Genetics of Phage Lysis

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

ABSTRACT We have been witnessing an increased interest in bacteriophage studies focused on their use as antibacterial agents to fight pathogenic bacteria. This interest is a consequence of the phages’ ability to lyse a bacterial host. Until recently, little was known about the mechanisms used by mycobacteriophages to induce lysis of their complex hosts. However, studies on Ms6-induced lysis have changed this scenario and provided new insights into the mechanisms of bacteriophage-induced lysis. Specific lysis protein genes have been identified in mycobacteriophage genomes, reflecting the particular mycobacterial cell envelope composition. These include enzymes that target mycolic acid–containing lipids and proteins that participate in the secretion of the phage endolysin, functioning as chaperone-like proteins. This chapter focuses on the current knowledge of mycobacteriophage-induced lysis, starting with an overview of phage lysis and basic features of the lysis players.

Bacteriophages, or simply phages—the viruses that infect bacteria—are the most abundant biological entities on earth, playing a fundamental role in bacterial ecology and evolution (1, 2). To survive they need to infect sensitive bacteria, where they replicate and produce new viral particles. An infectious cycle starts with adsorption of the phage particle to the surface of a specific bacterial host, followed by penetration of the phage genome into the cytoplasm. Once inside the host cell, the genetic information carried in the viral genome is responsible for its own replication and for the synthesis of the components to make new phage particles. These newly assembled progeny virions now need to be released into the environment where host bacteria are potentially available for new infection cycles. Except for filamentous bacteriophages, which are released from their hosts without affecting the cell viability (3), all other phages must lyse the infected bacteria to liberate the virion progeny to the extracellular milieu (4, 5).

Bacterial lysis is thus the final event of a phage lytic cycle. To accomplish this step phages have to overcome the bacterial cell barriers, especially the peptidoglycan (PG) layer, a rigid and stable structure that allows the bacterial envelope to support internal osmotic pressure. For this, phages may use two basic strategies. Phages with single-stranded small genomes, exemplified by phages Qβ (ssRNA) and φX172 (ssDNA), synthesize a single lysis protein termed “amurin,” which causes lysis by interfering with PG biosynthesis (4, 6, 7). Double-stranded DNA phages, which represent more than 95% of known bacterial viruses (8), use the holin-endolysin strategy, employing at least two essential proteins whose coordinated action results in well-timed and swift host cell lysis. Endolysins are enzymes that target the integrity of the PG layer, while holins are small membrane proteins that control the activation or its access to the murein at a precisely defined time (4, 5).

The mechanism of lysis performed by bacteriophage λ is by far the most well studied and was, for many years, considered to be universal (4). However, it is now becoming increasingly evident that phages of both Gram-positive and Gram-negative bacteria employ diverse lysis models, reflecting adaptations to their hosts. A remarkable example of such diversity is given by mycobacteriophages, which exhibit different lysis cassettes.
These phages infect mycobacterial hosts, bacteria that possess a complex cell envelope with a cell wall core composed of PG covalently linked to arabinogalactan, which is in turn linked to mycolic acids (9). To overcome the disadvantage that this complex envelope represents for a successful infective cycle, these viruses have evolved new lysis systems by acquiring specific lysis genes that confer a substantial selective advantage providing optimal lysis timing in response to environmental factors.

This chapter will focus on the current knowledge of mycobacteriophage-mediated lysis. An overview of bacteriophage lysis will first be presented to introduce the reader to the subject. Recent reviews discuss phage lysis in other systems (4, 5, 10).

THE LYSIS PLAYERS

Two major aspects are considered for successful bacteriophage-induced lysis: (i) compromise the PG layer and (ii) ensure the proper time of lysis. Determination of the precise time of lysis is achieved by holins. These are small membrane proteins with at least one predicted transmembrane domain (TMD) and a hydrophilic, highly charged C-terminus. According to the number of TMDs (three, two, or one), holins can be classified in classes I, II, or III, respectively (11, 12). Holins accumulate in the bacterial cytoplasmic membrane (CM) during late gene expression without disturbing membrane integrity (4, 13–16). Then at a precise, allele-specific time, the holin triggers to form holes that permeabilize the membrane, resulting in its disruption, and lysis begins (11, 12). According to the size of the holin-mediated hole, holins can be classified in two types: canonical holins (e.g., λ S105), which form large holes of near-micrometer scale (17), and pinholins (e.g., S2168 of phage 21), which form heptameric channels of ∼2 nm (18). The timing of lysis must be tightly regulated, since early lysis can release few or no phage particles, resulting in an unprofitable infection cycle, while a delay might compromise the availability of new bacterial hosts in the environment. Some phages also encode an antiholin, a negative regulator that controls the functions of all three classes of holins contributing to the precise regulation of lysis (11). In some cases the antiholin is encoded by the holin gene (19), while in others it is encoded by an independent gene (20–22).

