophage Ms6 starts at a GTG codon that overlaps the *gp1* TGA stop codon in a different reading frame, and is preceded in four nucleotides by a ribosome-binding site (RBS) sequence (5′-GGGAGCA-3′). It encodes a 384 amino acid polypeptide with significant similarity to mureinolytic enzymes (lysins) of several bacteriophages with *N*-acetyl-muramoyl-L-alanine amidase activity (Garcia *et al*., 2002). During an attempt to purify LysA as a C-terminal histidine-tagged fusion product (LysA-His_6), we detected the production of two proteins of 27 kDa ($Lysin_{241}$) and 43 kDa ($Lysin_{384}$), rather than a single polypeptide of 43 kDa, in *E. coli* extracts (Fig. 1A). We also observed that when the protein was tagged in the N-terminal domain, only the larger product ($Lysin_{384}$) reacted with the anti-His_6 antibody in addition to an unexpected 16 kDa product (Fig. 1D). The fact that a time-dependent decrease of the 43 kDa form with a concomitant increase of the labelled 27 kDa form was not observed, but both proteins seemed to be produced independently over the time (Fig. 1E), led us to consider that the larger and smaller proteins are not a precursor and an N-terminal

processed form respectively, but rather two independent translated products. Also supporting this notion, analysis of the amino acid sequence of Ms6 LysA did not predict an amino-terminal signal sequence or a peptidase cleavage site (Catalão *et al.*, 2010).

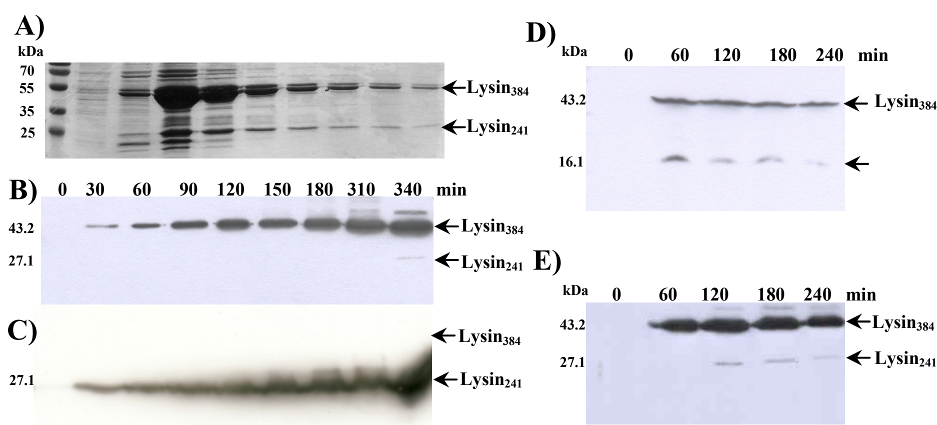


Figure 1. LysA expression in *E. coli*. A) Purified LysA-His_6 fractions after SDS-PAGE analysis and Coomassie Blue staining. LysA-His_6 was produced from pMJC41 in *E. coli* BL21 (DE3) after IPTG induction. Two polypeptides were observed after LysA purification with 43.2 kDa ($Lysin_{384}$-His_6) and 27.1 kDa ($Lysin_{241}$-His_6). **B)** LysA-His_6 synthesis from pMJC41 over the time is not always followed by $Lysin_{241}$-His_6 production. **C)** Synthesis of $Lysin_{241}$-His_6 from pMJC43. Removal of pET29b and LysA ($Lysin_{384}$-His_6) translational signals does not hinder $Lysin_{241}$-His_6 synthesis. **D)** Expression of His_6-LysA from pMG231A. **E)** Expression of LysA- His_6 from pMJC41: detection of C-terminal His_6tag LysA shows the production of $Lysin_{384}$-His_6 and $Lysin_{241}$-His_6. The molecular masses in kDa of $Lysin_{384}$ and $Lysin_{241}$ are indicated on the left; positions of both proteins are indicated by an arrow on the right. $Lysin_{384}$ and $Lysin_{241}$ were detected by Western-blotting with an anti-His_6 antibody, except for panel A.

Further examination of the *lysA* nucleotide sequence revealed a second potential overlapped gene in the same reading frame, possessing putative translational signals, a consensus Shine-Dalgarno sequence (5´-TGGAGGT-3´) utilized by Gram-positive bacteria including mycobacteria (Dale and Patki, 1990)

```
          Gp1
      →
     catcatcgacagcgcgctgcgctccatccccgtcctcggcggaatcctcgggagcaaacg

      →  Lysin384 (LysA)                      tgaaagctt (Ms6-LysAHis6TGA)
     gtg
   1    accacgaaagatcaagtcgcccaaatcaccatcgccgaagccaaggcgcgcggctac
      V  T  T  K  D  Q  V  A  Q  I  T  I  A  E  A  K  A  R  G  Y
  61 acccgcagcgaatgcctggcgatcatgtccaccttctaccaagagtccggctggaacgac
      T  R  S  E  C  L  A  I  M  S  T  F  Y  Q  E  S  G  W  N  D
 121 accatctgggacccgacccacaccacctacggcattgcccagcaggacggctcctaccca
      T  I  W  D  P  T  H  T  T  Y  G  I  A  Q  Q  D  G  S  Y  P
 181 caccgcttcgacggtgccgcagcccaaatcaaaggcttcttcgacaagctcgacgtgtgg
      H  R  F  D  G  A  A  A  Q  I  K  G  F  F  D  K  L  D  V  W
 241 cgcgccaaacccggtgccagcaccgatatatggctgaacatctgctggatgcagcaggcc
      R  A  K  P  G  A  S  T  D  I  W  L  N  I  C  W  M  Q  Q  A
 301 cccaactggcccagcgctgactactggtacgccaacggccgccgcgcctacctcaccgaa
      P  N  W  P  S  A  D  Y  W  Y  A  N  G  R  R  A  Y  L  T  E
 361 atcaagtcacgcatcgccaccgtcaccccatacctcgacaagtactggcccgccgatgga
      I  K  S  R  I  A  T  V  T  P  Y  L  D  K  Y  W  P  A  D  G

   →  Lysin241  tgg  cca (Ms6LysAHis6GTG→TGG)
                 gtg
 421 ggtaccgcc     cccgacgaaccacgccccgacttcaacgagtttccgatctggtcgaac
      G  T  A  V  P  D  E  P  R  P  D  F  N  E  F  P  I  W  S  N
 481 aacaacagcgcccgcagcggcaagcccaccatgttcctgatccacacccaagaaggcggc
      N  N  S  A  R  S  G  K  P  T  M  F  L  I  H  T  Q  E  G  G
 541 ggcggggacgctgccgccgagaacctggcgaagtggttccagaacggcaacggcgtctcg
      G  G  D  A  A  A  E  N  L  A  K  W  F  Q  N  G  N  G  V  S
 601 taccactacacgatctcccaagcgtccgatggtggtgtgacggtggtcgattgcgtcgac
      Y  H  Y  T  I  S  Q  A  S  D  G  G  V  T  V  V  D  C  V  D
 661 actgaccgcgccgcctggtctgtgggcaacgccaacagcatcagcatcaacctgtgcttc
      T  D  R  A  A  W  S  V  G  N  A  N  S  I  S  I  N  L  C  F
 721 gcggggtcgcgagcatcctggatgcgggatcagtggatgaagcagtccaacgcaatcgac
      A  G  S  R  A  S  W  M  R  D  Q  W  M  K  Q  S  N  A  I  D
 781 gtcgcagcctacctcgcggtgcaggacgcgaagaagtacggcttcaccccgctcgtggtg
      V  A  A  Y  L  A  V  Q  D  A  K  K  Y  G  F  T  P  L  V  V
 841 ccaccgccgtatacgaatgggcgacctggcatctcggaccaccggtgggtgaccgacgtg
      P  P  P  Y  T  N  G  R  P  G  I  S  D  H  R  W  V  T  D  V
 901 ttcaagtggggcactcacaccgacgtcggagactggttcccgtgggactacttcgccgaa
      F  K  W  G  T  H  T  D  V  G  D  W  F  P  W  D  Y  F  A  E
 961 cgggtcaaccactgggccaacggtggcaagaccgagcctgaaccgcccaaggtgaaacgc
      R  V  N  H  W  A  N  G  G  K  T  E  P  E  P  P  K  V  K  R
1021 ttcccggacgactggagtgaccgcgaaatcctcgtcgagatcctgcggcaactgcgcgga
      F  P  D  D  W  S  D  R  E  I  L  V  E  I  L  R  Q  L  R  G
1081 tacaacctcactggctggccgcagctcggcggaaaaaccctcgtggacgcggtagcagaa
      Y  N  L  T  G  W  P  Q  L  G  G  K  T  L  V  D  A  V  A  E
1141 ctgttgggccac     tga
      L  L  G  H
                 └──→ 5 histidines insertion (Ms6-LysAHis6 and Ms6Δgp1-LysAHis6)
```

Figure 2. Relevant features of the DNA sequence including and surrounding the Ms6 *lysA* gene. Putative RBS consensus sequences from $lysin_{384}$ and $lysin_{241}$ are shown in bold and underlined. Translational start and stop codons are superscripted and/or in bold. Amino acids residues of LysA are indicated below the nucleotide sequence. Amino acid substitutions and insertions to construct Ms6 *lysA* mutant phages are highlighted; Ms6-$LysAHis_{6TGA}$ has a stop codon and a HindIII restriction site, downstream of the *lysA* start codon which eliminates synthesis of $Lysin_{384;}$ substitution of the GTG codon by TGG at position 143 eliminates synthesis of $Lysin_{241}$ (Ms6-$LysAHis_{6GTG \rightarrow TGG}$); Ms6-$LysAHis_6$ (Catalão *et al*., 2010) and $Ms6_{\Delta gp1}$-$LysAHis_6$ have a five histidine insertion just before the TGA stop codon to generate a His_6tag C-terminal fusion with LysA.

and the associated translational start codon GTG, 6 nucleotides downstream (Fig. 2). Occurrence of an additional translation event at the predicted location would be compatible with the difference in apparent molecular masses (43 and 27 kDa) of the two proteins considering that an additional His_6tag was C-terminally fused to LysA to produce LysA-His_6. This was experimentally confirmed by sequencing the amino-terminal region of the 27 kDa protein ($Lysin_{241}$). This protein was obtained by expressing LysA ($Lysin_{384}$) in *E. coli* and isolating the smaller lysin from a polyvinylidene difluoride (PVDF) membrane. Indeed, the obtained N-terminal sequence, MPDEPRPD, matched the deduced sequence from residues 143 to 150. Of note is the fact that the larger product ($Lysin_{384}$) is in great excess when compared to the smaller protein ($Lysin_{241}$) and frequently $Lysin_{241}$ synthesis was not observed (Fig. 1B) when expressed in *E. coli*. This is not unexpected since synthesis of $Lysin_{384}$ is dependent on the expression vector pET29 transcription and translational signals and also indicates that the predicted RBS of $Lysin_{241}$ is recognized by *E. coli*. To further clarify these results, $Lysin_{384}$ lacking its own GTG start codon was cloned into the XbaI/HindIII sites of pET29 which removes the Shine-Dalgarno sequence and the start codons of the expression vector. We were expecting that if $Lysin_{241}$ results in fact from a new translation event and is not the

result of cleavage of $Lysin_{384}$, it should be synthesized from this construction. Western-blotting analysis revealed the production of a single polypeptide with 27 kDa corresponding to $Lysin_{241}$ (Fig 1C). This result unambiguously proves that $Lysin_{241}$ results from a new translation event and is independent of $Lysin_{384}$ synthesis.

To follow mycobacteriophage Ms6 LysA production in the course of *M. smegmatis* infection, infected cells were examined for lysin synthesis. To allow detection of the protein, mycobacteria were infected with the Ms6 derivative mutant, Ms6-$LysAHis_6$ phage (at an m.o.i. of 10) where the C-terminal 3´ end of *lysA* gene was fused to a sequence coding for a hexahistidine tag, allowing the production of a LysA-His_6 tagged protein (Catalão *et al.*, 2010). Samples were collected immediately before and every 30 minutes following infection until near the end of the Ms6 infection cycle. Protein extracts were prepared from such samples as described in Material and Methods and checked for the presence of LysA by immunoblotting. Western-blotting revealed the presence of two proteins rather than a single lysin band, with 27 kDa and 43 kDa corresponding to $Lysin_{241}$ and $Lysin_{384}$, respectively (Fig. 3A). Both proteins were first detected at 90 minutes postinfection with mobility indistinguishable from that exhibited by the protein forms of the *E. coli* expressed lysin. In addition, synthesis of both proteins was also followed during *M. smegmatis* infection with the mutant phage $Ms6_{\Delta gp1}$ that has a reduced burst-size (Catalão *et al.*, 2010). When we analysed the endolysin production during the infective cycle of the *gp1* deleted phage, we observed a decrease in $Lysin_{384}$ synthesis/stability although $Lysin_{241}$ amount was comparable to the levels detected for the wild-type phage (Fig. 3B). Similarly to what was observed during Ms6 infection of *M. smegmatis*, both endolysins were detected 90 minutes postinfection (Fig. 3A).

These results suggest that during *M. smegmatis* Ms6 infection, different products of the *lysA* gene are synthesized and result from the existence of two

translational elements that direct the production of a smaller ($Lysin_{241}$) and a larger ($Lysin_{384}$) endolysin rather than a processing event. In addition levels of synthesis and/or stability of $Lysin_{384}$ seem to be dependent on the Gp1 chaperone-like protein as already proposed (Catalão *et al.*, 2010).

