

Faculté des sciences

Molecular characterization of phage Deep-Purple infecting the *Bacillus cereus* group

Author : Manon Nuytten
Supervisor : Prof. Jacques Mahillon (UCLouvain / ELI / MIAE)
Scientific advisor : Audrey Leprince
Reviewers : Prof. Claude Bragard (UCLouvain / ELI / ELIM)
Prof. Henri Batoko (UCLouvain / LIBST)
Master thesis year : 2019 - 2020

Acknowledgements

I would like to start this report by expressing my gratitude to all the people that have made this work possible.

In the first place, I would like to warmly thank my supervisor, Professor Jacques Mahillon, for hosting me in his team and for his guidance for each stage of this research work.

I also acknowledge my reviewers, Professor Claude Bragard and Professor Henri Batoko, for accepting to be part of my master thesis jury.

I am extremely grateful to my scientific advisor, Audrey Leprince, for her precious help all year long. It has been a real pleasure to work and learn by your side. Thank you for your enthusiasm, your patience and your sympathy.

Many thanks to the entire MIAE team for the day-to-day help and for the friendly atmosphere in the lab. I have been delighted to share this year with you. I would like to especially thank Pauline Hinnekens and Marie Verhaegen for their friendly support.

Finally, I must express my profound gratitude to my family and friends for their unfailing support during this busy year. I wish to especially acknowledge my mother for her great encouragement since the beginning of my studies. Your pride has always been my greatest motivation.

Abstract

As bacterial parasites, phages are generating considerable interest due to their potential applications in the development of antimicrobials and detection tools. The understanding of crucial steps of phage infection cycle is fundamental for phage research. This includes the adsorption stage which involves the interaction of bacterial receptors with phages proteins termed Receptor Binding proteins (RBPs) and the lysis process occurring at the end of the phage infection cycle upon bacterial peptidoglycan degradation performed by phage endolysins. Previous studies have mainly focused on Gram-negative targeting phages and much less information is currently available on phages preying on Gram-positive bacteria. This work aimed to shed light on the adsorption and lysis processes of Deep-Purple, a tailed phage belonging to the *Siphoviridae* family and infecting the *Bacillus cereus* group.

Bioinformatic analyses of Deep-Purple tail proteins retained two putative RBP candidates namely gp28 and gp29. Gp28 was chosen for its putative evolved Dlt function, reported to feature carbohydrate binding modules able to interact with the bacterial surface and sometimes even replacing the RBP in some phages including the *Lactobacillus casei* phage J-1. Gp29 was investigated for its putative tail fiber function and due to its C-terminal intramolecular chaperone domain (IMC), structurally similar to the IMC of *Salmonella* phage S16 RBP. Both candidates were tagged with GFP and recombinant proteins were successfully produced in *Escherichia coli* cells and purified. Their capacity to bind to *Bacillus cereus* was assessed in a cell wall decoration assay. This allowed to observe that both proteins are capable of interacting with *B. cereus* cells. Further investigations are required to determine if both are *bona fide* RBP and their precise roles in phage adsorption.

Gp32 was investigated as sole potential endolysin candidate because of its sequence similarity with other endolysins. It displayed a putative N-terminal domain associated with 1,4-N-acetylmuramidase activity (EAD, Enzymatically Active Domain) and a C-terminal SH3 domain putatively involved in host recognition (CBD, Cell wall Binding Domain). This work managed to produce and purify gp32 protein and its lytic activity was confirmed by turbidity reduction assay on *B. cereus* cells and extracted peptidoglycan. The spectrum of activity of the protein was assessed and revealed that all tested members within the *B. cereus* group were sensitive to gp32 and some tested members outside of the *B. cereus* group also showed endolysin susceptibility. Then, the lytic activity of gp32 was assessed under varying conditions showing that gp32 was effective in conditions ranging from pH 4 to 8 and under 37°C but was highly sensitive to NaCl concentrations. Full-length gp32 and its CBD were tagged with GFP in order to evaluate their binding capacity in a cell wall decoration assay. These experiments confirmed their capacity to tightly bind to bacteria and corroborated the activity spectrum assessed before. Finally, the spectrum of activity of gp32 was compared to Deep-Purple spectrum determined via top agar experiments and revealed that endolysin can act upon phage resistant bacteria.