Endolysins are PG hydrolases that are produced during the late phase of gene expression and are designed to cleave one or more of the covalent bonds in the PG meshwork, compromising the integrity of the cell wall. They can be classified into five major types (Fig. 1) depending on the specific bonds that are attacked: (i) N-acetylmuramidases (generically termed lysozymes) and (ii) lytic transglycosylases, both cleaving the β-1,4 glycosidic bond between N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG); (iii) endo-β-N-acetylgalactosaminidases, cleaving the other glycosidic bond between NAG and NAM; (iv) N-acetyl-muramoyl-l-alanine amidases, which hydrolyze the amide bond between NAM and l-alanine residues in the oligopeptide cross-linking chains, and (v) endopeptidases, which attack the peptide bonds in the same chains (23). Endopeptidases can be further divided into two groups: enzymes that cut within the stem peptide of the PG and those that attack bonds within the interpeptide bridge (if present) (24). Most endolysins from phages that infect Gram-negative bacteria are relatively small globular proteins composed of only a single catalytic domain, while endolysins from Gram-positive-infecting phages are typically modular, having one or more catalytic domains at the N-terminus and a cell wall–binding domain at the C-terminus (23, 24). These structural differences reflect the variations in cell wall architecture between these major bacterial groups (for an extensive review on bacteriophage endolysins see references 23–26).

The access of endolysins to the PG may occur in two ways: canonical endolysins, such as λR, accumulate fully folded and active in the cytoplasm during late gene expression, until the holin forms holes in the CM large enough to allow the passage of the endolysin through them; secreted endolysins are exported as inactive enzymes in a holin-independent manner, using the host Sec system. Two types of secreted endolysins have been described: (i) endolysins with a Sec-dependent SP sequence, which are processed by the leader peptidase during translocation of the enzyme to the cell wall compartment, e.g., Lys44 of Oenococcus phage f0g44 (27), and (ii) endolysins endowed with a signal arrest release (SAR) sequence, which is not removed during translocation, retaining the endolysins in the periplasm as inactive forms tethered to the CM, e.g., R21 of phage 21 and Lys of phage P1 (28, 29). Although in the former case it is not known how endolysins are activated, in the latter, activation is a consequence of the membrane depolarization by either a canonical holin or a pinholin, which allows the escape of the SAR endolysin from the bilayer and refolding into an active form, ready to attack the PG (30–32).

In addition to the essential lysis proteins, phages infecting Gram-negative hosts encode a third functional class of lysis proteins, the spanins. These proteins span the entire periplasm, allowing the fusion of the inner
membrane with the outer membrane (OM) at a third stage of host lysis, where the first stage is disruption of the CM by the holin, followed by degradation of the cell wall by the endolysin. Spanins, which may be synthesized as a sole protein (T1 Gp11) or as two subunits (λ Rz and Rz1 proteins), are responsible for elimination of the third and final barrier to phage release, the OM (33–36).

MYCOBACTERIOPHAGE-MEDIATED LYSIS
Mycobacteriophages are a remarkably diverse group of viruses that infect mycobacteria. Currently, more than 3,400 mycobacteriophages have been isolated, with more than 270 complete genome sequences available in GenBank (http://www.phagesdb.org). Most of them have Mycobacterium smegmatis as their host, and they all share a double-stranded DNA genome (37–39). As dsDNA-tailed phages, mycobacteriophages have to overcome the complex envelope of mycobacteria to achieve lysis of their hosts and release the progeny virions at the end of a lytic cycle. Until recently, little was known about the mechanisms underlying mycobacteriophage-induced lysis of mycobacteria, but studies on mycobacteriophage Ms6 (40–45) provided new insights into the way phages achieve lysis of their hosts. The first report came from the work of Garcia et al. (46), who described the genetic organization of the lysis module of mycobacteriophage Ms6. A detailed analysis of the Ms6 lysis model is described below.
The Lysis Model of Mycobacteriophage Ms6

Mycobacteriophage Ms6 is a temperate phage, isolated from *M. smegmatis* strain HB5688 (47). The Ms6 lysis cassette is composed of five genes (Fig. 2) clustered downstream of two σ^70-like promoters. This promoter region (P_lys) is separated from the first lysis gene by a leader sequence of 214 bp, in which was detected a transcription termination signal, suggesting that an antitermination mechanism is involved in the regulation of Ms6 lysis gene transcription (46). Although the complete genome sequence is not yet available, according to the organization of the lysis genes, Ms6 seems to be closely related to phages included in cluster F, subcluster F1 (10, 37). Like all dsDNA phages, Ms6 uses the holin-endolysin strategy to achieve lysis of its host; however, the model of lysis is different from those described for other bacteriophages. Perhaps one of the most striking features of Ms6 is that, in addition to encoding a specialized function related to the nature of the mycobacteria cell envelope, it displays a peculiar mechanism of endolysin export. The access of the Ms6 endolysin (LysA) to the PG was shown to be independent of the holin function; however, LysA does not have a predicted Sec-type SP or a SAR sequence, as do the secreted endolysins described above.