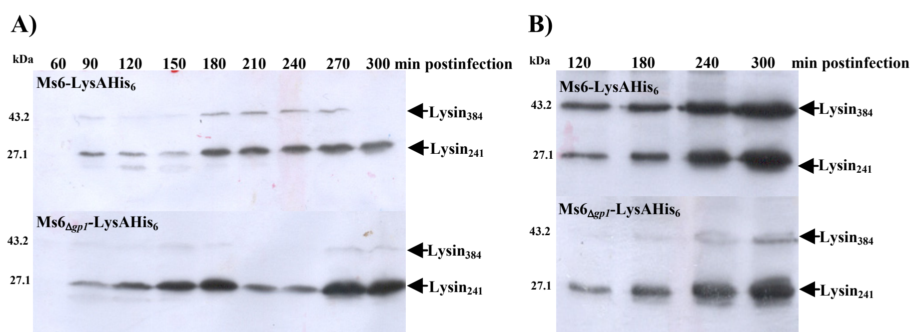


Figure 3. Time course of $Lysin_{384}$ and $Lysin_{241}$ synthesis during Ms6 infection of *M. smegmatis*. Lysin production in *M. smegmatis* was analysed after infection with Ms6-$LysAHis_6$ or $Ms6_{\Delta gp1}$-$LysAHis_6$ at an m.o.i of 10. Extracts were prepared from samples taken at 30-min intervals as described in Material and Methods. Samples were analysed by Western-blotting and $Lysin_{384}$ and $Lysin_{241}$ synthesis was detected with an anti-His_6 monoclonal antibody. **A)** $Lysin_{384}$ and $Lysin_{241}$ synthesis is first detected 90 minutes postinfection both in Ms6-$LysAHis_6$ (upper panel) and $Ms6_{\Delta gp1}$-$LysAHis_6$ (lower panel) mutant phages. Only the results for 60 to 300 min postinfection are shown. **B)** $Lysin_{384}$ is synthesized to near undetectable levels during $Ms6_{\Delta gp1}$-$LysAHis_6$ infection of *M. smegmatis* (lower panel) whereas $Lysin_{241}$ production is comparable to the wild-type phage. The molecular masses in kDa of $Lysin_{384}$ and $Lysin_{241}$ are indicated on the left; positions of both proteins are indicated by an arrow on the right.

Occurrence of Ms6 Lysin A-like proteins in mycobacteriophages

All of the mycobacteriophage genomes sequenced so far possess a putative *lysA*-like gene and they have been recently grouped in mycobacteriophage gene

family Pham66-1 (Payne *et al*., 2009; Hatfull *et al*., 2010), although this group of enzymes is not restricted to phages that infect mycobacteria (Hatfull *et al*., 2010). The Lysin A (LysA) family of proteins appears to be a particularly highly diverse and interesting group of lytic enzymes composed of subgenic modules with reasonably defined boundaries (Hatfull *et al*., 2006) and containing amidase, glycosidase or peptidase motifs (Payne *et al*., 2009). The extensive sequence divergence between mycobacteriophage-encoded and other endolysins may reflect the peculiar composition and structural features of the mycobacterial cell wall. While phages Ms6 and TM4 encode enzymes with an “Amid-2” type domain, others such as D29 and Bxb1 employ hydrolases with lysozyme-like domains to bring about host cell lysis (São-José *et al*., 2003; Hatfull *et al*., 2010). A search for conserved domains showed that Ms6 LysA holds a central peptidoglycan recognition protein (PGRP) super-family conserved domain (cl02712) with its 3 sites (amidase catalytic site, a Zn binding site, and a substrate binding site), localized between amino acid residues 170 and 300 (Fig. 4B). PGRPs are pattern recognition receptors that bind, and in certain cases, hydrolyze peptidoglycan of bacterial cell walls. This family includes Zn-dependent *N*-acetylmuramoyl-L-alanine amidases (EC: 3.5.1.28) which cleave the amide bond between *N*-acetylmuramoyl and L-amino acids, preferentially D-lactyl-L-Ala, in bacterial cell walls.

The mycobacteriophage Ms6 lysis module is closely related to the lysis module of phages belonging to cluster F, subcluster F1, which includes phages PMC, Llij, Che8, Boomer, Fruitloop, Pacc40, Ramsey and Tweety (Hatfull *et al*., 2010). A BLASTp search for Ms6 LysA homologues identified similar proteins among phages of this subcluster, except for the endolysin of mycobacteriophage Pacc40, and produced significant alignments with the N-terminal region of endolysins belonging to mycobacteriophages Corndog, Phyler, Phaedrus and Pipefish. Examination of the endolysin DNA sequence of these bacteriophages

revealed the existence in all of them of a second potential translation site positioned in close proximity to the beginning of the central PGRP domain (Fig. 4A, 4B). As we recently described, the accessory lysis protein Gp1 binds the N-terminal domain of LysA (Catalão *et al.*, 2010) and *gp1* was implicated in lysis because of its linkage to *lysA* and the demonstration of its requirement for mycobacteria efficient lysis. In view of these results, we considered whether the existence of two translational sites within *lysA* could be related with the interaction Gp1/LysA. Owing to their related amino acid sequences, the Gp1-like proteins have been recently grouped in the mycobacteriophage gene family Pham1480 (Hatfull *et al.*, 2010). In addition to subcluster F1, already mentioned, homologues of Ms6 Gp1 were also identified in phages belonging to subcluster A1 (Bethlehem, Bxb1, U2, DD5, Jasper, KBG, Lockley and Solon). Further examination of the endolysin DNA sequence of these bacteriophages revealed the existence of two potential translational sites in the *lysA* gene of bacteriophages Bethlehem, Bxb1, U2, DD5, KBG and Solon (Fig. 4A). While in some phages of subcluster A1 (Bethlehem, Bxb1, Jasper, Lockley and U2) and of subcluster F1 (Che8, Tweety, PMC, Llij, Fruitloop, Pacc40), *gp1*-like genes are overlapped with the lysin gene, in phages DD5, KBG and Solon (subcluster A1) and in phages Boomer and Ramsey (subcluster F1), Pham1480 is separated from the lysin gene by one intervening gene that code for putative homing endonuclease HNH motifs (Hatfull *et al.*, 2010). Although in these mycobacteriophages *gp1* and *lysA* are closely linked, a question arises as, while several mycobacteriophages seem to have kept somewhere in their genomes genes encoding proteins that could perform functions analogous to Gp1, others simply lack *gp1*.

A)

	SD	Start codon
Consensus	**AGGAGGT-------4-7 ---→ATG**	

Subcluster F1

Ms6	acaagtactggcccgccga**tggaggt**accgcc	**gtg**ccc	**435**
Che8	gccctcaacgaaaccacac**cggagga**acct	**gtg**acc	**519**
Tweety	acaagtactggcccgccga**tggaggt**accgcc	**gtg**ccc	**435**
PMC	acaagtactggcccgccga**tggaggt**accgcc	**gtg**ccc	**435**
Llij	acaagtactggcccgccga**tggaggt**accgcc	**gtg**ccc	**435**
Boomer	tcacaaccgcctaccacac**cggaggt**tccc	**atg**ccc	**528**
Fruitloop	acaagtactggcccgccga**tggaggt**accgcc	**gtg**ccc	**435**
Ramsey	gccctcaacgaaaccacac**cggagga**acct	**gtg**acc	**519**
Pacc40	actactacaaccagctcgc**aggagag**gacgcaaac	**gtg**gga	**447**

Subcluster A1

Bxb1	gcaacgactggtcgtcgcc**caagga**tgcg	**atg**cac	**502**
U2	gcaacgattggacgtcgcc**gaagga**cgcg	**atg**cat	**502**
Bethlehem	gcaacgactggacgtcgcc**gaagga**cgcg	**atg**cat	**502**
Solon	gtaatgactggaacagtcc**aaaaga**tgcg	**atg**cac	**384**
KBG	gtaatgactggaacagtcc**aaaaga**tgcg	**atg**cac	**384**
DD5	gtaatgactggaacagtcc**aaaaga**tgcg	**atg**cac	**384**

Subcluster B3

Pipefish	acccttcgacaccgacacc**aggaggc**acagctccc	**gtg**gtc	**483**
Phaedrus	acccttcgaccccgacacc**aggaggc**acagcgccc	**gtg**gtc	**483**

Unclustered

Phyler	acccttcgaccccgacacc**aggaggc**acagcgccc	**gtg**gtc	**525**
Corndog	agtactggcccagcaccac**gggagga**aca	**gtg**gcc	**450**

B)

Subcluster F1

```
gp30Llij           -----------MTTKDQVAQITIAEAKARGYTRSECLAIMSTFYQESG---WNDTIWDPT 46
gp30PMC            -----------MTTKDQVAQITIAEAKARGYTRSECLAIMSTFYQESG---WNDTIWDPT 46
gp2Ms6             -----------MTTKDQVAQITIAEAKARGYTRSECLAIMSTFYQESG---WNDTIWDPT 46
gp29Fruitloop      -----------MTTKDQVAQITIAEAKARGYTRSECLAIMSTFYQESG---WNDTIWDPT 46
gp30Tweety         -----------MTTKDQVAQITIAEAKARGYARSECLAVMSTFYQESG---WNDTIWDPT 46
gp32Che8           MSFTWFRPEGPLRTREQVAREVHAVSLARGLDELATVIALMTISTEVGTGIGDDRKWWCP 60
gp32Boomer         MSFTWFRPEGPLRTREQVAREVHAVSLARGLDELATVIALMTISTEVGTGTGDDRKWWCP 60
gp32Ramsey         MSFTWFRPEGPLRTREQVAREVHAVSLARGLDELATVIALMTISTEVGTGTGDDRKWWCP 60
                              : *::***: . * : ***  .   :  : *:  * *    :*  *  .

gp30Llij           HTTYGIAQQDGSYPHRFDGAAAQIKGFFDKLDVWRAKPGASTDIWLNICWMQQAPN---- 102
gp30PMC            HTTYGIAQQDGSYPHRFDGAAAQIKGFFDKLDVWRAKPGASTDIWLNICWMQQAPN---- 102
gp2Ms6             HTTYGIAQQDGSYPHRFDGAAAQIKGFFDKLDVWRAKPGASTDIWLNICWMQQAPN---- 102
gp29Fruitloop      HTTYGIAQQDGSYPHRFDGAAAQIKGFFDKLDVWRAKPGASTDIWLNICWMQQAPN---- 102
gp30Tweety         HTTYGIAQQDGSYPHRFDGAAAQIKGFFDKLDVWRAKPGASTDIWLNICWMQQAPN---- 102
gp32Che8           ANDRVPATKN--YPHDSRSDDNRSSGYFQQQPGPNGEPWWGTP--ENMMTLPQAANTFLE 116
gp32Boomer         ANDRVPATKN--YPHDSRSDDNRSSGYFQQQPGPNGEPWWGTP--ENMMTLPQAANTFLE 116
gp32Ramsey         ANDRVPATKN--YPHDSRSDDNRSSGYFQQQPGPNGEPWWGTP--ENMMTLPQAANTFLE 116
                    .    * ::  ***   .   : .*:*::    ..:*  .*    *:  : **.*

gp30Llij           WPSADYWYANGRRAYLTEIKSRIAT---VTPYLDKY---WPADG---G---TAVPD---- 146
gp30PMC            WPSADYWYANGRRAYLTEIKSRIAT---VTPYLDKY---WPADG---G---TAVPD---- 146
gp2Ms6             WPSADYWYANGRRAYLTEIKSRIAT---VTPYLDKY---WPADG---G---TAVPD---- 146
gp29Fruitloop      WPSADYWYANGRRAYLTEIKSRIAT---VTPYLDKY---WPADG---G---TAVPD---- 146
gp30Tweety         WPSADYWYANGRRAYLTEIKSRITT---VTPYLDKY---WPADG---G---TAVPD---- 146
gp32Che8           RLSDDYRRAANNPRLAGEFAQRVQQSEFPDRYADKWDEAWSVLRRALN---ETTPEEPVT 173
gp32Boomer         RLSDDYRRAANNPRLAGEFAQRVQQSEFPDRYADKWDEAWSVLRRALGSQPPTTPEVPMP 176
gp32Ramsey         RLSDDYRRAANNPRLAGEFAQRVQQSAYPDRYADKWDEAWSVLRRALN---ETTPEEPVT 173
                     * **  * ..     *: .*:        * **:   *..     .    :.*:

gp30Llij           EPRPDFNEFPIWSAN--NSSRSG-KPTMFLIHTQEGGGG-DAAAENLAKWFQN-ANGVSY 201
gp30PMC            EPRPDFNEFPIWSAN--NSSRSG-KPTMFLIHTQEGGGG-DAAAENLAKWFQN-ANGVSY 201
gp2Ms6             EPRPDFNEFPIWSNN--NSARSG-KPTMFLIHTQEGGGG-DAAAENLAKWFQN-GNGVSY 201
gp29Fruitloop      EPRPDFNEFPIWSNN--NSARSG-KPTMFLIHTQEGGGG-DAAAENLAKWFQN-GNGVSY 201
gp30Tweety         EPRPDFNEFPIWSAN--NSSRSG-KPTMFLIHTQEGGGG-DAAAENLAKWFQN-SNGVSY 201
gp32Che8           ENRPAYNEFPIWSAN--NSARSG-KPTMFLIHTQEGGGG-DAAAENLAKWFQN-ANGVSY 228
gp32Boomer         ENRPPYNEFPIWSENHYNTKRAANDIDAFLLHTSEGFVGRDDAAEALSLWYQPRSRQVAY 236
gp32Ramsey         ENRPAYNEFPIWSAN--NSARSG-KPTMFLIHTQEGGGG-DAAAENLAKWFQN-ANGVSY 228
                   * ** :******* *  *: *:. .   **:**.**  * * *** *: *:*  .. *:*

gp30Llij           HYTISQASDGGVTVVDCVDTDRAAWSVGNANSISINLCFAGSRAAWSREQWM-KQSNAID 260
gp30PMC            HYTISQASDGGVTVVDCVDTDRAAWSVGNANSISINLCFAGSRAAWSREQWM-KQSNAID 260
gp2Ms6             HYTISQASDGGVTVVDCVDTDRAAWSVGNANSISINLCFAGSRASWMRDQWM-KQSNAID 260
gp29Fruitloop      HYTISQASDGGVTVVDCVDTDRAAWSVGNANSISINLCFAGSRASWMRDQWM-KQSNAID 260
gp30Tweety         HYTISQASDGGVTVVDCVDTDRAAWSVGNANSISINLCFAGSRASWTLDQWM-KQSNAID 260
gp32Che8           HYTISQASDGGVTVVDCVDTDRAAWSVGNANSISINLCFAGSRAAWSRDQWM-KQSNAID 287
gp32Boomer         HYAVSQASDGGVTVVDNVDTDYASWSALSANGRSINLCFAGTRAAWSRNEWLGKFGNAID 296
gp32Ramsey         HYTISQASDGGVTVVDCVDTDRAAWSVGNANSISINLCFAGSRASWTLDQWM-KQSNAID 287
                   **::************ **** *:**. .**. ********:**:*  ::*: * .****

gp30Llij           VAAYLAVQDAKKYGFEPLVVPPPYVNGHPGISDHRWVTDVFKWGTHTDVGDWFPWDYFTE 320
gp30PMC            VAAYLAVQDAKKYGFEPLVVPPPYVNGHPGISDHRWVTDVFKWGTHTDVGDWFPWDYFTE 320
gp2Ms6             VAAYLAVQDAKKYGFTPLVVPPPYTNGRPGISDHRWVTDVFKWGTHTDVGDWFPWDYFAE 320
gp29Fruitloop      VAAYLAVQDAKKYGFTPLVVPPPYTNGRPGISDHRWVTDVFKWGTHTDVGDWFPWDYFAE 320
gp30Tweety         VAAYLAVQDAKKYGFEPLVVPPPYTNGRPGISDHRWVTDVFKWGTHTDVGDWFPWDYFAE 320
gp32Che8           VAAYLAVQDAKKYGFEPLVVPPPYVNGRPGISDHRWVTDVFKWGTHTDVGDWFPWDYFTE 347
gp32Boomer         VAAYLAVQDCKKYNIPTKVIAPPYTGRLPGITDHRYVTQILKDGTHTDVGDGFPWDYFTE 356
gp32Ramsey         VAAYLAVQDAKKYGFEPLVVPPPYVNGRPGISDHRWVTDVFKWGTHTDVGAWFPWDYFTE 347
                   *********.***.: . *:.***..  ***:***:**:::* *******  ******:*
```

```
gp30Llij          RVAFWANGGASEPEPPKVKRFPDDWTDRELAVETLRQQRGYALDGWPQLGGRTVVDVLGA 380
gp30PMC           RVAFWANGGASEPEPPKVKRFPDDWTDRELAVETLRQQRGYALDGWPQLGGRTVVDVLGA 380
gp2Ms6            RVNHWANGGKTEPEPPKVKRFPDDWSDREILVEILRQLRGYNLTGWPQLGGKTLVDAVAE 380
gp29Fruitloop     RVNHWANGGKTEPEPPKVKRFPDDWSDREILVEILRQLRGYTLTGWPQLGGKTLVDAVAE 380
gp30Tweety        RVNHWANGGKTEPEPPKVKRFPDDWSDREILVEILRQLRGYNLTGWPQLGGKTLVDAVAD 380
gp32Che8          RVNHWAAGGKTEPEPPKVKRFPDDWTDRELAVETLRQQRGYALDGWPQLGGRTVVDALGA 407
gp32Boomer        RVNHWAAGDKTEPEPPKVKRFPDDWTDRELGVETLRQQRGYTLNGWPQLGGRTVVDVLGA 416
gp32Ramsey        RVNHWANGGKTEPELPKVKRFPDDWTDREILVEILRQLRGYNLTGWPQLGGKTLVDAIAE 407
                  ** .** *. :*** **********:***: ** *** *** * *******:*:**.:.

gp30Llij          IGEKLG-IEGCYDVKGKS----- 397
gp30PMC           IGEKLG-IEGCYDVKGKS----- 397
gp2Ms6            LLGH------------------- 384
gp29Fruitloop     LLGH------------------- 384
gp30Tweety        LRTDIIDLQGAIEHGEITLGGAQ 403
gp32Che8          IGAKLG-IEGCYDVKGKS----- 424
gp32Boomer        IGEKLG-VEGCYDVKDKS----- 433
gp32Ramsey        IRN---------VVCDK------ 415
                  :
```