All in one, our data allowed the identification of two tail proteins interacting with *B. cereus* cells during adsorption and collected considerable information regarding the endolysin activity of Deep-Purple phage.

Abbreviation list

aa	amino acid
ATP	Adenosine Triphosphate
BHI	Brain Heart Infusion
BLAST	Basic Local Alignment Search Tool
BppU	Baseplate upper protein
Cas	CRISPR associated
CBD	Cell wall Binding domain
CDS	Coding DNA Sequence
CHAP	Cysteine, Histidine-dependent Amidohydrolase/Peptidase
ChBr	Choline binding module
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
ddH₂O	Double-distilled water
Dit	Distal tail protein
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double-stranded DNA
DTT	Dithiotheritol
EAD	Enzymatically Active Domain
GFP	Green Fluorescent Protein
GlcNAc	N-acetylglucosamine
GO	Gene Ontology
gp	Gene product
HF	High Fidelity
IMC	Intramolecular Chaperone
IPTG	Isopropyl- β -D-thiogalactopyranoside
kb	Kilobase
LacI	Operon lactose inhibitor
LacO	Operon lactose operator
L-ala	L-alanine
LB	Lysogeny Broth
LPS	Lipopolysaccharide
LTA	Lipoteichoic acid
LTF	Long Tail Fiber
MurNAc	N-acetylmuramic acid
Omp	Outer membrane protein
O/N	Overnight
ORF	Open Reading Frame
PAM	Proto-spacer Adjacent Motif
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction

PG	Peptidoglycan
RNA	Ribonucleic acid
rpm	Revolution per minute
RT	Room Temperature
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
<i>s.l.</i>	<i>Sensu lato</i>
SOC	Super Optimal broth with Catabolic repression
sp.	unspecified species
spp.	Several unspecified species
<i>s.s.</i>	<i>Sensu stricto</i>
ssDNA	Single-stranded DNA
STF	Short Tail Fiber
Tal	Tail-associated lysin
Tm	Melting temperature
TMD	Transmembrane domain
TMP	Tape Measure Protein
tRNA	Transfer ribonucleic acid
WTA	Wall Teichoic Acid

Contents

1	Introduction	5
1.1	Bacteriophages	5
1.1.1	Biology and classification	6
1.1.2	Life cycles	7
1.1.3	Phages applications	8
1.1.4	Phage and bacterial coevolution	9
1.2	Phage adsorption	11
1.2.1	Overall process	11
1.2.2	Tail proteins	11
1.2.3	Bacterial receptors	14
1.3	Host cell lysis	16
1.3.1	Overall process	16
1.3.2	Endolysins	16
1.3.3	Holins	19
1.4	Deep-Purple, the Siphovirus, infecting <i>B. cereus</i>	20
1.4.1	The <i>B. cereus</i> group	20
1.4.2	Phages preying on <i>B. cereus</i>	21
1.4.3	Phage Deep-Purple	21
2	Material and methods	24
2.1	Strains and plasmids	24
2.1.1	Competent cells preparation	27
2.2	Culture media and conditions	27
2.3	Molecular methods	28
2.3.1	Cloning	28
2.3.2	PCR screening for Deep-Purple presence in <i>Bacillus</i> strains	31
2.4	Recombinant protein expression	31
2.4.1	Induction	31
2.4.2	Solubility assay	32
2.4.3	Protein gel electrophoresis	32
2.4.4	Purification under native conditions	33
2.5	Cell wall extraction	33
2.6	Endolysin characterization	34
2.6.1	Turbidity reduction assay	34
2.7	Cell wall decoration assay	34
2.8	Top agar	35
2.9	Bioinformatics	35

3	Results and discussion	36
3.1	Bioinformatic analyses of Deep-Purple genome	36
3.1.1	Focus on tail proteins	36
3.1.2	Lysis cassette	38
3.2	Deep-Purple lifestyle	38
3.3	RBP identification	39
3.4	Endolysin	42
3.4.1	Molecular cloning and protein induction	42
3.4.2	Gp32 lytic activity	44
3.4.3	Cell wall decoration assay	47
3.4.4	Gp32 activity summary	49
4	Conclusions and perspectives	50

Chapter 1

Introduction

As natural bacterial parasites, phages can find various applications in the detection and the biocontrol of pathogens.