Catalão et al. (42) have shown that LysA export is assisted by a chaperone-like protein (Gp1) encoded by the first gene of the Ms6 lysis cassette. Analysis of the physical and structural properties of Gp1 shows that it shares the properties of molecular chaperones, particularly type III secretion system (TTS) chaperones: gp1, positioned immediately upstream of the endolysin gene and overlapped with it, encodes a small protein of 8.3 kDa with an acidic isoelectric point of 4.6. In common with molecular chaperones, Gp1 has the ability to oligomerize and was shown to interact with its effector, an interaction that encompasses the N-terminal region of the chaperone and the first 60 amino acids of the Ms6 endolysin (42, 45). Evidence for a role in LysA translocation comes from experiments in *M. smegmatis* that showed an increased alkaline phosphatase activity of a LysA-PhoA′ hybrid protein only in the presence of Gp1 (42). Doubts about the in vivo role of Gp1 in lysis were resolved when deletion of gp1 from the Ms6 genome resulted in alterations in the phage lysis phenotype. Although not essential for plaque formation, and thus not essential for lysis, Gp1 is necessary to achieve efficient lysis, since its absence resulted in a decrease of ∼70% in the burst size (42).

Interestingly, and although no secretion signals were predicted in the amino acid sequence of Ms6 LysA, experiments in *Escherichia coli*, using either a thermosensitive secA strain or the SecA inhibitor sodium azide, showed that lysis is blocked in cells expressing Gp1 and LysA, indicating that the *E. coli* sec machinery is involved in Ms6 LysA translocation, as it is with the secretory endolysins described so far (27, 28, 42). However, the involvement of the mycobacteria sec system in Ms6 LysA delivery was not yet determined. The Sec-dependent export pathway is highly conserved among different bacteria. Mycobacteria have two homologues of secA: secA1 and secA2 (48); secA1 is essential in *M. smegmatis*, while secA2 is nonessential for viability. Depletion of secA1 leads to a loss of cell viability. Consequently, the attempts made so far to evaluate the dependence on the *M. smegmatis* SecA1 for Ms6-induced lysis were not successful, so the direct involvement of SecA1 in Ms6 LysA translocation remains to be demonstrated. Although the dependence of SecA2 for export of some proteins has been described (49), infection of an *M. smegmatis* ΔsecA2 mutant strain with Ms6 wild type did not alter the phage plaque ability (42).

Ms6 LysA has a central PG recognition protein conserved domain localized between amino acid residues

**FIGURE 2** Genetic organization of the Ms6 lysis cassette. Genes are drawn to scale with gene names indicated. Segments of *lysA* generating the full-length Lysin_{384} and the N-terminal truncated version Lysin_{241} are indicated separately. The promoter region P_lys is separated from *gp1* by a leader sequence (L). The arrow indicates direction of the transcription from the promoter region P_lys. † indicates the localization of a transcription termination signal. Adapted from reference 46 with permission. doi:10.1128/microbiolspec.MGM2-0017-2013.f2
168 and 312, containing an amidase-2 domain (pfam01510) (44). Its N-acetyl-L-muramoyl-L-alanine amidase activity was recently reported; LysA was shown to cleave the bond between 1-Ala of the stem peptide and the lactyl moiety of the muramic acid residues of muramyl pentapeptide and to release up to 70% of the diaminopimelic acid present in the isolated mycobacterial cell wall (50). Interestingly, a recent analysis of the lysA gene shows that it generates two proteins designated Lysin384 and Lysin241 according to the size of the produced polypeptides. Catalão et al. (44) have shown that Lysin241 is not a mature form of Lysin384, but rather is a result of a second translation event from an initiation codon, in the same reading frame, positioned 143 amino acids away from the first initiation codon. Not surprisingly, both proteins have PG hydrolase activity, since Lysin241 retains the PG recognition protein domain.