Figure 4. Occurrence of Ms6 LysA-like proteins in mycobacteriophages. A) Potential translational signals in mycobacteriophages *lysA*-like genes. Sequences corresponding to the putative translational signals (Shine-Dalgarno sequences and associated translational start codons) are shown in bold. The *E. coli* consensus for each of these signals is also shown for comparison. Numbers refer to nucleotide positions. **B)** CLUSTALW alignment of Ms6 LysA deduced amino acid sequence and putative LysA protein sequences of mycobacteriophages included in subcluster F1. Mycobacteriophages: Llij gp30 (YP655026), PMC gp30 (YP655791), Ms6 gp2 (AAG48318), Fruitloop gp29 (YP002241714), Tweety gp30 (YP001469263), Che8 gp32 (NP817370), Boomer gp32 (YP002014248) and Ramsey gp32 (YP002241819); the primary accession numbers of the UniProtKB/TrEMBL database are given in parentheses. Identical (*), highly similar (:) and similar (.) amino acids are indicated. Dashes represent gaps introduced by CLUSTALW to optimize the alignment. The PGRP superfamily conserved domain is highlighted on a grey background. Numbers refer to amino acid positions. Predicted start codons are shown in bold.

Role of Lysin$_{384}$ and Lysin$_{241}$ during **M. smegmatis** *phage infection*

To better understand the contribution of LysA (Lysin$_{384}$) to the Ms6 infection cycle, we first investigated whether it is an essential function for host lysis. We used the Bacteriophage Recombineering of Electroporated DNA (BRED)

strategy to delete *lysA* gene (*gp*2) from the mycobacteriophage Ms6 genome. A 200 bp substrate containing a 100 bp sequence flanking *lysA* was designed to introduce a 1089 bp internal-deletion in Ms6 *lysA*, fusing 19 codons at the 5´ and 3´ends of the gene to maintain the *lysB* RBS and minimize effects on expression of the adjacent *lys*B gene, as well as avoiding genetic polarity. Co-electroporation of the 200 bp deletion substrate and Ms6 genomic DNA into a *M. smegmatis* recombineering strain yielded 73 primary plaques. 25 out of the 73 plaques were screened by DADA-PCR and three contained detectable levels of the mutant phage mixed with the wild-type. Although we predicted *lysA* to be essential [as recently the LysA of mycobacteriophage Giles was demonstrated to be essential for lytic growth (Marinelli *et al*., 2008)] the mutant presumably grows in the mixed plaque through assistance of wild-type helper phage. To demonstrate Ms6 *lysA* essentiality, serial dilutions of the three mixed plaques were re-plated and the secondary lysate was screened by DADA-PCR. The mutant allele is no further present in the lysate for all three mixed plaques, which indicates that the mutant is not viable and that LysA is indeed essential for Ms6 propagation (data not shown).

As a result of synthesis of $Lysin_{241}$ and $Lysin_{384}$ during phage infection, we next investigated the influence of each endolysin form in phage growth parameters. Once more, we took advantage on the fact that the BRED recombineering strategy has already been described to efficiently introduce base changes that confer an amino acid substitution (Marinelli *et al*., 2008). To eliminate synthesis of $Lysin_{384}$ or $Lysin_{241}$ we designed two complementary oligonucleotides that introduce a stop codon and a HindIII restriction site downstream of the start codon of $Lysin_{384}$, or two complementary oligonucleotides that modify the $Lysin_{241}$ GTG start codon (valine) to TGG (tryptophan), and introduce an MscI restriction site. Oligonucleotides were co-transformed with Ms6-$LysAHis_6$ genomic DNA; 18 primary plaques were screened by PCR and HindIII or MscI digestion, one of which was clearly a mixed plaque. Serial dilutions of the mixed plaques were re-

plated and both the secondary lysate and individual secondary plaques were screened by PCR with the same pair of primers followed by HindIII or MscI restriction; as the mutant allele is still present in the lysate (Fig. 5), this indicates that both Ms6-LysAHis$_{6TGA}$ and Ms6-LysAHis$_{6GTG\rightarrow TGG}$ mutants are viable. In addition, two pure mutant phages of Ms6-LysAHis$_{6TGA}$ and Ms6-LysAHis$_{6GTG\rightarrow TGG}$ were readily identified by screening 18 plaques from secondary plating. These results demonstrate that for plaque formation only one of the two LysA forms, Lysin$_{384}$ or Lysin$_{241}$ is required.

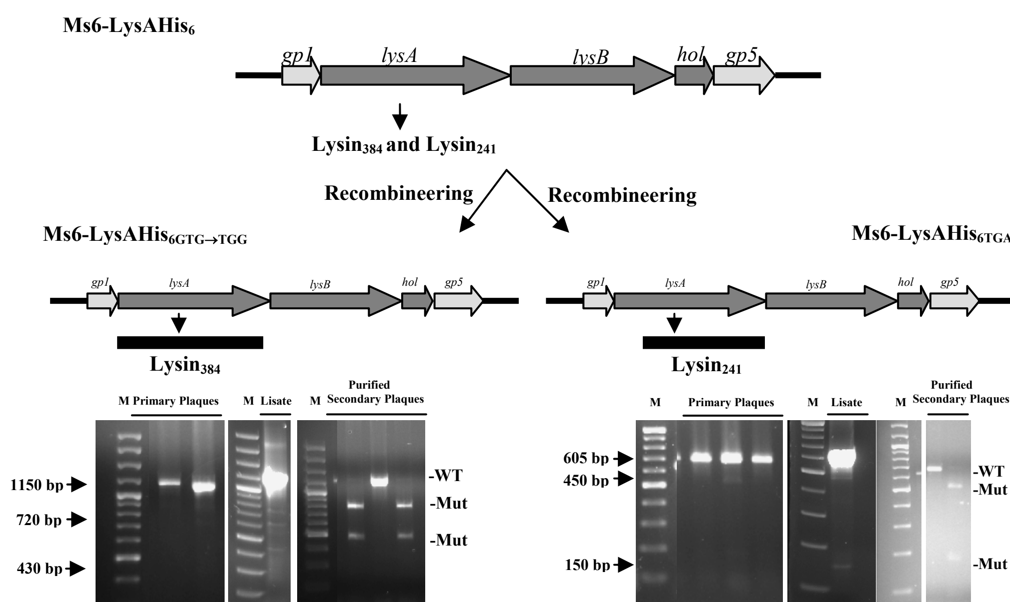


Figure 5. Strategy for construction of Ms6 *lysA* mutant derivatives. Two complementary oligonucleotides that modify *lysin*$_{241}$ GTG start codon (valine) to TGG (tryptophan) and introduce an MscI restriction site, or two complementary oligonucleotides that introduce a stop codon and a HindIII restriction site downstream of the start codon of *lysin*$_{384}$, were co-transformed with Ms6-LysAHis$_6$ genomic DNA; primary plaques were recovered and screened by PCR and MscI or HindIII digestion to identify a mixed plaque

containing wild-type and mutant phages DNA. The mixed primary plaque was diluted and plated; the lysate was screened to check for phage viability, and purified secondary plaques were screened to identify pure mutant phages of Ms6-LysAHis$_{6GTG \rightarrow TGG}$ and Ms6-LysAHis$_{6TGA}$, expressing only Lysin$_{384}$ or Lysin$_{241}$, respectively.

Nevertheless, we considered whether the absence of one of the two endolysin forms during *M. smegmatis* phage infection (Fig. 6) could confer an altered lysis phenotype.

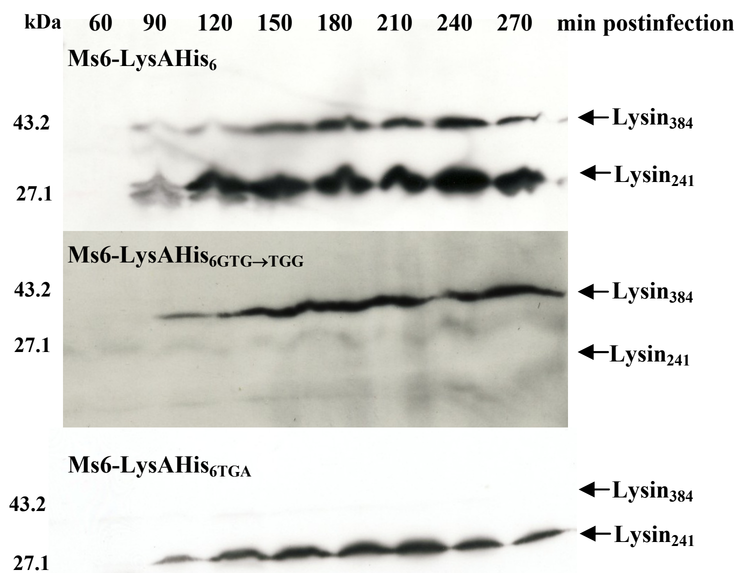
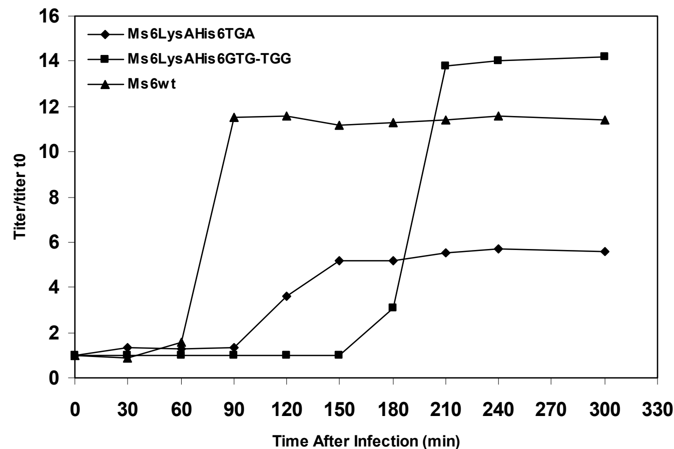


Figure 6. Time course of Lysin$_{384}$ and Lysin$_{241}$ synthesis during Ms6-LysAHis$_6$, Ms6-LysAHis$_{6GTG \rightarrow TGG}$ or Ms6-LysAHis$_{6TGA}$ infection of *M. smegmatis*. Lysin production in *M. smegmatis* was analysed after infection at an m.o.i. of 10. Extracts were prepared from samples taken at 30-min intervals as described in Material and Methods. Samples were analysed by Western-blotting and Lysin$_{384}$ and Lysin$_{241}$ synthesis was detected as already described. Both Lysin$_{384}$ and Lysin$_{241}$ synthesis could be detected beginning 90 minutes postinfection in Ms6-LysAHis$_6$ (upper panel). In Ms6-LysAHis$_{6GTG \rightarrow TGG}$ and Ms6-

LysAHis$_{6TG}$, (lower panels) only Lysin$_{384}$ or Lysin$_{241}$ synthesis could be detected also beginning 90 min postinfection, respectively. Only the results for 60 to 270 min postinfection are shown. The molecular masses in kDa of Lysin$_{384}$ and Lysin$_{241}$ are indicated on the left; positions of both proteins are indicated by an arrow on the right.