For all phages, the infection process starts with the adsorption step, mediated by the irreversible attachment of Receptor Binding Proteins (RBPs) to bacterial receptors. This specific interaction influences the phage host spectrum. After this crucial step, the genetic material of phages is injected into the bacterial host. At this point, phages can perform two types of life cycles : the lytic cycle (performed by virulent phages) or the lysogenic cycle (performed by temperate phages). Endolysins are enzymes used by phages at the end of their infection cycle to lyse bacteria by breaking down the peptidoglycan layer.

Besides phages, the enzymes that they produce have a great potential for biocontrol and detection purposes. Indeed, RBP can be used as biosensors allowing the detection of bacteria when fused to GFP (Green Fluorescent Protein). Additionally, endolysins can be applied externally to lyse Gram-positive bacteria and can therefore be used as antimicrobial agent. The main advantage of endolysins is that they are less susceptible to face bacterial resistance because they target bonds that are essential for the bacteria. Moreover, the development of protein engineering strategies can optimize them for specific applications.

1.1 Bacteriophages

Bacteriophages are the most abundant and diverse entities on Earth with an estimated global population superior to 10^{31} particles. As natural obligate intracellular parasites, they can thrive in virtually all reservoirs populated by bacteria and archaea. They form a very dynamic population with a turning over every few days (Hatfull, 2015).

Phages were first reported by Frederick W. Twort in 1915 that observed lysed colonies of micrococci and managed to isolate the lysis causative agent (Twort, 1915). Two years later, Felix d’Herelle described an agent able to lyse bacteria causing dysentery that he named bacteriophages (for "bacterial eaters", commonly shortened as phages) (d’Herelle, 1917). Phages were rapidly used in the medical field to control human pathogens but western countries lost interest for phage therapy after antibiotics discovery. Only countries suffering from poor access to antibiotics (*e.g.* Georgia as part of the former USSR) kept going with phage research (Fruciano and Bourne, 2007). However, western countries are nowadays turning back to phage research not only for their molecular and biotechnological applications but also as an alternative to tackle emerging and spreading bacterial resistance. Additionally, the use of phages as biological model allowed the understanding of core biological processes including the establishment of the central dogma of molecular biology (information sequentially passed from DNA to RNA to protein). Furthermore, phage research allowed the discovery of several techniques vital to modern biology including sequencing and CRISPR-Cas among others (Salmond and Fineran, 2015).

1.1.1 Biology and classification

Phages specificity can be limited to one bacterial species or a subset of the same species. All bacterial and archaeal genera are known to be vulnerable to many phages (Sharma et al., 2017).

Phages are classified in a hierarchical system defining one order and ten families. The regulating body for the viral taxonomic system (*i.e.* International Committee on Taxonomy of Viruses) carries out phage particles characterization using parameters such as host spectrum, physical characteristics, type of genomic material, genome size and resistance to organic solvents. Most phages described to date (96%) belong to the *Caudoviridae* order, possessing a double-stranded DNA (dsDNA) genome packaged within a proteinaceous capsid connected to a tail (Ackermann, 2011). Based on tail morphology, the *Caudoviridae* order is divided into three families : *Myoviridae* (*e.g.* *Escherichia coli* phage T4), possessing contractile tails ; *Siphoviridae* (*e.g.* *E. coli* phage λ), featuring long non-contractile tails and *Podoviridae* (*e.g.* *E. coli* phage T7), with short stub-like tails (Fig. 1.1) (Ackermann, 2006).