What is puzzling is that both forms of LysA are necessary for complete and efficient lysis of M. smegmatis. Although deletion of the complete lysA nucleotide sequence revealed it to be essential for M. smegmatis lysis, Ms6 mutants producing only one of the forms of LysA were shown to be viable, albeit defective in the normal timing, progression, and completion of host cell lysis. Lack of Lysin384 resulted in a lysis delay of 30 min and in a reduction in the number of phage particles released, while in the absence of Lysin241, lysis starts 90 min later with no significant effect on the number of phage particles released (44). Worthy of note is the fact that although Lysin241 keeps the enzyme catalytic domain, the N-terminal region that interacts with Gp1 is absent. However, a tight association between Gp1 and Lysin384 exists, since Catalão et al. observed that during an Ms6 derivative mutant, where the two forms of endolysin are expressed with a hexastidine tag at the C-terminus, Ms6 mutants producing only one of the forms of LysA complement a λS defective mutant (44). A more recent characterization of Ms6 Gp4 function suggested that Gp4 might function as a pinholin, since the first TMD of Gp4 shares structural characteristics with class II holins. A gene encoding a small protein of 77 amino acids, Gp4, is present in MS6, and its nucleotide sequence is identical to that of the E. coli amino acid residue of mu-ramyl pentapeptide and to release up to 70% of the diaminopimelic acid present in the isolated mycobacterial cell wall (50). Interestingly, a recent analysis of the lysA gene shows that it generates two proteins designated Lysin384 and Lysin241 according to the size of the produced polypeptides. Catalão et al. (44) have shown that Lysin241 is not a mature form of Lysin384, but rather is a result of a second translation event from an initiation codon, in the same reading frame, positioned 143 amino acids away from the first initiation codon. Not surprisingly, both proteins have PG hydrolase activity, since Lysin241 retains the PG recognition protein domain.

What is puzzling is that both forms of LysA are necessary for complete and efficient lysis of M. smegmatis. Although deletion of the complete lysA nucleotide sequence revealed it to be essential for M. smegmatis lysis, Ms6 mutants producing only one of the forms of LysA were shown to be viable, albeit defective in the normal timing, progression, and completion of host cell lysis. Lack of Lysin384 resulted in a lysis delay of 30 min and in a reduction in the number of phage particles released, while in the absence of Lysin241, lysis starts 90 min later with no significant effect on the number of phage particles released (44). Worthy of note is the fact that although Lysin241 keeps the enzyme catalytic domain, the N-terminal region that interacts with Gp1 is absent. However, a tight association between Gp1 and Lysin384 exists, since Catalão et al. observed that during an Ms6 infection, the synthesis/stability of the larger endolysin is dependent on Gp1 production (44).

Despite the fact that both proteins are necessary for efficient lysis, Lysin241 seems to be much more active than Lysin384, as revealed by zymogram assays. Contrary to what is observed in an M. smegmatis infection with an Ms6 derivative mutant, where the two forms of endolysin are expressed with a hexastidine tag at the C-terminus, in E. coli, expression from the full lysA gene results in almost undetectable levels of Lysin241. However, even a low level of Lysin241 showed high hydrolytic activity on Micrococcus luteus lyophilized cells (44). The same observation was reported for the endolysin produced by the mycobacteriophage Corndog (51). The authors suggested that the N-terminus, present in the larger form of the endolysin, could somehow inhibit catalysis in the zymogram assay. Interestingly, E. coli crude extracts containing Lysin384 or Lysin241 were both shown to inhibit the growth of several bacteria “from without” (44). Although it is clear that Gp1 is essential for translocation of Lysin384, it is not known at this time how Lysin241 accesses the PG or why Ms6 synthesizes two endolysins.

Even though Ms6 Lysin384 is exported in a holin-independent manner, lysis of M. smegmatis does not occur until holin triggering. It is now evident that secreted endolysins, once positioned in the extracytoplasmic environment, must be kept inactive until the proper time of lysis (5, 31, 32). In fact, compromising the murein layer at a time well before the new phage particles are assembled would result in no or too little progeny release. Thus, even for phages encoding secreted endolysins, holins are still crucial for determining the timing of lysis. Indeed, all phages that encode secreted endolysins also appear to encode a holin-like protein (5). Holin function is thus confined to controlling the access of the endolysin to its target, either by allowing its passage through holes formed in the CM—as happens in the "class II" model of lysis (5)—or by allowing the activation of endolysin already positioned close to the murein layer at the time of dissipation of the membrane potential—exemplified by the P1 model of lysis (18, 28, 30).

In Ms6, the potential role of the holin function in endolysin activation was supported using nisin, a permeabilizing compound that triggers CM depolarization. In contrast to what happens with M. smegmatis cells expressing only Gp1 or LysA, addition of nisin to cells expressing both proteins resulted in complete lysis. This means that the endolysin was already positioned next to its target, the PG, at lysis onset, since the pore diameter produced by nisin (2 nm) should not allow the passage of a protein as large as Ms6 endolysin (42).

However, how Ms6 endolysin is kept inactive until the proper time of lysis is a question that remains to be elucidated.