To address this question one step growth curves and determination of phage growth parameters (latent period, rise period and burst size) were carried out to compare the phages infection cycle. The one step growth experiment (Fig. 7A) shows that phage Ms6-LysAHis$_{6TGA}$ has slightly prolonged latent period of 140 minutes (corresponding to 90 minutes represented in Fig. 7A plus the initial 50 min of adsorption) when compared to the latent period of 110 min of the wild-type phage, and a rise period of 60 minutes, which means that a delay exists in the detection of phage release from cells infected with Ms6-LysAHis$_{6TGA}$. In addition, Lysin$_{384}$ is important for infective particles release with its absence leading to smaller size phage plaques (Fig. 7B, panel 2) when compared to Ms6$_{wt}$ phage (Fig. 7B, panel 1). On the other hand, Ms6-LysAHis$_{6GTG \rightarrow TGG}$ phage release is only detected 230 min after *M. smegmatis* infection. This indicates that similar to Lysin$_{384}$, Lysin$_{241}$ has an obvious function in completion of lysis, although it does not have an apparent effect in phage release.

A)



B)

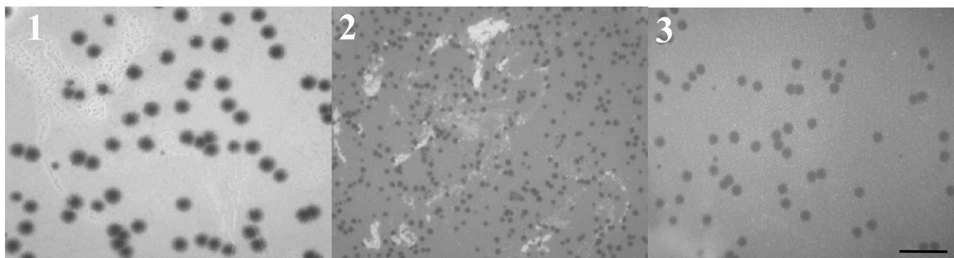


Figure 7. A) One step growth curves of mycobacteriophage Ms6 and lysin-mutant derivatives. For each curve the titers measured were divided by the titer at t=0 for normalization (titer/titer t_0). **B)** Plating ability of the lysin-mutant bacteriophages. **1.** *M. smegmatis* infected with $Ms6_{wt}$ phage; **2.** *M. smegmatis* infected with Ms6-$LysAHis_{6TGA}$; **3.** *M. smegmatis* infected with Ms6-$LysAHis_{6GTG \rightarrow TGG}$. Scale bar indicates 1 cm.

Examination of Ms6-$LysAHis_{6GTG \rightarrow TGG}$ phage plaques shows that although not smaller in size, plaques are more turbid than wild-type Ms6 probably due to a partial host cell lysis (Fig. 7B, panel 3). These results strongly suggest that even though only one of the two LysA forms, $Lysin_{384}$ or $Lysin_{241}$, is required to accomplish host cell lysis, both enzymes are necessary for complete and efficient lysis of *M. smegmatis*.

$Lysin_{384}$ and $Lysin_{241}$ are cell wall-degrading enzymes that support peptidoglycan hydrolytic activity

$Lysin_{384}$ (LysA) has been previously described as not affecting *E. coli* growth rate unless permeabilization of the plasma membrane by chloroform addition which results in immediate lysis (Garcia *et al.*, 2002). To follow growth and viability of *E. coli* strains expressing $Lysin_{241}$, the *lysA* DNA fragment corresponding to *$lysin_{241}$* was cloned into pQE30 vector and $Lysin_{241}$ was expressed under the control of a regulated T5 bacteriophage promoter. As shown in Fig. 8A,

although induction of $Lysin_{241}$ did not result in *E. coli* lysis unless chloroform was added, growth seems to be halted over the induction period. Nevertheless, the exact mechanism of toxicity is unknown.

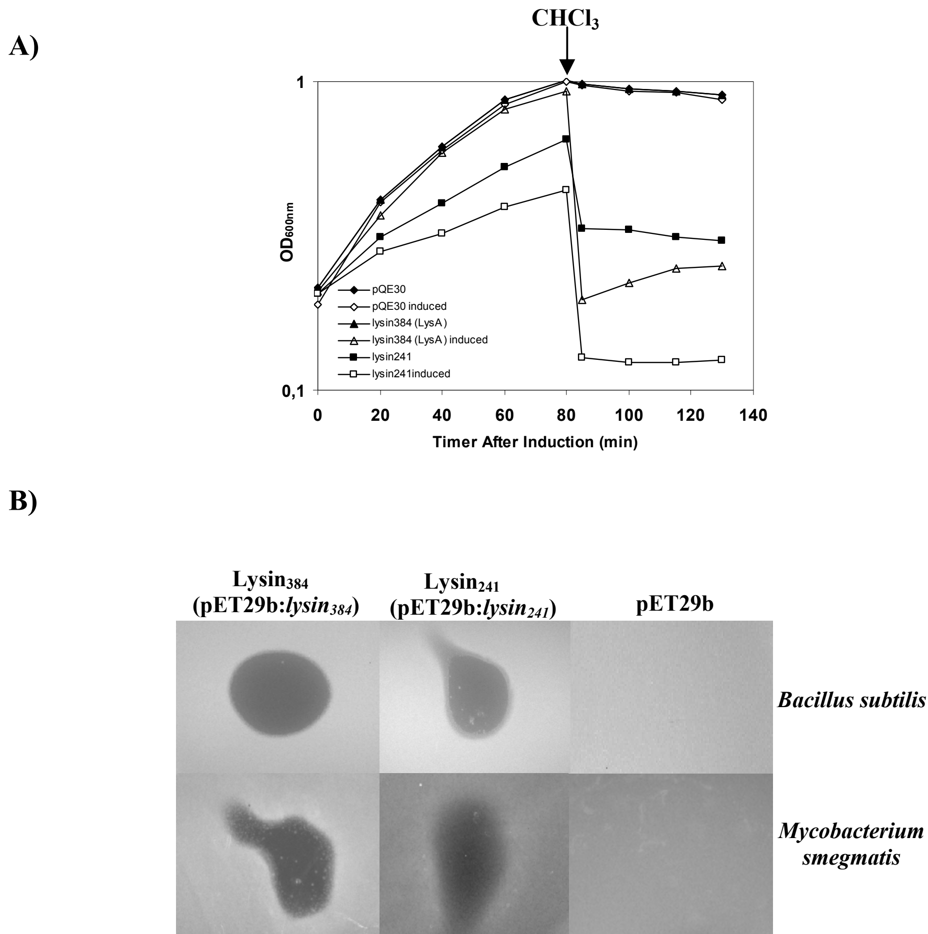


Figure 8. Lytic activity of $Lysin_{384}$ and $Lysin_{241}$. A) Effect of lysin expression on growth and viability of *E. coli*. *E. coli* JM109 cells carrying plasmid pQE30 containing no insert or cloned genes (*lysin*$_{384}$ or *lysin*$_{241}$) were grown in LB broth at 37 ºC to an OD_{600nm} of 0.2. At time zero, transcription of cloned lysis genes was induced by addition of 1 mM IPTG. At the time indicated by the arrow, 2% $CHCl_3$ was added to the cultures. Results are averages of triplicate experiments. **B)** Effect of $Lysin_{384}$ or $Lysin_{241}$ activity on lawns of *B. subtilis*

(upper panel) and *M. smegmatis* (lower panel). 20 µl of *E. coli*:pMJC41 or *E. coli*:pMJC42 extracts containing $Lysin_{384}$ or $Lysin_{241}$ were spotted onto the bacterial lawn of the test strain and incubated overnight at 37 ºC. After overnight incubation, the presence of a clear zone was examined. *E. coli*:pET29b induced extract was used as a negative control.

As mentioned before, Ms6 LysA holds a PGRP domain and its hydrolase activity was already demonstrated; the purified protein was shown to cleave the bond between L-Ala and D-muramic acid and to release up to 70% of the diaminopimelic acid present in isolated mycobacterial cell walls which confirmed the amidase activity of the enzyme (Piechota *et al.*, unpublished results). To more directly assess the enzymatic activity of $Lysin_{384}$ and $Lysin_{241}$ we tested their ability to generate a zone of clearing in a zymogram assay. Lytic activity in lysin-producing *E. coli* extracts was checked by *in situ* protein renaturation after SDS-PAGE, using gel-incorporated autoclaved *M. luteus* cells as the substrate. As shown in Fig. 9 both $Lysin_{384}$ and $Lysin_{241}$ have hydrolase activity. Peptidoglycan hydrolysis has been also demonstrated for other three mycobacteriophage LysA proteins (gp32 Che8, gp236 Bxz1 and gp69 Corndog) in zymograms (Payne *et al.*, 2009).

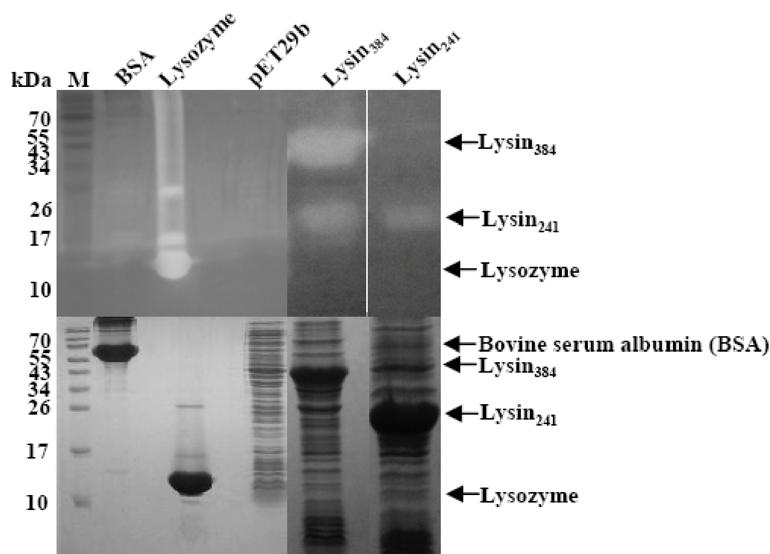


Figure 9. Peptidoglycan hydrolysis by *E. coli*-produced $Lysin_{384}$ and $Lysin_{241}$ in *M. luteus* cells. Lytic activity of lysin extracts was assessed by *in situ* renaturation after SDS-PAGE using a gel matrix containing *M. luteus* cells as substrate (upper panel). Peptidoglycan hydrolysis by renatured proteins within the gel produces clear zones that no longer stain with methylene blue. Lysozyme and bovine serum albumin (BSA) represent positive and negative controls, respectively. A cell-free control gel was run in parallel and stained with Coomassie blue (lower panel). The molecular masses in kDa of bovine serum albumin (BSA), lysozyme, $Lysin_{384}$ and $Lysin_{241}$ are indicated on the left; positions of proteins are indicated by an arrow on the right.

Despite the fact that Ms6 LysA is produced by a phage that infects a bacteria with a complex envelope and taking into consideration the potential application of bacteriophages and purified endolysins, the antibacterial activities of bacteriophages $Ms6_{wt}$, $Ms6\text{-}LysAHis_{6TGA}$, $Ms6\text{-}LysAHis_{6GTG \rightarrow TGG}$ and of the endolysins $Lysin_{384}$ and $Lysin_{241}$, were determined following the analysis of clearing formation on viable bacterial strains. Results are summarized in Table 3.

Table 3. Antibacterial activity of Ms6 and derivative mutants and its lytic proteins.

Isolate or strain	Bacteriophage activity*			Lytic protein activity*	
	$Ms6_{wt}$	$Ms6\text{-}LysAHis_{6TGA}$	$Ms6\text{-}LysA His_{6GTG \rightarrow TGG}$	$Lysin_{384}$	$Lysin_{241}$
Mycobacterium smegmatis	++	++	++	++	++
Mycobacterium vaccae	+	+	-	++	++
Mycobacterium aurum	++	++	-	++	++
Mycobacterium fortuitum	+	+	-	-	-
Enterococcus faecium	-	-	-	-	-
Enterococcus faecalis	-	-	-	-	-
Streptococcus group A	-	-	-	+	+
Micrococcus luteus	-	-	-	++	++
Micrococcus pyogenes	-	-	-	++	++
Bacillus subtillis	-	-	-	++	++
Bacillus pumilus	-	-	-	++	++
Staphylococcus aureus	-	-	-	+	+
Staphylococcus epidermidis	-	-	-	+	+
Escherichia coli	-	-	-	+	+

*Designations refer to bacterial lawn clearance and are as follows: ++, clearance observable at the site of bacteriophage or lytic protein application; + partial clearance observable at the site of bacteriophage or lytic protein application; -, no clearance.

E. coli extracts containing $Lysin_{384}$ or $Lysin_{241}$ showed strong activity against all mycobacteria tested except *M. fortuitum*, as well as against various Gram-positive bacteria, *M. luteus*, *M. pyogenes*, *B. subtilis* and *B. pumilus*. $Lysin_{384}$ and $Lysin_{241}$ exhibited weak activity towards *Streptococcus pyogenes*, *S. aureus*, *S. epidermidis* and the Gram-negative bacterium *E. coli.* For each strain tested, a crude extract of *E. coli* BL21 (DE3):pET29b was used as a control, and no lysis halo was observed, whatever the substrate (Fig. 8B).