The phages population exhibits a high diversity of sizes and shapes. Generally, phage particle sizes range from 24 to 200 nm. Proteinaceous capsids mostly feature icosahedral shape and allow the packaging of the genetic material (ssDNA, dsDNA or RNA). Phage tail consists of a hollow tube allowing the transfer of phage genetic material during infection. Additionally, some phages feature plate, fibers and spikes attached to their tail, helping them during attachment to the bacterial cell (Sharma et al., 2017). Phage genome sizes range from 4 to 600 kb and in most cases, the bigger the genome, the more complicated the phage (Brussow and Hendrix, 2002). Generally, phages feature a modular genome organization with structural proteins encoded in the late gene region in a conserved order in relation with the assembly process (Dunne et al., 2018b).

Tailless phages are classified based on their shape, the nature of their envelope or of their genome (ssDNA, dsDNA, RNA).

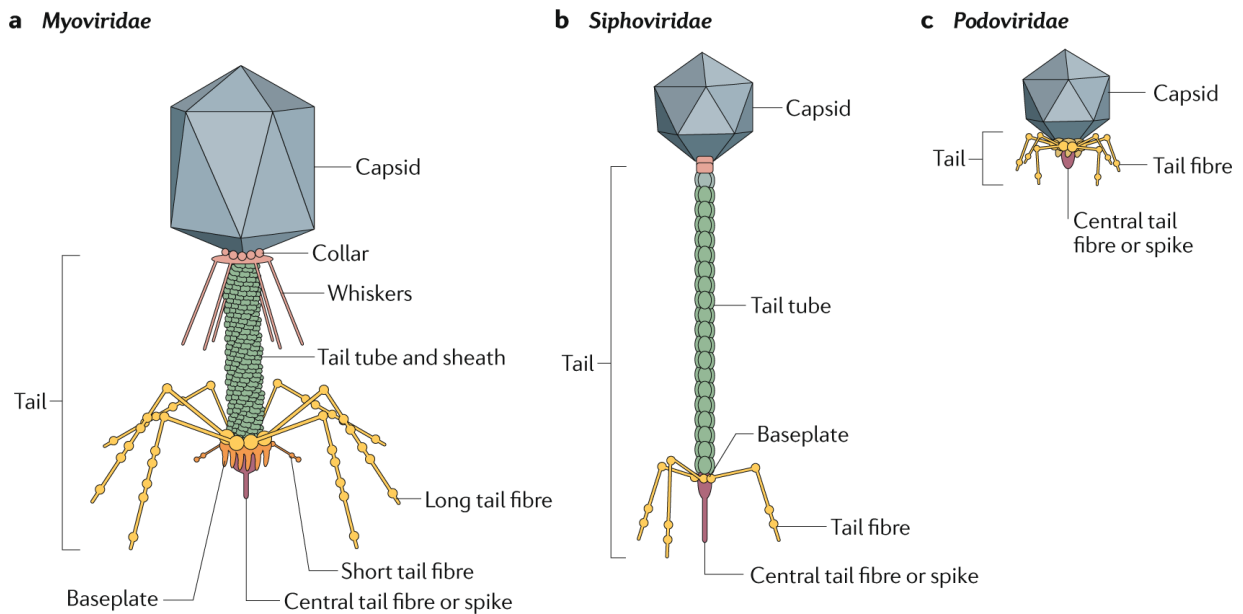


Figure 1.1: Representative structures of tailed phages (*Caudoviridae* order) ; a : *Myoviridae* family featuring contractile tail sheath ; b : *Siphoviridae* family featuring long non-contractile tail ; c : *Podoviridae* family with short stub-like tail and no baseplate structure (Nobrega et al., 2018).