Achieving the Proper Time of Lysis

In Ms6, the regulation of mycobacteria lysis timing also seems to display peculiar features. Achievement of the correct lysis timing was shown to be dependent on the interaction and concerted action of two proteins with holin-like features (43), even though the presence of both is not absolutely required for M. smegmatis lysis. Gene gp4 encodes a small protein of 77 amino acids, sharing structural characteristics with class II holins. A holin function was also supported by Gp4’s ability to complement a λS defective mutant (46). A more recent characterization of Ms6 Gp4 function suggested that Gp4 might function as a pinholin, since the first TMD has characteristics of a SAR domain with a high per-
percentage of weakly hydrophobic or polar residues (43). The presence of a SAR domain followed by a typical TMD is characteristic of the previously described pinholins, such as the holin of phage 21, St2168 (18, 27, 52). Coexpression of Ms6 LysA and Gp4 in E. coli does not induce bacterial lysis (46); however, changing Ms6 gp4 by the mycobacteriophage D29 holin gene (gp11) results in E. coli lysis (43). This is consistent with the idea that, in contrast to the D29 holin, Ms6 Gp4 forms holes too small to allow passage of the 43-kDa Ms6 endolysin. Unexpectedly, deletion of gp4 from the Ms6 genome results in earlier lysis timing by about 30 min, a lysis phenotype that seems to be more consistent with antiholin function for Gp4 rather than holin (43). The last gene (gp5) of the Ms6 lysis module was also shown to encode a holin-like protein. Gp5 is a 124-amino-acid protein with a predicted TMD at the N-terminus and a highly charged C-terminus, fitting the structural characteristics of class III holins; however, it does not complement an Sλ defective mutant. Nevertheless, it was shown that Gp5 has a regulatory role in the timing of lysis, since a deletion of gp5 from the Ms6 genome resulted in a viable phage, but with a delayed time of lysis (43), excluding, in light of current knowledge, the possibility of functioning as an antiholin. It has been described elsewhere that null mutation in antiholin genes causes acceleration of the onset of lysis, while null mutation in holin genes results in a delay of the timing of lysis (4, 11, 53).

The observation that Gp4 interacts with Gp5 supports the notion that, in Ms6, the holin function is a result of the combined action of Gp4 and Gp5, contributing to the precise adjustment of the timing of hole formation and to keep the infected cell productive, allowing the assembly of more virions (43). The regulation of the timing of lysis as the result of activity of a complex is not restricted to a phage that infects a mycobacterial host. Similarly, it was proposed that the holin function of the Bacillus subtilis prophage PBSX is a result of two holin-like proteins, XhIA and XhIB, that associate in the membrane to form a holin functional unit (34).

It is now clear that mycobacteriophage Ms6 insults the first cell barriers to phage release through a holin-endolysin strategy, but the access of the endolysin to its target is different from all phage-mediated lysis described so far. Figure 3 represents the proposed Ms6 model to overcome the first cell barriers to phage release: the CM and the cell wall. However, mycobacteriophages still have to face a third barrier: the mycobacterial OM. Although mycobacteria are classified as Gram positive,

![FIGURE 3 Model for Ms6 endolysin export. Export of the full-length endolysin (Lysin_{384}) is assisted by the chaperone Gp1 through the Sec translocase. Once in the cell wall compartment, it is kept in an inactive state until the holin complex Gp4/Gp5 dissipates the membrane potential. The endolysin activation is schematically represented by the change of the enzyme spherical configuration to a “pacman” shape. Lysin_{384} is an N-terminally truncated version of Lysin_{384}. (?) indicates that export of this shorter version to the extracytoplasmatic environment is not known. PG, peptidoglycan; CM, cytoplasmic membrane; Cyt, cytoplasm. Adapted from reference 10 with permission. doi:10.1128/microbiolspec.MGM2-0017-2013.f3](image-url)
they have a complex cell envelope, which includes an OM. This is a lipid bilayer, consisting of an inner leaflet rich in mycolic acids that are covalently bound to the PG-arabinogalactan complex via an ester linkage, and an outer leaflet mainly composed of glycolipids, phospholipids, and species-specific lipids (55, 56).

**Overcoming the Last Barrier to Phage Release**

Similar to what has been described for bacteriophage λ and other phages infecting Gram-negative hosts, which encode spanins to overcome the Gram-negative OM, it is not surprising that mycobacteriophages encode additional lysis proteins targeting the mycobacterial OM. In fact, mycobacteriophage Ms6 synthesizes an additional lysis protein, encoded by *lysB*, a gene positioned immediately downstream of the *lysA* gene (46), which was shown to have lipolytic activity (40). The Ms6 LysB amino acid sequence contains a pentapeptide, GYSQG, which matches the conserved G-X-S-X-G motif characteristic of lipolytic enzymes. Biochemical characterization of the enzyme showed that Ms6 LysB hydrolyzes both esterase and lipase substrates, showing a higher affinity for long-chain substrates, a characteristic that is in agreement with the highly diverse lipid content of the mycobacteria cell envelope. Importantly, Ms6 LysB was shown to cleave the ester bond between mycolic acids and the arabinogalactan of the mycolyl arabinogalactan-peptidoglycan (mAGP) complex, releasing free mycolic acids (41). However, the activity of Ms6 LysB is not restricted to mAGP. Ms6 LysB was also shown to hydrolyze other mycobacterial lipid components of the cell envelope, particularly the trehalose dimycolate (TDM, a glycolipid involved in the virulence of pathogenic species) of both fast- *M. smegmatis* and slow-growing mycobacteria (*Mycobacterium bovis* BCG and *Mycobacterium tuberculosis* H37Ra), indicating that Ms6 LysB activity is not species-specific (41). Nevertheless, due to the importance of the mAGP complex to stability of the mycobacteria cell envelope, it was thus proposed that the mAGP complex is the major target for Ms6 LysB activity. Breaking this linkage allows the separation of the OM from the cell wall, disrupting the last barrier to phage release.