Halo formation by mycobacteriophage Ms6

When Ms6 was first isolated, it was noticed that it forms plaques in *M. smegmatis* lawns with halos surrounding them. Formation of the halo begins after two days of incubation, once the infected area is fully formed. Halo formation around phage plaques seems to be unusual among mycobacteriophages. Nevertheless, this phenomenon has been described for, at least, mycobacteriophages Bxb1 (Mediavilla *et al*., 2000) and Che12 (Kumar *et al.*, 2008) and has been observed for a number of other phages, particularly those infecting encapsulated and/or mucoid Enterobacteriaceae (Stirm *et al*., 1971) and Gram-positive bacteria; plaques of phage A511 infecting *Listeria monocytogenes* were shown to form clearly visible secondary lysis zones caused by release of Ply511, a hydrophilic amidase endolysin that diffuses from the centre of the plaque and hydrolyses the surrounding cells “from without” (Loessner *et al*., 2005).

Bxb1 forms plaques with large halos, and formation begins after 3 days of incubation. Upon further incubation, the halo grows in size and spreads out across the *M. smegmatis* lawn until, after 18 days it encompasses a large part of the entire lawn. This enzymatic activity was attributed to a structural component of Bxb1

virions or to Gp8 (Mediavilla *et al*., 2000), however its origin remains unknown. The Gp8 protein is the LysA-like protein of Bxb1 that, as already mentioned, possesses within its DNA sequence two potential translation sites (Fig. 4A) that could direct the synthesis of two endolysin forms as it happens for Ms6 LysA. We considered whether this gene arrangement could be implicated in halo formation as both D29 and L5 phages, whose virion particles are indistinguishable from Bxb1, do not form halos around the lysis plaques. Furthermore, the two translational sites gene organization could not be identified in the DNA sequence of L5 or D29 *gp10*. To address this question, *M. smegmatis* cells in the exponential growth phase were infected with phages D29, $Ms6_{wt}$, $Ms6\text{-}LysAHis_{6TGA}$ or $Ms6\text{-}LysAHis_{6GTG \rightarrow TGG}$. As expected, halo formation around D29 phage plaques was not observed (Fig. 10). $Ms6_{wt}$ forms turbid plaques showing size variation with partially clear halos surrounding them (Fig. 10). When we analysed halo formation in *M. smegmatis* infected with phages $Ms6\text{-}LysAHis_{6TGA}$ or $Ms6\text{-}LysAHis_{6GTG \rightarrow TGG}$, we observed that although this phenomenon can still be detected, halo size and turbidity is distinct from the ones observed after *M. smegmatis* $Ms6_{wt}$ infection. Fig. 10 shows that both phages $Ms6\text{-}LysAHis_{6TGA}$ and $Ms6\text{-}LysAHis_{6GTG \rightarrow TGG}$ form smaller and much more turbid halos than the wild-type phage. We assumed that halo formation may result from the diffusion of $Lysin_{384}$ and $Lysin_{241}$ from the $Ms6_{wt}$ phage infective centre which leads to "lysis from without" in surrounding cells. As both endolysins forms of LysA are essential to obtain an effective lysis of *M. smegmatis*, halo turbidity could result from partial lysis of cells surrounding the phage plaques. However, these results do not rule out the existence in the Ms6 genome of other genes whose products could be implicated in the halo formation phenomenon.

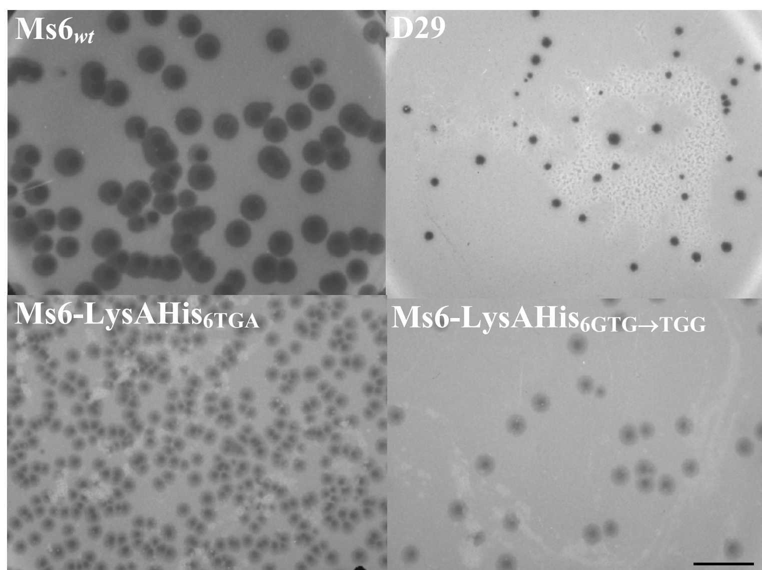


Figure 10. Halo formation by mycobacteriophage Ms6 and lysin-mutant derivatives. Serial dilutions of the bacteriophages stocks containing ~10^{10} particles ml^{-1} were prepared and 100 µl of the 10^{-8} dilution was plated with 200 µl of an exponential growing culture of *M. smegmatis* as top agar lawns. Incubation was continued for 3 days at 37 ºC. Mycobacteriophage D29 was used as a negative control of halo formation (Mediavilla *et al*., 2000). Scale bar represents 1 cm.

Discussion

In the present work, we provide evidence that the mycobacteriophage Ms6 *lysA* gene encodes two overlapped lysin proteins ($Lysin_{384}$ and $Lysin_{241}$), in the same reading frame. Our group has recently described that the N-terminal domain of $Lysin_{384}$ is necessary and sufficient to directly interact with Gp1, a chaperone-like protein located upstream of LysA, that assists the translocation of the endolysin across the cytoplasmic membrane in a holin-independent way (Catalão *et al*., 2010). Interestingly, however, during an attempt to overproduce LysA in a purified form as a C-terminal histidine-tagged fusion product, we detected the co-purification of two polypeptides with 43.2 and 27.1 kDa, the larger one

corresponding to the predicted Lysin$_{384}$ molecular mass. These data raised the possibility that the shorter protein could be the mature form of Lysin$_{384}$. If a cleavage event had occurred, it would implicate the removal of 143 amino acids of the enzyme corresponding to 16 kDa. Although unusual, as generally mycobacterial signal peptides (SP) length is 32 residues (Wiker *et al.*, 2000), comparable to the lengths of SPs from Gram-positive bacteria (von Heijne and Abrahmsén, 1989), cleavage of a large segment of 143 amino acids that functions as a signal peptide has already been described for the *Staphylococcus simulans* lysostaphin (Recsei *et al.*, 1987). Indeed, a 16 kDa protein was detected by Western-blotting even though no sequence signals or cleavage sites were detected by bioinformatic approaches in the N-terminal region of LysA. Further investigation is needed to clarify the meaning of this result. Examination of the *lysA* nucleotide sequence revealed the existence of a second possible lysin gene overlapped with *lysA* in the same reading frame and preceded by a consensus Shine and Dalgarno sequence. To rule out the occurrence of a cleavage event, as observed for the endolysin of bacteriophage fOg44 (São-José *et al.*, 2000) we used different experimental approaches that confirmed the existence in *lysA* of two translation events, both in *E. coli* and during *M. smegmatis* infection. Further examination of the literature revealed that this gene arrangement is very uncommon among bacteriophage endolysins; however, at least one exception was found for the bacteriophage ϕvML3 endolysin gene that encodes two proteins, a larger lysin that has homology with lysozymes and a smaller lysin protein that has some features resembling those of a holin (Shearman *et al.*, 1994). However, a BLASTn search for Ms6 *lysA* homologues revealed that this peculiar endolysin gene arrangement is widespread in mycobacteriophages, in particular among those that possess a *gp1*-like gene (Pham1480) (Hatfull *et al.*, 2010), which suggests that Gp1 may confer a selective advantage for host cell lysis under different environmental conditions. In addition, two putative translational signals were also identified in endolysin genes

belonging to four mycobacteriophages that do not possess Gp1 homologues (Phyler, Phaedrus, Pipefish and Corndog) but possess N-terminal related Ms6 LysA sequences. The lack of representation of Pham1480 upstream of *lysA* in these phages could result from loss of *gp1*-like gene in these genomes. Furthermore, in three mycobacteriophages that possess Gp1 similar proteins but unrelated Ms6 LysA enzyme (TM4, Jasper and Lockley), this *lysA* gene arrangement was not observed which suggests that Pham1480 in these mycobacteriophages might result from recent acquisition by horizontal genetic exchange (Hatfull *et al*., 2010). These data also support the idea that all of these genomes have been in genetic communication, as Pham1480 is restricted to mycobacteriophages (Hatfull *et al*., 2010), and reflect the highly sequence diversity and modular nature of mycobacteriophage genomes that are characteristically mosaic comprising modules (frequently containing just a single gene) or cassettes, many of which shared by other phage genomes (Pedulla *et al*., 2003; Hatfull, 2005). We observed that synthesis and/or stability of the larger endolysin ($Lysin_{384}$) is highly dependent on Gp1 production. A reasonable explanation is that in the absence of its chaperone, the endolysin becomes unstable in the cytoplasm, or that an efficient translation of LysA is more or less dependent upon translation of the adjacent *gp1*-coding region as suggested by overlapping stop/start codons. In fact, when analysing LysA production during the life infective cycle of both Ms6-LysAHis$_6$ and $Ms6_{\Delta gp1}$-LysAHis$_6$ phages, we observed a decrease in the $Lysin_{384}$ levels in cells infected with $Ms6_{\Delta gp1}$, although $Lysin_{241}$ synthesis remains apparently unaffected. However, this is not the result of a polar effect at the transcriptional level as infection of *M. smegmatis* cells expressing the wild-type Gp1 protein with $Ms6_{\Delta gp1}$, leads to a reversion of the lysis defect (Catalão *et al*., 2010). Construction of Ms6 mutant phages deleted in *lysA* or defective for $Lysin_{384}$ or $Lysin_{241}$ synthesis showed that LysA is essential for host cell lysis. In fact, as pointed out by R. Young (2005) endolysins are always essential (for dsDNA phages) in terms of plaque-forming

ability, whereas holins may be not; indeed, for mycobacteriophage Ms6, LysA is the only lysis function that can not be suppressed and is indispensable for lysis, even though deletion of the additional lysis genes (*gp1*, *lysB*, *gp4* and *gp5*) may result in poor phage viability and severe lysis defects (Catalão *et al*., 2010; Catalão and Gil, unpublished results). Suppression of $Lysin_{384}$ or $Lysin_{241}$ synthesis does not result in a non-lysis phenotype as both proteins contain the PGRP domain. However, lack of $Lysin_{384}$ or $Lysin_{241}$ in phage virion results in an altered lysis phenotype; analysis of the $Ms6\text{-}LysAHis_{6TGA}$ (defective for $Lysin_{384}$ synthesis) and $Ms6\text{-}LysAHis_{6GTG \rightarrow TGG}$ (defective for $Lysin_{241}$ synthesis) phage growth parameters revealed that, whereas $Lysin_{384}$ is necessary to achieve a normal burst of infective phages, $Lysin_{241}$ has an important function in the progression and complete host cell lysis. These results suggest that while $Lysin_{384}$ is continuously secreted to the periplasm during phage assembly in order to guarantee rapid cell lysis, $Lysin_{241}$ acts at a later stage of infection (when lysis is triggered) and only once an adequate number of progeny virions is reached intracellularly, allowing completion of host cell lysis. At this time it is unknown if $Lysin_{384}$ activation is dependent on holin function or if instead, it remains fully active upon translocation across the cytoplasmic membrane assisted by Gp1, independently of holin depolarization of the plasma membrane. Indeed, as what happens with SAR endolysins that spontaneously escapes the bilayer at a low rate, resulting in that cells eventually undergo lysis (Sun *et al*., 2009), $Lysin_{384}$ also possesses some lytic activity before lysis is triggered. However, the poor viability of an Ms6 derivative deleted in *gp4* and *gp5* holin genes showed that similarly to secretory SAR endolysins that are first secreted as enzymatically inactive enzymes anchored to the membrane by an N-terminal terminal SAR domain, the secreted Ms6 $Lysin_{384}$ must be under some extracytoplasmic regulatory mechanism that ensures that premature lysis does not take place. It is also possible that Gp1 plays a role in maintenance of $Lysin_{384}$ inactive state: Gp1 binding to the N-terminal domain may block substrate binding

or Gp1 may allow $Lysin_{384}$ to adopt an active conformation. Indeed, the fact that $Lysin_{384}$ is detected almost exclusively in the presence of Gp1 suggests that Gp1 might affect the conformation of the endolysin either during or after translation. Chaperone-synthesis dependence has been already described for some lipases (Hobson *et al*., 1993; Frenken *et al*., 1993a; Frenken *et al*., 1993b; Kok *et al*., 1995). The energized state of the cytoplasmic membrane was also described as being implicated both in autolysins activation (Jolliffe *et al*., 1981; Blackman *et al*., 1998; Smith *et al*., 2000; Patton *et al*., 2006) and secretory endolysins activation (São-José *et al*., 2000; Xu *et al*., 2004; Nascimento *et al*., 2008). Recently an autoinhibition and activation mechanism was proposed for Auto, a virulence-associated peptidoglycan hydrolase of *Listeria monocytogenes*. The enzyme is autoinhibited when newly secreted but activated by proteolytic cleavage of the N-terminal α-helix that physically blocks the substrate-binding cleft (Bublitz *et al*., 2009). Work is currently ongoing to try to understand the mechanisms underlying $Lysin_{384}$ regulation.