1.1.2 Life cycles

After bacterial recognition, phages inject their genetic material into the host. Then, they replicate inside the bacterial cell by performing either a lytic or lysogenic life cycle. In the lytic cycle, virulent phages hijack the bacterial metabolism using protein-synthesizing and energy-generating systems to produce new virions that are released after cell lysis. In the lysogenic cycle, temperate phages coexist with their host during many generations in a dormant form known as prophage (either integrated into the host genome or residing as a plasmidial molecule) (Howard-Varona et al., 2017). Schematic representation of both life cycles can be seen in Fig. 1.2.

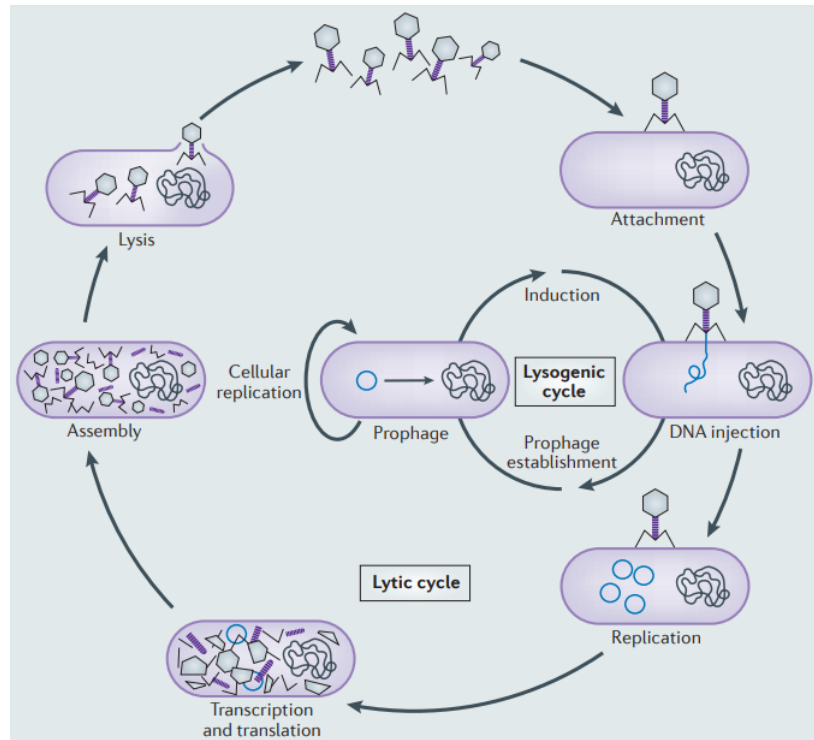


Figure 1.2: Schematic representation of phage life cycles. In the lytic cycle, the phage genome is replicated and transcribed by hijacking the host machinery. New virions are assembled and released after cell lysis in order to infect new bacterial cells. In the lysogenic cycle, the phage (termed as prophage) and the host genome coexist inside the bacterial cell for many generations. Infections are not productive, no virions are formed unless induction of a lytic cycle (Salmond and Fineran, 2015).

Lytic cycle

During lytic cycle, virulent phages infect host cells in order to produce phage progeny that will eventually be released after host cell lysis and be able to infect neighboring cells. After entry of the viral genome, phage encoded genes are expressed in the bacterial cytoplasm to take over the host metabolism therefore allowing the production of new virions (Young, 2014). During the latent period, phage particles assembly takes place. For tailed phages, this follows a four-step process beginning with the assembly of a prohead, followed by the ATP-dependent packaging of DNA then maturation of the prohead and final attachment of the neck and tail proteins (Ackermann, 2007). Once a threshold number of assembled virions is reached, late phage genes including holin and endolysin (for classical dsDNA phages) are expressed (Nicastrò et al., 2016). Holins perforate the host cytoplasmic membrane which allows the endolysin to get access to the peptidoglycan layer for its degradation (Young, 2014).