Analogs can be made between the function of LysB against the mycobacterial OM and the λ Rz/Rz1 homologues or spanins of Gram-negative hosts, which mediate the third and final step of host lysis by fusing the inner and outer membranes (33, 35, 57, 58). The Rz and Rz1 λ proteins were long considered to be auxiliary lysis proteins since, in laboratory conditions, RzRz1-null mutants are still able to cause lysis unless the OM is artificially stabilized with milimolar concentrations of divalent cations (59). Recently, Young and his collaborators showed that λ lysis actually requires functioning RzRz1 genes without artificial stabilization of the OM and thus could no longer be considered auxiliary lysis proteins (36).

Under normal infection conditions, LysB is not essential for lysis, since a deletion of *lysB* from the Ms6 (60) genome results in a mutant phage that retains the ability to form plaques. If there are other factors that influence the lysis dependence on LysB, it is not currently known. Despite their different modes of action, both spanins and Ms6 LysB mediate the third and final step in host lysis by eliminating the last barrier to phage release, after holins and endolysins have disrupted the CM and PG, respectively. At this time it is unknown how Ms6 LysB is localized to its substrate, since no signal sequences allowing its transport across the cell barriers have been identified (40). One can hypothesize that LysB might reach the OM simply by diffusion after PG breakdown by LysA.

**Diversity of the Mycobacteriophage Lysis Systems**

The availability of more than 270 mycobacteriophage complete genome sequences shows that they are highly diverse; however, some are more closely related than others. Based on nucleotide sequence comparison, these phages have been grouped into clusters and subclusters (for a more detailed explanation see reference 71). The majority of data regarding the lysis systems of mycobacteriophages, other than Ms6, are restricted to bioinformatic analysis of their genomes. Description and comparison of more than 80 phage genome sequences (37, 38, 61–63) provides insights into lysis cassette diversity (Fig. 4).

All mycobacteriophages described to date encode an endolysin. A bioinformatic analysis of 224 sequenced mycobacteriophage endolysins shows that they are highly diverse and modular in nature, with a large number of domain organizations (38, 51). Interestingly, most of them seem to be composed of three domains: a C-terminal domain that is likely to be associated with binding to the cell wall and two catalytic domains—an N-terminal domain with putative peptidase activity and a central catalytic domain that specifies a glycoside hydrolase (muramidase, transglycosylase) or an amidase (51). This is in contrast to the majority of endolysins from phages infecting Gram-positive or Gram-negative bacteria. In general, endolysins from phages infecting Gram-positive hosts are composed of an N-terminal...
**Lysis Genes Organization**

<table>
<thead>
<tr>
<th>Cluster/Subcluster</th>
<th>Phage</th>
<th>Lysis Genes Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Bethlehem</td>
<td>gp7 lysA lysB</td>
</tr>
<tr>
<td>A1</td>
<td>DD5</td>
<td>gp8 HNH lysA lysB</td>
</tr>
<tr>
<td>A2</td>
<td>D29</td>
<td>lysA hol lysB</td>
</tr>
<tr>
<td>B1</td>
<td>Chah</td>
<td>lysA lysB</td>
</tr>
<tr>
<td>B2</td>
<td>Rosebush</td>
<td>lysA hol</td>
</tr>
<tr>
<td>B5</td>
<td>Acadian</td>
<td>lysA gp48 lysB</td>
</tr>
<tr>
<td>C2</td>
<td>Myrna</td>
<td>lysA</td>
</tr>
<tr>
<td>D</td>
<td>PBI</td>
<td>lysA gp36 gp37 gp38 lysB</td>
</tr>
<tr>
<td>D</td>
<td>Troll4</td>
<td>lysA gp36 gp37 gp38 gp39 lysB</td>
</tr>
<tr>
<td>E</td>
<td>Cjw1</td>
<td>lysA hol gp34 lysB</td>
</tr>
<tr>
<td>F1</td>
<td>Ms6</td>
<td>gp1 lysA lysB</td>
</tr>
<tr>
<td>F1</td>
<td>Boomer</td>
<td>gp30 gp31 lysA lysB</td>
</tr>
<tr>
<td>F2</td>
<td>Che9d</td>
<td>lysA lysB hol gp38</td>
</tr>
<tr>
<td>G</td>
<td>Halo</td>
<td>lysA lysB hol gp30</td>
</tr>
<tr>
<td>J</td>
<td>Omega</td>
<td>lysA gp51 gp52 lysB hol gp55</td>
</tr>
<tr>
<td>N</td>
<td>Charlie</td>
<td>gp27 lysA hol gp30</td>
</tr>
</tbody>
</table>

**Ms6 gp1 homologues**

**Holin**

**Mycobil-arabinogalactan esterase**

**Endolysin**

**Holin-like**

**Unknown function**
catalytic domain and a cell wall–binding domain positioned at the C-terminus, while endolysins from phages infecting Gram-negative hosts are mostly small single-domain proteins, usually without a specific cell wall–binding domain module (24).