An interesting observation was the fact that Ms6 forms turbid plaques surrounded by a clear zone of apparent bacterial lysis. This phenomenon is widely observed among bacteriophages that infect Gram-positive hosts and results from "lysis from without" of bacteria, as a result of endolysin diffusion from phage plaques that kills uninfected cells (Loessner, 2005). Although uncommon in mycobacteriophages, halo formation surrounding phage plaques that expand to cover a large area of uninfected cells has been reported for mycobacteriophage Bxb1 (Mediavilla *et al*., 2000). This phenomenon was attributed either to enzyme activity associated with a component of the phage tails that modifies the external features of the non-growing mycobacterial cells as phage particles diffuse from phage plaques, or to the activity of products encoded by genes *8* and *30*. *gp8* encodes the Bxb1 LysA protein and features the same gene arrangement of Ms6 *lysA* with two potential translation sites from which could result the synthesis of a

larger and a smaller endolysin protein. Based on these observations we investigated the contribution of $Lysin_{384}$ and $Lysin_{241}$ for Ms6 halo formation. An exponential growing culture of *M. smegmatis* was infected with the wild-type phage or the Ms6 *lysA*-mutant derivatives Ms6-$LysAHis_{6TGA}$ and Ms6-$LysAHis_{6GTG\rightarrow TGG}$ in the same experimental conditions; halo formation was observed beginning two days after incubation. We observed that both mutant phages form smaller and more turbid halos than the wild-type phage, suggesting that both $Lysin_{384}$ and $Lysin_{241}$ contribute for the "lysis from without" of the surrounding bacteria. However, we can not exclude the existence of other Ms6 proteins implicated in this phenomenon. Currently we do not know if Ms6 encodes Bxb1 *gp30* homologues, a gene coding for a putative protein with D-ala-D-ala carboxypeptidase activity (Mediavilla *et al.*, 2000). A gene (*gp19*) encoding a putative similar protein has been recently identified in mycobacteriophage Ardmore and shows strong identity to a number of genes encoded in mycobacteriophages included in subcluster F1. *gp19* could be predicted to encode a protein involved in host cell lysis although its location, quite far from the lysis cassette, the beta-lactamase superfamily domain and the linkage to the tail assembly function group do not support this hypothesis (Henry *et al.*, 2010). Remarkably, $Lysin_{384}$ and $Lysin_{241}$ enzymes showed lytic activity from without, against most of the mycobacteria tested (that included *M. smegmatis*, *M. vaccae*, *M. aurum* and *M. fortuitum*) in addition to several Gram-positive bacteria, contrary to what was previously thought (Payne *et al.*, 2009). Although *Mycobacterium* spp. are Gram-positive bacteria included in the suborder of *Corynebacterineae*, the envelope of this bacterial group is composed of a typical plasma membrane surrounded by a cell wall core, which, in turn, is surrounded by an outer membrane layer equivalent to the outer membrane of Gram-negative bacteria (Hoffman *et al.*, 2008; Zuber *et al.*, 2008). The cell wall consists of peptidoglycan covalently bound to arabinogalactan, which itself is covalently bound to mycolic acids (Daffé and Drapper, 1998; Zuber *et al.*, 2008). This

envelope is unusual in that it is very rich in lipids, and unlike other Gram-positive microorganisms, mycobacteria possess an outer permeability barrier that confers their characteristic resistance to solutes that includes many antibiotics and therapeutic agents (Trias and Benz, 1994; Brennan and Nikaido, 1995; Draper, 1998). The lower number of porins spanning the outer membrane compared to Gram-negative bacteria and the exceptional length of the pores are two major determinants of the low permeability of the outer membrane of *M. smegmatis* for hydrophilic solutes (Niederweis, 2003). Unexpectedly, however, we observed that different mycobacteria species are susceptible to exogenously added $Lysin_{384}$ and $Lysin_{241,}$ despite their mycolic-acid-rich outer membrane. However, when Ms6 LysB lipolytic enzyme was added from the outside it was unable to disrupt the outer membrane sufficiently to kill the mycobacteria (Gil *et al*., 2010). Even though it is unlikely that $Lysin_{384}$ or $Lysin_{241}$ can diffuse through water-filled channels, the porins (Nikaido, 1994; Niederweis, 2003), as typically only molecules with masses up to 600 Da can pass through the pores (Draper and Daffé, 2005), lysin access to the peptidoglycan may occur during cell division and septal peptidoglycan biogenesis. Exogenously applied phage-encoded endolysins have been shown to possess effective antimicrobial activity (Fischetti, 2008) against Gram-positive bacterial pathogens (Loeffler *et al*., 2001; Schuch *et al*., 2002; Loeffler and Fischetti, 2003; Loeffler *et al*., 2003).

Acknowledgments

We would like to thank Dr. Graham Hatfull, Dr. Julia van Kessel and Dr. Laura Marinelli (University of Pittsburgh, USA) for supplying plasmid pJV53 and for technical assistance with the recombineering experiments. This work was supported by funds provided by FCT (Fundação para a Ciência e Tecnologia) - PTDC/SAU-FCF/73017/2006. Maria João Catalão and Filipa Gil are the recipients of a fellowship from FCT (SFRH/BD/24452/2005) and (SFRH/BD/29167/2006).

References

Blackman S. A., Smith T. J. and Foster S. J. (1998) The role of autolysins during vegetative growth of *Bacillus subtilis* 168. *Microbiology* **144(Pt 1):**73-82.

Borysowski J., Weber-Dabrowska B. and Gorski A. (2006) Bacteriophage endolysins as a novel class of antibacterial agents. *Exp Biol Med* (Maywood) **231(4):**366-377.

Brennan P. J. and Nikaido H. (1995) The envelope of mycobacteria. *Annu Rev Biochem* **64:**29-63.

Bublitz M., Polle L., Holland C., Heinz D. W., Nimtz M. and Schubert W. D. (2009) Structural basis for autoinhibition and activation of Auto, a virulence-associated peptidoglycan hydrolase of *Listeria monocytogenes*. *Mol Microbiol* **71(6):**1509-1522.

Catalão M. J., Gil F., Moniz-Pereira J. and Pimentel M. (2010) A chaperone-like protein is involved in the endolysin delivery to the peptidoglycan. *Mol Microbiol* (In final revision).

Cheng Q., Nelson D., Zhu S. and Fischetti V. A. (2005) Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. *Antimicrob Agents Chemother* **49(1):**111-117.

Daffé M. and Draper P. (1998) The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* **39:**131-203.

Dale J. W. and Patki A. (1990) Mycobacterial gene expression and regulation. *In*, Molecular biology of mycobacteria. McFaden J. (Ed.) Surrey University Press, London, p.173-198.

Davison S., Couture-Tosi E., Candela T., Mock M. and Fouet A. (2005) Identification of the *Bacillus anthracis* (gamma) phage receptor. *J Bacteriol* **187(19):**6742-6749.

Draper P. (1998) The outer parts of the mycobacterial envelope as permeability barriers. *Front Biosci* **3:**D1253-1261.

Draper P. and Daffé M. (2005) The cell envelope of *Mycobacterium tuberculosis* with special reference to the capsule and outer permeability barrier. *In*, Tuberculosis and the tubercle bacillus. Stewart T. Cole *et al*. (Ed). ASM Press, Washington D. C., p. 261-272.

Fischetti V. A. (2008) Bacteriophage lysins as effective antibacterials. *Curr Opin Microbiol* **11(5):**393-400.

Frenken L. G., Bos J. W., Visser C., Müller W., Tommassen J. and Verrips C. T. (1993a) An accessory gene, *lipB*, required for the production of active *Pseudomonas glumae* lipase. *Mol Microbiol* **9(3):**579-589.

Frenken L. G., de Groot A., Tommassen J. and Verrips C. T. (1993b) Role of the *lipB* gene product in the folding of the secreted lipase of *Pseudomonas glumae*. *Mol Microbiol* **9(3):**591-599.

Garcia M., Pimentel M. and Moniz-Pereira J. (2002) Expression of mycobacteriophage Ms6 lysis genes is driven by two sigma (70)-like promoters and is dependent on a

transcription termination signal present in the leader RNA. *J Bacteriol* **184:**3034-3043.

Gil F., Catalão M. J., Moniz-Pereira J., Leandro P., McNeil M. and Pimentel M. (2008) The lytic cassette of mycobacteriophage Ms6 encodes an enzyme with lipolytic activity. *Microbiology* **154(Pt 5):**1364-1371.

Gil F., Grzegorzewicz A. E., Catalão M. J., Vital J., McNeil M. and Pimentel M. (2010) The mycobacteriophage Ms6 LysB specifically targets the outer membrane of *Mycobacterium smegmatis*. *Microbiology* **156:**1497-1504.

Gründling A., Manson M. D. and Young R. (2001) Holins kill without warning. *Proc Natl Acad Sci USA* **98:**9348-9352.

Hatfull G. F. (2005) Mycobacteriophages: pathogenesis and applications. *In*, Phages: their role in bacterial pathogenesis and biotechnology. Waldor M. K., Friedman D. I. and S. Adhya (Ed.) ASM Press, Washington D.C., p. 238-255.

Hatfull G. F., Pedulla M. L., Jacobs-Sera D., Cichon P. M., Foley A., Ford M. E., Gonda R. M., Houtz J. M., Hryckowian A. J., Kelchner V. A., Namburi S. Pajcini K. V., Popovich M. G., Schleicher D. T., Simanek B. Z., Smith A. L., Zdanowicz G. M., Kumar V., Peebles C. L., Jacobs W. R. Jr., Lawrence J. G. and Hendrix R. W. (2006) Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genetics* **2:**0835-0847.

Hatfull G. F., Jacobs-Sera D., Lawrence J. G., Pope W. H., Russell D. A., Ko C. C., Weber R.J., Patel M. C., Germane K. L., Edgar R. H., Hoyte N. N., Bowman C. A., Tantoco A. T., Paladin E. C., Myers M. S., Smith A. L., Grace M. S., Pham T. T., O'Brien M. B., Vogelsberger A. M., Hryckowian A. J., Wynalek J. L., Donis-Keller H., Bogel M. W., Peebles C. L., Cresawn S. G. and Hendrix R. W. (2010) Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. *J Mol Biol* **397(1):**119-143.

von Heijne G. and Abrahmsén L. (1989) Species-specific variation in signal peptide design. Implications for protein secretion in foreign hosts. *FEBS Lett* **244(2):**439-446.

Henry M., O´Sullivan O., Sleator R. D., Coffey A., Ross R. P., McAuliffe O., O´Mahony J. M. (2010) *In silico* analysis of Ardmore, a novel mycobacteriophage isolated from soil. *Gene* **453(1-2):**9-23.

Hobson A. H., Buckley C. M., Aamand J. L., Jørgensen S. T., Diderichsen B. and McConnell D. J. (1993) Activation of a bacterial lipase by its chaperone. *Proc Natl Acad Sci USA* **90(12):**5682-5686.

Hoffmann C., Leis A., Niederweis M., Plitzko J. M. and Engelhardt H. (2008) Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci USA* **105(10):**3963-3967.

Jolliffe L. K., Doyle R. J. and Streips U. N. (1981) The energized membrane and cellular autolysis in *Bacillus subtilis*. *Cell* **25(3):**753-763.

Kakikawa M., Yokoi K., Kimoto H., Nakano M., Kawasaki K., Taketo A. and Kodaira K. (2002) Molecular analysis of the lysis protein Lys encoded by *Lactobacillus plantarum* phage ϕg1e. *Gene* **299:**227-234.

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

Kok R. G., van Thor J. J., Nugteren-Roodzant I. M., Vosman B. and Hellingwerf K. J. (1995) Characterization of lipase-deficient mutants of *Acinetobacter calcoaceticus* BD413: identification of a periplasmic lipase chaperone essential for the production of extracellular lipase. *J Bacteriol* **177(11):**3295-3307.

Kumar V., Loganathan P., Sivaramakrishnan G., Kriakov J., Dusthakeer A., Subramanyam B., Chan J., Jacobs W. R. Jr. and Paranji Rama N. (2008) Characterization of temperate phage Che12 and construction of a new tool for diagnosis of tuberculosis. *Tuberculosis* (Edinb) **88(6):**616-623.

Laemmli U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227(5259):**680-685.

Loeffler, J. M., Nelson, D. and Fischetti V. A. (2001) Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell hydrolase. *Science* **294:**2170-2172.

Loeffler J. M. and Fischetti V. A. (2003) Synergistic lethal effect of a combination of phage lytic enzymes with different activities on penicillin-sensitive and -resistant *Streptococcus pneumoniae* strains. *Antimicrob Agents Chemother* **47(1):**375-377.

Loeffler J. M., Djurkovic S. and Fischetti V. A. (2003) Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia. *Infect Immun* **71(11):**6199-6204.

Loessner M. J. (2005) Bacteriophage endolysins: current state of research and applications. *Curr Opin Microbiol* **8:**480-487.

Marinelli L. J., Piuri M., Swigonová Z., Balachandran A., Oldfield L. M., van Kessel J. C. and Hatfull G. F. (2008) BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. *PLoS One* **3(12):**e3957.

Mediavilla J., Jain S., Kriakov J., Ford M. E., Duda R. L., Jacobs W. R. Jr, Hendrix R. W. and Hatfull G. F. (2000) Genome organization and characterization of mycobacteriophage Bxb1. *Mol Microbiol* **38(5):**955-970.

Nascimento J. G., Guerreiro-Pereira M. C., Costa S. F., São-José C. and Santos M. A. (2008) Nisin-triggered activity of Lys44, the secreted endolysin from *Oenococcus oeni* phage fOg44. *J Bacteriol* **190(1):**457-461.

Navarre W. W., Ton-That H., Faull K. F. and Schneewind O. (1999) Multiple enzymatic activities of the murein hydrolase from staphylococcal phage ϕ11. Identification of a D-alanyl-glycine endopeptidase activity. *J Biol Chem* **274(22):**15847-15856.

Niederweis M. (2003) Mycobacterial porins-new channel proteins in unique outer membranes. *Mol Microbiol* **49(5):**1167-1177.

Nikaido H. (1994) Porins and specific diffusion channels in bacterial outer membranes. *J Biol Chem* **269(6):**3905-3908.

Patton T. G., Yang S. J. and Bayles K. W. (2006) The role of proton motive force in expression of the *Staphylococcus aureus cid* and *irg* operons. *Mol Microbiol* **59(5):**1395-1404.

Payne K., Sun Q., Sacchettini J. and Hatfull G. F. (2009) Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. *Mol Microbiol* **73(3):**367-381.

Pedulla M. L., Ford M. E., Houtz J. M., Karthikeyan T., Wadsworth C., Lewis J. A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N. R., Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S., Kriakov J., Lawrence J. G., Jacobs W. R. Jr., Hendrix R. W. and Hatfull G. F. (2003) Origins of highly mosaic mycobacteriophage genomes. *Cell* **113:**171-182.