Lysogenic cycle

Temperate phages can establish a stable relationship with the bacterial host termed lysogeny. During lysogeny, the genome of temperate phages, termed prophage, replicates together with the host DNA, either as a linear or circular plasmid-like state or integrated into the bacterial chromosome (integration requires an integrase which binds homologous segments of phage and bacterial DNA, resulting in site-specific recombination (Hyman et al., 2012)) (Salmond and Fineran, 2015). In lysogenic cycles, prophages need to be replicated sufficiently to be inherited by daughter host cells after host replication. Lysogenic infections are not productive, meaning that no virions are produced due to the phage repressors activity. In response to danger signals or spontaneously, the repressors activity can stop, therefore allowing prophages to exit the lysogenic state and perform lytic cycle. This switch is commonly termed induction (Nicastrò et al., 2016).

1.1.3 Phages applications

Recently, phages have become an important tool in the biotechnological field allowing drugs designing, delivery of proteins of interest and detection of pathogenic bacteria among other applications (Nicastrò et al., 2016). Phages and their encoded proteins have applications in a plethora of domains including medicine, food safety, aquaculture, agriculture and biosensor development (Schmelcher et al., 2012).

Phages as antimicrobial agents

As already mentioned, endolysins are phage enzymes that degrade the peptidoglycan of the bacterial host from within during late stages of infection. Due to the absence of an outer membrane in the Gram-positive bacterial cell wall, endolysins can degrade the peptidoglycan from the outside and therefore be used as antimicrobials. This is especially interesting when considering emerging and spreading resistance of bacteria against classical antibiotics. Furthermore, protein engineering strategies such as domain swapping and random mutagenesis can modify their properties and consequently potentially optimize them for specific applications (Schmelcher et al., 2012). Recent findings involving peptides with outer membrane-disrupting properties fused to endolysins could allow the tackling of Gram-negative bacteria (Schmelcher et al., 2012). The main advantages of endolysins compared to classical antimicrobials is their high specificity for some peptidoglycan types (restricted to certain bacterial genus, species or serotype peptidoglycan types) and their less susceptibility to bacterial resistance (Schmelcher et al., 2010).

However, the development of phage therapy is slowed down in European countries due to the lack of an appropriate regulatory framework (Huys et al., 2013). However, the "Phagoburn" clinical trial involving three European countries and investigating the safety and efficacy of phages for the treatment of burn wounds infected with *E. coli* and *Pseudomonas aeruginosa* has shown encouraging results (Servick, 2016). In food safety and agriculture fields, faster applications of phage technologies have emerged, with several products already commercially available for the tackling of *Listeria monocytogenes* in processed foods (*i.e.* LISTEX P100 (Chibeu et al., 2013) and ListShield (Yang et al., 2017)).

Phages for detection

Classical culture-based diagnostic protocols are the most widely used techniques for the detection of pathogens due to their sensitivity and because they yield colonies that can be further subjected to other tests. However, the main drawback of this protocol is time requirement as it often takes 48 to 72 h to obtain preliminary results (Hagens and Loessner, 2007). Numerous culture-independent methods have been developed in an attempt to design a faster detection tool (*e.g.* PCR-based, immunological and mass spectrometry techniques). However, those rapid

methods require time-consuming pre-enrichment steps, are generally more expensive and PCR-based techniques cannot differentiate death and living cells. Phages and their enzymes, especially the RBP and Cell wall Binding domain (CBD) of endolysins (responsible for host recognition), represent ideal detection tools thanks to their high specificity, natural signal-amplifying properties, inexpensive production and robustness. Moreover, phages are, in principle, able to make the difference between dead and living cells (Schmelcher and Loessner, 2014).

Phages particles can be used in phage amplification assay. To this end, the tested samples are mixed with phages targeting the bacteria tested. Viricide is added after infection time lapse to ensure that extracellular phage particles are inactivated, therefore only keeping phages engaged in the infection process. Helper cells are added after neutralization of the viricide to allow signal amplification by infection of those cells by progeny phages as the number of cells present in the tested samples are generally insufficient to yield bacterial lawns. Final plating of mixture in a soft agar overlay results in formation of countable plaques corresponding to the initial number of target cells in the sample (Schmelcher and Loessner, 2014).