Some mycobacteriophage endolysins contain both glycosidase and amidase domains, such as Che9d Gp35 and Wildcat GP49 (51). Of note is the fact that for the majority of the mycobacteriophage endolysins, the peptidase motifs were identified through similarities to known peptidases. Except for the M23 domain, peptidase domains were not readily identified by conserved domain searches (51). Although PG hydrolase activity was demonstrated, by zymogram analysis, for the endolysins of TM4, D29, Che8, Bxz1, and Corndog (58, 64), apart from the Ms6 endolysin (50), none of the predicted proteins were, so far, characterized experimentally.

Apart from Ms6, little is known about how endolysins access the extracytoplasmic compartment. Ms6 gp1 homologues have been identified in other mycobacteriophage genomes (Fig. 4), particularly in the lysis cassettes of phages that belong to subclusters A1 and F1 (37, 38). Taking into consideration the high similarity between the lysis genes among phages of the same cluster, one can speculate that phages belonging to subcluster F1 will use the Ms6 strategy to deliver LysAs to their PG substrate. If that is also the case of phages belonging to cluster A1, which although sharing Gp1 homologues, have a different lysis cassette organization, it is a question that deserves further investigation. However, Gp1 homologues are missing from the majority of mycobacteriophage genomes, suggesting that in these phages LysA export occurs in a different way. Similar to Ms6, lysA secondary gene products seem also to occur in other mycobacteriophage genomes. Zymogram analysis of E. coli extracts containing LysA proteins of Corndog (Gp69), Bxz1 (Gp236), and Che8 (Gp32) show that in addition to the full-length protein, smaller products can also hydrolyze the PG (51). Recently, Payne and Hatfull (51) reported that the sole expression of D29, L5, and Kostya endolysins in M. smegmatis mc2155, under the control of the acetamidase promoter, resulted in lysis. This is an intriguing observation since the holin gene was excluded from the expression clone, and none of the proteins is predicted to have secretion signals. As suggested by the authors, it is unlikely that upon a L5 infection, lysis would be holin independent, since a predicted holin gene is localized immediately downstream of the lysA gene. For any degree of confidence, the apparent holin-independent lysis must be evaluated in the phage infection context.

In the majority of mycobacteriophages sequenced so far, no holin function was experimentally determined. Homologues of Ms6 Gp4 and Gp5 were found in other mycobacteriophages, with the highest identity within subcluster F1. However, holin proteins are highly diverse and do not share high sequence identity (11), and thus it is more difficult to identify genes encoding putative holins. Among all mycobacteriophage genomes available in the Genbank database, few have annotated holin genes; however, genes encoding putative proteins with predicted TMDs can be identified in the vicinity of endolysin genes, making them good candidates to encode the lysis-TMD proteins. Whenever a holin gene has been assigned, it is closely linked and downstream to lysA.

Like mycobacteriophage Ms6, the majority of mycobacteriophages also encode a LysB protein. Structure and amino acid sequence comparisons show that LysB proteins are also diverse, although they all share the characteristic motif, GXSXG, of lipolytic enzymes (40, 58, 65). A mycolylarabinogalactan esterase activity was also demonstrated for the LysB protein of D29 (58). These authors have also determined the crystal structure of D29 LysB, revealing structural similarities to cutinases (enzymes that also cleave esterase and lipase substrates). At the time of the last mycobacteriophage genome analysis, lysB genes were identified in 76 out of 80 phage genomes (38). When present, the lysB gene is always positioned downstream of lysA—in some cases immediately downstream, while in other cases it is separated from lysA by no more than four genes (Fig. 4). With few exceptions (38), all genes positioned between lysA and lysB also encode lysis proteins, specifically putative holin proteins (Fig. 4). The high incidence of lysB genes in mycobacteriophage lysis cassettes suggests that LysB proteins may have an important role in lysis. In fact, Payne et al. (58) showed that a deletion of gene lysB from the phage Giles genome delayed the time of lysis by 30 min. It is suggested that this delay results from a defective break of the cell barriers in the absence

**FIGURE 4** Diversity of mycobacteriophage lysis cassettes. The illustration shows representatives of mycobacteriophages with diverse genome organization. Not previously assigned holin-like genes display white bars that represent the number and location of putative TMD coding sequences. Adapted from reference 10 with permission. doi:10.1128/microbiolspec.MGM2-0017-2013.f4
of LysB. This explanation derives support from the observation that after an infection of \textit{M. smegmatis} with a \(\Delta\)lysB mutant of phage Giles, 45\% of the phage particles remained associated with unlysed cells in contrast to the 10\% observed in a wild-type infection. It seems, therefore, that in the former, the new phage particles are trapped within cell debris as a result of a deficient elimination of the OM, resulting in a delay of their release into the environment.