Piechota C., Gil F., Mahapatra S., Crick D. C., Yufang M., Huang H., Scherman M. S., Jones V., Pavelka M. S., Moniz-Pereira J., Pimentel M. and McNeil M. R. (2005) Mycobacteria phage Ms6 amidase reveals highly amidated mycobacterial peptidoglycan. Unpublished results.

Piuri M. and Hatfull G. F. (2006) A peptidoglycan hydrolase motif within the mycobacteriophage TM4 tape measure protein promotes efficient infection of stationary phase cells. *Mol Microbiol* **62(6):**1569-1585.

Portugal I., Anes E. and Moniz-Pereira J. (1989) Temperate mycobacteriophage from *M. smegmatis*. *Acta Leprol* **7:**243-244.

Pritchard D. G., Dong S., Baker J. R. and Engler J. A. (2004) The bifunctional peptidoglycan lysin of *Streptococcus agalactiae* bacteriophage B30. *Microbiology* **150(Pt 7):**2079-2087.

Recsei P. A., Gruss A. D. and Novick R. P. (1987) Cloning, sequence, and expression of the lysostaphin gene from *Staphylococcus simulans*. *Proc Natl Acad Sci USA* **84(5):**1127-1131.

Sambrook J. and Russell D.W. (2001). Molecular Cloning: a Laboratory Manual, 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Press.

São-José C., Parreira R., Vieira G. and Santos M. 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. *J Bacteriol* **178:**5823-5831.

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

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

Schuch, R., Nelson, D. and Fischetti V. A. (2002) A bacteriolytic agent that detects and kills *Bacillus anthracis*. *Nature* **418:**884-889.

Shearman C. A., Jury K. L. and Gasson M. J. (1994) Controlled expression and structural organization of a *Lactococcus lactis* bacteriophage lysin encoded by two overlapping genes. *Appl Environ Microbiol* **60(9):**3063-3073.

Smith T. J., Blackman S. A. and Foster S. J. (2000) Autolysins of *Bacillus subtilis*: multiple enzymes with multiple functions. *Microbiology*. **146(Pt 2):**249-262.

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

Son J. S., Jun S. Y., Kim E. B., Park J. E., Paik H. R., Yoon S. J., Kang S. H. and Choi Y. J. (2009) Complete genome sequence of a newly isolated lytic bacteriophage, EFAP-1 of *Enterococcus faecalis*, and antibacterial activity of its endolysin EFAL-1. *J Appl Microbiol* **108(5):**1769-1779.

Stirm S., Bessler W., Fehmel F. and Freund-Mölbert E. (1971) Bacteriophage particles with endo-glycosidase activity. *J Virol* **8(3):**343-346.

Sun Q., Kuty G. F., Arockiasamy A., Xu M., Young R. and Sacchettini J. C. (2009) Regulation of a muralytic enzyme by dynamic membrane topology. *Nat Struct Mol Biol* **16(11):**1192-1194.

Trias J. and Benz R. (1994) Permeability of the cell wall of *Mycobacterium smegmatis*. *Mol Microbiol* **14(2):**283-290.

Wang I.-N., Smith D. L. and Young R. (2000) Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* **54:**799-825.

Wiker H. G., Wilson M. A. and Schoolnik G. K. (2000) Extracytoplasmic proteins of *Mycobacterium tuberculosis* - mature secreted proteins often start with aspartic acid and proline. *Microbiology* **146(Pt 7):**1525-1533.

Xu M., Struck D. K., Deaton J., Wang I. and Young R. (2004) A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. *Proc Natl Acad Sci USA* **101:**6415-6420.

Xu M., Arulandu A., Struck D. K., Swanson S., Sacchettini J. C. and Young R. (2005) Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. *Science* **307(5706):**113-117.

Yokoi K. J., Kawahigashi N., Uchida M., Sugahara K., Shinohara M., Kawasaki K., Nakamura S., Taketo A. and Kodaira K. (2005) The two-component cell lysis genes *holWMY* and *lysWMY* of the *Staphylococcus warneri* M phage ϕWMY: cloning, sequencing, expression, and mutational analysis in *Escherichia coli*. *Gene* **351:**97-108.

Young R., Wang I.-N. and Roof W. D. (2000) Phages will out: strategies of host cell lysis. *Trends Microbiol* **8:**120-127.

Young R. (2002) Bacteriophage holins: deadly diversity. *J Mol Microbiol Biotechnol* **4:**21-36.

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

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

Zuber B., Chami M., Houssin C., Dubochet J., Griffiths G. and Daffé M. (2008) Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J Bacteriol* **190(16):**5672-5680.

5

Concluding Remarks

The results presented in this thesis further underscore the role of the mycobacteriophage Ms6 Gp1 accessory lysis protein during mycobacteria infection and allowed us to conclude that:

- Gp1 is a chaperone-like protein that interacts with the endolysin ($Lysin_{384}$) and participates in secretion to the extracytoplasmatic space in a holin-independent manner. The presence of this mycobacterium-specific lysis factor, may confer a selective advantage as an alternative to lysis exclusively holin-dependent.
- Hol is not essential for lysis and has structural features of a pinholin. A second holin-like protein is encoded by *gp5*, whose combined action with Hol is required to effect host cell lysis at the correct and programmed timing, although not essential for mycobacteria lysis.
- *lysA* encodes two endolysins ($Lysin_{384}$ and $Lysin_{241}$). Even though only $Lysin_{384}$ interacts with Gp1, both forms of LysA are required for the normal timing, progression and completion of host cell lysis. $Lysin_{241}$ results from the use of an internal, in-frame translation initiation site within the *lysA* gene.

This work represents a step forward in understanding the mechanisms involved in the molecular pathways of mycobacteriophage lysis. In this chapter, we integrate our new findings, discuss some questions and consider future perspectives in light of the recently discovered lysis models that include secretory endolysins endowed with signal peptides or signal-arrest-release (SAR) sequences and a new class of holins encoded by phages with SAR endolysins, pinholins, whose function is confined to membrane depolarization in order to fulfil their role in controlling the timing of lysis. To accomplish a successful infective cycle, dsDNA bacteriophages must lyse their host cells in order to liberate the progeny virions. The main barrier to host lysis is the continuous meshwork of peptidoglycan that allows the bacterial envelope to withstand internal osmotic pressure (Young and Wang, 2006). These viruses have evolved sophisticated lysis mechanisms involving the action of one protein, the holin, which imposes a strict temporal

program on the muralytic lytic enzyme, the endolysin assuring an optimum balance between the duration of the lytic cycle (which if delayed would compromise the opportunity to infect new hosts) and effective progeny yield (which would be too low if lysis was premature) (Young, 2002). Mycobacteriophages are phages which specifically target mycobacteria. The interest in these phages derives in large part from the medical significance and biological idiosyncrasies of their hosts as several mycobacterial species are important human and animal pathogens still causing major public health concerns, with the most notorious being *M. tuberculosis* and *Mycobacterium leprae*, the causative agents of tuberculosis and leprosy, respectively (Hatfull and Jacobs Jr., 1994; Hatfull, 2006). The mycobacteriophage Ms6 is a temperate dsDNA bacteriophage with an unusual lytic cassette: in addition to the endolysin-holin lysis system encoded by genes *lysA* (*gp2*) and *hol* (*gp4*), the Ms6 lytic cassette comprises three accessory lysis proteins encoded by genes *gp1*, *gp3* (*lysB*) and *gp5* (Garcia *et al*., 2002). Mycobacteria are Gram-positive bacteria that have evolved a complex cell wall, comprising a peptidoglycan-arabinogalactan polymer with covalently bound mycolic acids of considerable size, a variety of extractable lipids, and pore-forming proteins. The cell wall provides an extraordinary efficient permeability barrier to noxious compounds and contributes to the high intrinsic resistance of mycobacteria to many drugs (Brennan and Nikaido, 1995). To overcome the disadvantage that a complex cell wall may represents for a successful infective cycle, mycobacteriophages have evolved new lysis strategies by acquiring, through their evolution, genes that likely confer a substantial selective advantage over those without them, by providing faster and more complete lysis (Payne *et al*., 2009). A remarkably example is the existence, in 56 of the 60 completely sequenced mycobacteriophage genomes, of the *lysB* gene that encodes a mycolylarabinogalactan esterase that hydrolysis the mycolylarabinogalactan bond and compromises the stability of the mycobacterial outer membrane by severing its linkage to the underlying arabinogalactan-

peptidoglycan layer (Payne *et al*., 2009; Gil *et al*., 2010). In addition, several mycobacteriophages encode within its genome two accessory lysis genes (*gp1* and *gp5* Ms6 homologues) that are restricted to these group of phages (Hatfull *et al*., 2010), and are expected to encode mycobacterium-specific lysis factors. With this work we intended to unravel a potential role for Gp1 and its relationships with the additional Ms6 lysis proteins in the context of mycobacteria infection by mycobacteriophages. Several lines of evidence showed that *gp1*, the first gene of Ms6 lysis module, encodes a secretion chaperone-like protein that binds the endolysin, assists the export to the extracytoplasmic environment independently of holin function and is required to accomplish an efficient lysis of *M. smegmatis* and a productive burst size (Catalão *et al.*, 2010). It has already been described for some bacteriophages of both Gram-negative and Gram-positive hosts that endolysins are not always dependent on holins for export (São-José *et al*, 2000; Xu *et al.*, 2004; Xu *et al.*, 2005; Sun *et al*., 2009) and rely on secretory signals, N-terminal cleavable signal peptides or SAR domains, to target the peptidoglycan. Even though mycobacteriophage endolysins endowed with signal peptides or SAR domains have not yet been described, the physical association observed between *gp1* and *lysA* and represented by overlapping stop/start codons suggested a strong selective pressure for the linkage of these genes (Summer *et al*., 2007). Indeed, the N-terminal 60 amino acids residues of Ms6 LysA are necessary and sufficient for Gp1 binding and are essential for export across the cytoplasmic membrane of *E. coli* in a *sec*-dependent way. These are exciting findings as no other protein has ever been shown to interact with a lysin in any system, phage or otherwise, except for the transcriptional regulatory interaction between RNA polymerase and lysozyme of bacteriophage T7 (Moffatt and Studier, 1987), and presupposes an unprecedented molecular phenomenon, the activity of a specific chaperone for a Sec secreted protein. Mycobacteria encode two homologues of *secA* within its genome, *secA1* and *secA2*. Whereas *secA1* is an essential gene in *M. smegmatis*,

secA2 is essential neither for viability nor for export (Braustein *et al*., 2001). As deletion of *secA1* is a high lethal event for mycobacteria (Braustein *et al*., 2001), the direct involvement of *M. smegmatis* Sec*A1* in the translocation of LysA during Ms6 infection of the host cell remains elusive. Nevertheless, even though LysA lacks an N-terminal signal sequence or a transmembrane domain and, it is known that Sec-exported proteins are synthesized as precursors with a signal sequence at the N-terminus (Rigel and Braunstein, 2008), Gp1 may be a potential candidate to engage the Sec system. Further investigation is needed to understand how does the complex Gp1-LysA interacts with the Sec system for secretion. We believe that requirement of Gp1 to obtain an efficient lysis of *M. smegmatis* results from the lack of ability of the Ms6 holin to provide an efficient way for endolysin to reach the peptidoglycan. Indeed, examination of the predicted protein sequence of Hol showed that the first TMD has characteristics of a SAR domain with a high percentage of hydrophobic or polar residues. The presence of a SAR domain followed by a typical TMD suggests that Ms6 Hol is a pinholin (Park *et al*., 2007; Summer *et al*., 2010) and needs only to depolarize the membrane in order to fulfil its role in controlling the time of lysis. Indeed, an Ms6 derivative mutant defective for Hol synthesis was still able to infect and lyse *M. smegmatis*, although with reduced latent period and burst size. In addition to Ms6 Hol, the accessory lysis protein Gp5 also presented characteristics of holins with an N-terminal TMD and a very highly charged and hydrophilic C-terminal domain. Determination of infection parameters of the holin genes-deleted mutant phages showed that although nonessential for phage viability the concerted action of these two proteins, Hol and Gp5, could play the role of a holin. Moreover, expression of both proteins is necessary to effect host cell lysis at the correct and programmed timing and their interaction may contribute to very precise adjustment of the timing of hole formation in order to keep the infected cell productive, allowing the assembly of more virions. Remarkably, in Ms6 (and conceivably, related mycobacteriophages)

the presence of the endolysin in addition to one of the lytic genes, *gp1*, *hol* or *gp5*, is sufficient for a lysis phenotype. However, this results in dramatic changes in phage fitness during the infective cycle and in lower viability of the mutant phages. The presence of the mycobacterium-specific lysis factors Gp1 and Gp5 may confer a selective advantage not only for fitness under different conditions but also as an alternative to lysis exclusively holin-dependent: it has been shown that single missense changes within the holin proteins can have a profound effect on both the process of host lysis and its timing, unpredictably advancing or retarding the instant of triggering (Young, 2002). Notwithstanding, the exact mechanism by which the lysis effectors Lysin A and Lysin B are localized to their substrates remains elusive in many mycobacteriophages; additional holin genes have yet to be identified bioinformatically or by functional studies as in the majority of the mycobacteriophage genomes sequenced there is no closely linked and easily recognizable holin gene (Payne *et al*., 2009; Hatfull *et al*., 2010). While there are few gene Phams represented in all 60 completely sequenced mycobacteriophage genomes, one of these is the *lysA* phamily (Pham66) (Payne *et al*., 2009; Hatfull *et al*., 2010). Examination of the Ms6 *lysA* DNA sequence revealed the existence of a second possible lysin gene ($lysin_{241}$) overlapped with *lysA* ($lysin_{384}$) in the same reading frame and preceded by a consensus Shine and Dalgarno sequence. While this gene arrangement was found to be uncommon among phage endolysins, a BLASTn search for Ms6 *lysA* homologues revealed that this peculiar endolysin gene arrangement is widespread in mycobacteriophages, in particular among those that possess a *gp1*-like gene (Pham1480) overlapped or in close proximity to *lysA*. Interestingly, however, two putative translational initiation signals were also identified in endolysins belonging to four mycobacteriophages that do not possess Gp1 homologues (Phyler, Phaedrus, Pipefish and Corndog) but that possess N-terminal related Ms6 LysA sequences. The lack of representation of Pham1480 upstream of *lysA* in these phages could result from loss or lack of acquisition of