Besides unmodified phage particles, engineered phage particles carrying a specific reporter gene can also be used for bacterial detection in reporter phage assays. Upon infection of a target cell, the phage reporter gene is expressed and allows detection of its product (*e.g.* by measurement of bioluminescence, fluorescence or enzymatic conversion of a chromogenic substrate). The main advantages of this technique is its high speed, sensitivity and the only detection of living cells. Yet, their construction requires detailed genetic knowledge of the phage (Schmelcher and Loessner, 2014).

In addition, phage proteins (especially RBPs and CBDs of endolysins) can be exploited as affinity molecules for the coating of beads consequently allowing the specific immobilization and detection of pathogens (*i.e.* CBD-coated beads for detection of Gram-positive *L. monocytogenes* (Walcher et al., 2010) and RBP-coated beads for the detection of Gram-negative *Campylobacter jejuni* (Singh et al., 2011)) (Kretzer et al., 2007).

At the end of the phage replication cycles, endolysins provoke the death of the host cells, not only releasing new progeny virions but also intracellular components that can be used as intracellular marker for the measurement of bacterial lysis (*e.g.* ATP (Squirrell et al., 2002) ; β -D-galactosidase (Neufeld et al., 2003)).

1.1.4 Phage and bacterial coevolution

Bacterial resistance

Phage and bacterial coevolution has led to the emergence of several bacterial resistance mechanisms.

Some bacterial strains have developed mechanisms to prevent phage adsorption to their receptors by blocking of the phage receptors, production of capsule or slime layers made out of exopolysaccharides or production of competitive inhibitor. The latter involves competition for the binding to the host receptors as phage receptors generally also play an important role in bacterial metabolism (*e.g.* substances intake) (Labrie et al., 2010). In addition, mutation or loss of phage receptors may also confer resistance to phage adsorption (Hyman and Abedon, 2010).

Superinfection exclusion systems allow some bacterial strain to prevent phage DNA entry, therefore conferring them immunity to these phages. These proteins are putatively associated with bacterial membrane and are encoded by prophage genes therefore suggesting that this system is more important for phage-phage interactions rather than phage-host interactions. The coliphage T4 has two superinfection exclusion systems encoded by *imm* and *sp* to prevent subsequent infection by other T-even-like phages (Fig. 1.3) (Labrie et al., 2010).

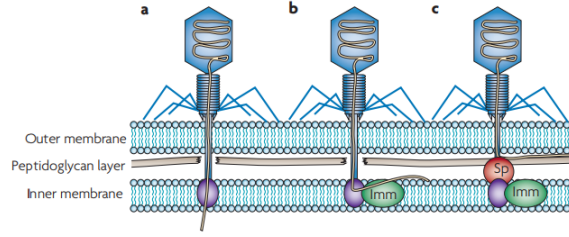


Figure 1.3: Schematic representation of T4 superinfection exclusion system. a. Normal phage infection. b. T4 phage protein Imm blocks the entry of phage DNA into the cytoplasm c. T4 phage protein Sp blocks the degradation of the peptidoglycan therefore blocking the invading phage DNA between the PG layer and the outer membrane (Labrie et al., 2010).

Restriction-modification systems are used by numerous bacteria to prevent phage infection by cleavage of invading phage DNA. When un-methylated phage DNA enters the bacteria, it will be mostly recognized by the bacterial restriction enzymes and rapidly degraded. More rarely, it can be methylated by a bacterial methylase to avoid restriction, therefore allowing the phage infection cycle to occur (Labrie et al., 2010).

The bacterial CRISPR–Cas systems confer adaptive immunity against foreign DNA such as phage DNA. CRISPR–Cas loci are generally composed of 21 to 48 bp direct repeats interspaced by non-repetitive spacers of similar length and flanked by *cas* genes in variable numbers. Phage insensitive strains have acquired at least one repeat-spacer unit on their CRISPR locus identical to the proto-spacer sequence found in the phage genome, conferring them immunity to that phage (Mojica et al., 2009).