Similar to what happens with phages that infect Gram-negative hosts, lysis induced by mycobacteriophages seems to be a three-step event, with LysB playing a role in the third step: elimination of the OM layer. For the few phages that do not encode a LysB protein (Che12, Rosebush, Quirzula, Myrna, and Charlie) (38, 39, 61), it was suggested that they have evolved a mechanism to utilize a host-encoded cutinase-like protein that would replace the function of LysB (37).

It is interesting to note that \(lysB\) equivalents were recently described in phages infecting other bacterial groups that have in common with mycobacteria a thick mycolic-acid-containing outer layer covalently linked to the PG—the mycolata group. Examples are the \textit{Rhodococcus equi} phages DocB7, Pepy6, and Poco6 (66). This supports the notion that the presence of LysB-like proteins is a selective advantage for phages infecting members of the mycolata group.

In conclusion, like phages that infect Gram-negative hosts, lysis induced by mycobacteriophages is also a three-step event (Fig. 5). The first step is achieved by holins that compromise the CM, allowing the endolysins to hydrolase, in a second step, the PG layer. Finally, LysB proteins disrupt the mycobacterial OM, eliminating the last barrier to mycobacteriophage release.

**APPLICATIONS OF THE LYSIS PROPERTIES**

The capacity of bacteriophages to lyse their hosts has been intensively explored, with the main goal of elimination of pathogenic bacteria. These studies are driven either by the use of bacteriophage itself as a therapeutic agent (phage therapy) (67) or for exploring the enzymes (enzymbiotics) involved in cell lysis for enzyme therapy (26, 67). Phage-encoded endolysins have shown high potential as antibacterial agents against a number of Gram-positive bacteria. At first glance, mycobacteriophages or their lysis enzymes might be of limited use as therapeutic agents to fight tuberculosis, in part due to the inaccessibility of phages to the pathogenic bacteria located intracellularly. On the other hand, the presence of

**FIGURE 5** Targets of Ms6 lysis proteins. Schematic representation of the mycobacteria cell envelope, where the target layer of each protein is indicated by an arrow. Arab, arabinan; CM, cytoplasmic membrane; Gal, galactan; LAM, lipoarabinomannan; OM, outer membrane; P, protein; PG, peptidoglycan; PIMs, phosphatidylinositol mannosides; PLs, phospholipids; Po, porin; Pp, periplasm; TDM, trehalose dimycolate; TMM, trehalose monomycolate. doi:10.1128/microbiolspec.MGM2-0017-2013.f5
an OM in the mycobacterial cell envelope constitutes a barrier for endolysins to access the PG from without. However, some studies on the use of mycobacteriophage lysis properties have recently been reported. Broxmeier et al. (68) have shown the potential of mycobacteriophage TM4 delivered by the nonpathogenic M. smegmatis to kill Mycobacterium avium and M. tuberculosis within macrophages. An in vivo assay with mice infected with M. avium showed that treatment with M. smegmatis transiently infected with TM4 significantly reduced the number of M. avium cells in the spleen (69). More recently Trigo et al. (70) demonstrated for the first time the potential of phage therapy against a Mycobacterium ulcerans infection. These authors showed in the mouse footpad model that a single subcutaneous injection of the lytic mycobacteriophage D29 can effectively decrease the proliferation of M. ulcerans.

CONCLUDING REMARKS

With the increased knowledge of phage lysis, it has become evident that phages have evolved sophisticated solutions to improve their survival in the biosphere. The studies on the Ms6 lysis model have provided new insights into the mechanisms of bacteriophage-induced lysis. In order to obtain the highest level of fitness, mycobacteriophages have evolved specific lytic functions by acquiring additional and specific lysis genes that confer a selective advantage by allowing efficient disruption of the complex cell barriers. It seems clear that the genetic diversity observed among the mycobacteriophage lysis cassettes reflects the ability of mycobacteriophages to adapt to different environmental conditions. The huge number of isolated mycobacteriophages opens exciting research opportunities to explore the mycobacteriophage-induced lysis pathways, which can give us new clues into mycobacterial secretion systems. Identification of the lysis enzymes that target specific linkages in the cell envelope, and thus compromise the mycobacteria viability, also opens new perspectives to explore these enzymes as new tools to destroy mycobacteria cells.

REFERENCES