gp1-like genes in these genomes. In addition, in three mycobacteriophages that possess Gp1 similar proteins but unrelated Ms6 LysA enzyme (TM4, Jasper and Lockley) this *lysA* gene arrangement was not observed which suggests that Pham1480 in these mycobacteriophages results from recent acquisition by horizontal genetic exchange (Hatfull *et al*., 2010) and may confer a selective advantage for host cell lysis under different environmental conditions. $Lysin_{384}$ can be produced in *E. coli* when multiple copies of the *lysA* gene are present under the control of a strong promoter and accumulates intracellularly in an active conformation even in the absence of a functional *gp1* gene which suggests that Gp1 is not a specific foldase. However, during *M. smegmatis* infection, synthesis and/or stability of the larger endolysin ($Lysin_{384}$) were found to be highly dependent on Gp1 translation. This suggests that Gp1 is not involved in transcription regulation but in a post-transcriptional event, associating transiently for stabilization of unfolded or partially folded endolysin intermediates thereby preventing misfolding or aggregation due to improper or premature intra- or interchain interactions that would occur in the absence of the chaperone. In addition, Gp1 chaperone activity may help maintaining $Lysin_{384}$ in a secretion-competent state; however, it is unlikely that it participates in $Lysin_{384}$ final function. Suppression of $Lysin_{384}$ or $Lysin_{241}$ synthesis does not result in a non-lysis phage phenotype. However, $Lysin_{384}$ is necessary to achieve a normal burst of infective phages whereas $Lysin_{241}$ has an important function in the progression and complete host cell lysis. We propose a model where $Lysin_{384}$ is continuously secreted to the periplasm through N-terminal binding to Gp1 secretion chaperone during phage assembly in order to guarantee rapid cell lysis whereas $Lysin_{241}$ acts at a later stage of infection (when lysis is triggered by the concerted action of Hol and Gp5) and only once an adequate number of progeny virions is reached intracellularly, allowing completion of host cell lysis (Fig. 1).

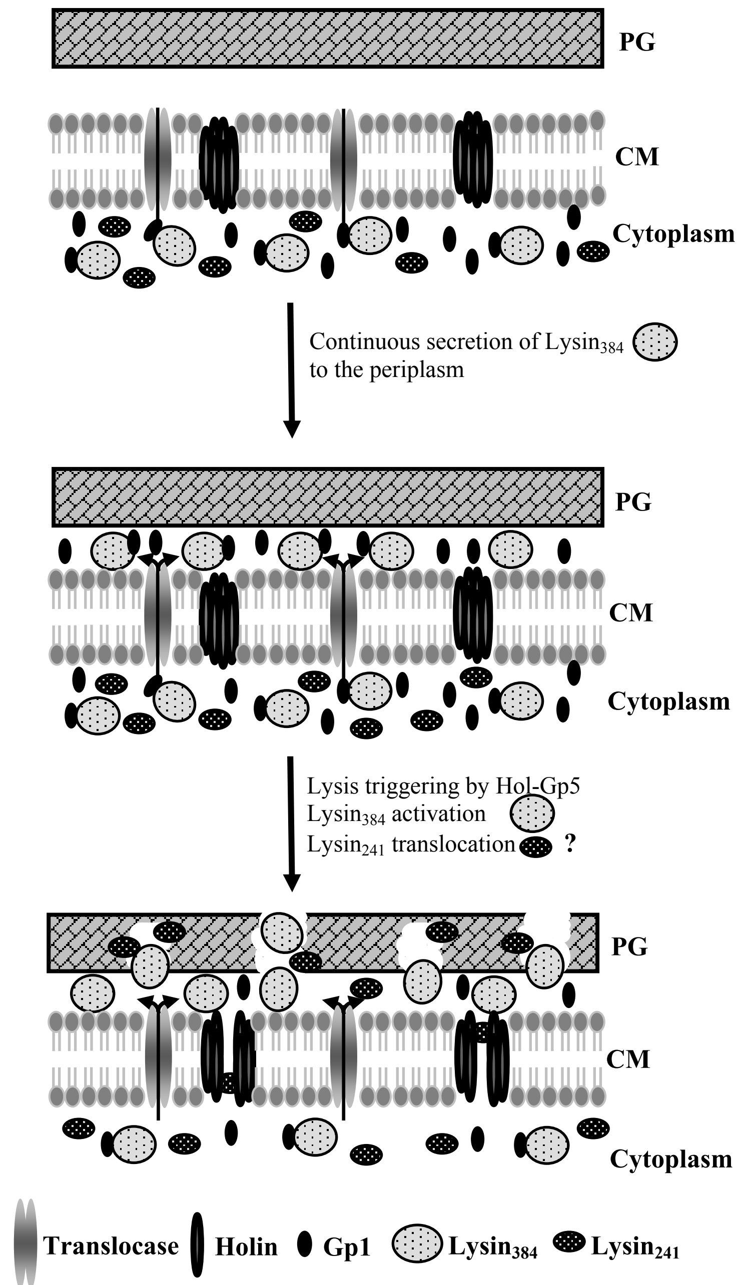


Figure 1. ***Proposed model for mycobacteriophage Ms6 lysis of mycobacteria.*** During

phage assembly, Lysin$_{384}$ N-terminal binding to Gp1 secretion-chaperone allows continuous secretion to the periplasm where it presumably remains inactive by an unknown mechanism. Triggering of host cell lysis by the concerted action of Hol and Gp5 activates Lysin$_{384}$ and allows access of Lysin$_{241}$ to the periplasm for a rapid and efficient completion of host cell lysis. See text for more complete description. **PG**-Peptidoglycan; **CM**-cytoplasmic membrane.

At this time it is unknown how Lysin$_{384}$ is activated. The secreted Ms6 Lysin$_{384}$ must be under some extracytoplasmic regulatory mechanism that ensures that premature lysis does not take place; nisin triggering of *M. smegmatis* host cell lysis after Gp1 and LysA induction suggests that the energized state of the cytoplasmic membrane is implicated in LysA activation as described for other secretory endolysins (São-José *et al.*, 2000; Xu *et al.*, 2004; Nascimento *et al.*, 2008) or autolysins (Jolliffe *et al.*, 1981; Blackman *et al.*, 1998; Smith *et al.*, 2000; Patton *et al.*, 2006). This regulation may also be dependent on Gp1 binding to the N-terminal domain of Lysin$_{384}$ which may block substrate binding, or Gp1 may allow Lysin$_{384}$ to adopt an active conformation. Some chaperones were already described as necessary to mask domains on the substrates that are required for their functions or interactions and that could otherwise interfere with the process of translocation, cause aggregation or be toxic for the bacterial cells (Lund, 2001; Francis *et al.*, 2002; Page and Parsot, 2002; Feldman and Cornelis, 2003).

Despite the fact that *M. smegmatis* possesses an outer membrane, several mycobacteria were shown to be susceptible to exogenously added Lysin$_{384}$ and Lysin$_{241}$ enzymes in contrast to what should be expected (Payne *et al.*, 2009) due to the presence of a mycolic-acid-rich outer permeability barrier that confers their characteristic resistance to hydrophilic solutes that includes many antibiotics and therapeutic agents (Trias and Benz, 1994; Brennan and Nikaido, 1995; Draper, 1998) suggesting that lysin access to the peptidoglycan may occur in growing cells during division. These data is in agreement with the ability of Ms6 to form turbid

plaques surrounded by a clear zone of apparent bacterial lysis. This phenomenon is widely observed among bacteriophages that infect Gram-positive hosts and results from “lysis from without” of bacteria as a result of endolysin diffusion from phage plaques that kills uninfected cells (Loessner, 2005).

As a final note, even though the exact mechanism involved in Ms6 endolysin translocation in mycobacteria is still elusive, with this thesis we hope to elucidate and contribute to a better understanding of the mechanisms of mycobacteriophage lysis of mycobacteria. The ultimate challenge will be to characterize the exact pathways by which the mycobacteriophage lysis effectors are positioned next to their targets which may result in the identification of mycobacteria new cell wall targets and in the potential therapeutic use of mycobacteriophage lysins.

References

Blackman S. A., Smith T. J. and Foster S. J. (1998) The role of autolysins during vegetative growth of *Bacillus subtilis* 168. *Microbiology* **144(Pt 1):**73-82.

Braunstein M., Brown A. M., Kurtz S., and Jacobs W. R. Jr. (2001) Two nonredundant SecA homologues function in Mycobacteria. *J Bacteriol* **183:**6979-6990.

Brennan P. J. and Nikaido H. (1995) The envelope of mycobacteria. *Annu Rev Biochem* **64:**29-63.

Catalão M. J., Gil F., Moniz-Pereira J. and Pimentel M. (2010) A chaperone-like protein is involved in the endolysin delivery to the peptidoglycan. *Mol Microbiol* (In final revision).

Draper P. (1998) The outer parts of the mycobacterial envelope as permeability barriers. *Front Biosci* **3:**D1253-1261.

Feldman M. F. and Cornelis G. R. (2003) The multitalented type III chaperones: all you can do with 15 kDa. *FEMS Microbiol Letters* **219:**151-158.

Francis M. S., Wolf-Watz H. and Forsberg A. (2002) Regulation of type III secretion systems. *Curr Opin Microbiol* **5:**166-172.

Garcia M., Pimentel M. and Moniz-Pereira J. (2002) Expression of mycobacteriophage Ms6 lysis genes is driven by two Sigma (70)-like promoters and is dependent on a transcription termination signal present in the leader RNA. *J Bacteriol* **184:**3034-3043.

Gil F., Grzegorzewicz A. E., Catalão M. J., Vital J., McNeil M. and Pimentel M. (2010) The mycobacteriophage Ms6 LysB specifically targets the outer membrane of *Mycobacterium smegmatis*. *Microbiology* **156:**1497-1504.

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

Hatfull G. F. (2006) Mycobacteriophages. *In*, The bacteriophages. Calendar R. (Ed.), 2nd ed. Oxford University Press, New York, p. 602-620.

Hatfull G. F., Jacobs-Sera D., Lawrence J. G., Pope W. H., Russell D. A., Ko C. C., Weber R. J., Patel M. C., Germane K. L., Edgar R. H., Hoyte N. N., Bowman C. A., Tantoco A. T., Paladin E. C., Myers M. S., Smith A. L., Grace M. S., Pham T. T., O'Brien M. B., Vogelsberger A. M., Hryckowian A. J., Wynalek J. L., Donis-Keller H., Bogel M. W., Peebles C. L., Cresawn S. G. and Hendrix R. W. (2010) Comparative Genomic Analysis of 60 Mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. *J Mol Biol* **397(1):**119-143.

Jolliffe L. K., Doyle R. J. and Streips U. N. (1981) The energized membrane and cellular autolysis in *Bacillus subtilis*. *Cell* **25(3):**753-763.

Loessner M. J. (2005) Bacteriophage endolysins: current state of research and applications. *Curr Opin Microbiol* **8:**480-487.

Lund P. A. (2001) Microbial Molecular Chaperones. *Advances Microb Physiol* **44:**93-140.

Moffatt B. A. and Studier F. W. (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49(2):**221-227.

Nascimento J. G., Guerreiro-Pereira M. C., Costa S. F., São-José C. and Santos M. A. (2008) Nisin-triggered activity of Lys44, the secreted endolysin from *Oenococcus oeni* phage fOg44. *J Bacteriol* **190:**457-461.

Page A. and Parsot C. (2002) Chaperones of the type III secretion pathway: jacks of all trades. *Mol Microbiol* **46:**1-11.

Park T., Struck D. K., Dankenbring C. A. and Young R. (2007) The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. *J Bacteriol* **189(24):**9135-9139.

Patton T. G., Yang S. J. and Bayles K. W. (2006) The role of proton motive force in expression of the *Staphylococcus aureus cid* and *irg* operons. *Mol Microbiol* **59(5):**1395-1404.

Payne K., Sun Q., Sacchettini J. and Hatfull G. F. (2009) Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. *Mol Microbiol* **73(3):**367-381.

Rigel N. W. and Braunstein M. (2008) A new twist on an old pathway - acessory Sec systems. *Mol Microbiol* **69:**291-302.

São-José C., Parreira R., Vieira G. and Santos M. 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. *J Bacteriol* **178:**5823-5831.

Smith T. J., Blackman S. A. and Foster S. J. (2000) Autolysins of *Bacillus subtilis*: multiple enzymes with multiple functions. *Microbiology* **146(Pt 2):**249-262.

Summer E. J., Berry J., Tran T. A., Niu L., Struck D. K. and Young R. (2007) *Rz/Rz1* lysis gene equivalents in phages of Gram-negative hosts. *J Mol Biol* **373(5):**1098-1112.

Summer E. J., Enderle C. J., Ahern S. J., Gill J. J., Torres C. P., Appel D. N., Black M. C., Young R. and Gonzalez C. F. (2010) Genomic and biological analysis of phage Xfas53 and related prophages of *Xylella fastidiosa*. *J Bacteriol* **192(1):**179-190.

Sun Q., Kuty G. F., Arockiasamy A., Xu M., Young R. and Sacchettini J. C. (2009) Regulation of a muralytic enzyme by dynamic membrane topology. *Nat Struct Mol Biol* **16(11):**1192-1194.

Trias J. and Benz R. (1994) Permeability of the cell wall of *Mycobacterium smegmatis*. *Mol Microbiol* **14(2):**283-290.

Xu M., Struck D. K., Deaton J., Wang I. and Young R. (2004) A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. *Proc Natl Acad Sci USA* **101:**6415-6420.

Xu M., Arulandu A., Struck D. K., Swanson S., Sacchettini J. C. and Young R. (2005) Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. *Science* **307(5706):**113-117.

Young R. (2002) Bacteriophage holins: deadly diversity. *J Mol Microbiol Biotechnol* **4:**21-36.

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