Finally, some bacteria have acquired phage abortive infection systems in which some intracellular bacterial proteins target crucial steps of phage multiplication (*i.e.* transcription, protein synthesis, maturation and host cell lysis). This abortion of phage infection also leads to the death of the infected cell, therefore representing a form of bacterial altruism (Hara-Kiri) as it allows the remaining non-infected bacterial population to escape phage infection. Their exact mechanisms are still poorly understood. One well-characterized abortive infection system is the Rex system found in *E. coli* phage λ . Upon phage infection, a phage protein–DNA complex is produced and this activates the intracellular sensor RexA. This activation triggers the ion channel RexB, which in turn induces the reduction of the membrane potential leading to a drop in ATP level. Finally, this triggers a decrease of macromolecules synthesis and thereby stops cell multiplication along with abortion of phage infection (Fig. 1.4) (Labrie et al., 2010).

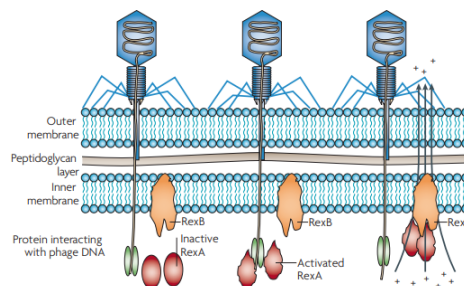


Figure 1.4: Schematic representation of Rex system. The sensor protein RexA is activated by the phage protein–DNA complex which in turn triggers the ion channel RexB that allows the passage of cations, therefore decreasing the membrane potential, resulting in death of the host cell and abortion of phage infection (Labrie et al., 2010).

Phage response to bacterial resistance

Phage diversity is mostly driven by their dynamic adaptation by selective pressure when facing resistant bacteria.

Phage can respond to host receptor mutation either by random mutation, random recombination with other entities or reduction in receptor recognition stringency (Hyman and Abedon, 2010). Inaccessibility to host receptor can be overcome by polysaccharide lyase or polysaccharide hydrolase activity that cleaves the exopolysaccharides layer covering host receptors (*e.g.* this enzymatic activity can be acquired by genetic transfer) (Labrie et al., 2010).

To cope with the restriction-modification systems, several phages have acquired some anti-restriction strategies. One of those involves the absence of endonuclease recognition sites in their genomes through acquisition of point mutations. Other strategies involve the presence of cognate methylase gene in their genome that methylates phage DNA or co-injection of phage protein masking restriction sites in their DNA (Labrie et al., 2010) ; (Samson et al., 2013).

Phage can overcome the CRISPR-Cas system by acquiring a point mutation in the targeted proto-spacer or in the conserved PAM region (proto-spacer adjacent motif). Phage escape can also occur when those sequences are deleted in the phage genome. Moreover, small specific phage proteins can interfere with the formation or action of the CRISPR-Cas complexes, therefore conferring CRISPR immunity (Samson et al., 2013).

Phages display numerous anti-abortive infection mechanisms. For example, in phage T4, the *rII* genes (*RIIA* and *RIIB*) prevent the *rex* gene products activity. However, *rII* does not confer complete protection as overexpression of *rexA* and *RexB* still leads to T4 exclusion (Labrie et al., 2010).

1.2 Phage adsorption

For all phages, the infection process is initiated by the adsorption process, which is mediated by the irreversible attachment of phage RBP to specific receptor at the bacterial surface. This crucial interaction influences the phage host range.

1.2.1 Overall process

A phage recognizes a potentially sensitive host through a series of interactions between phage RBP and host surface components, resulting in positioning of the phage for DNA injection into the host. Phage adsorption is generally described as a three-step process. The first step consists of random collisions between phage and host governed by Brownian motion, dispersion, diffusion or flow. This is followed by the reversible binding of the phage to bacterial surface components. At this stage, the phage keeps the ability to desorb from the host, allowing the phage to stay close to the cell surface as it searches for a specific receptor. Finally, the phage will bind irreversibly to the bacterial receptor, which triggers conformational rearrangements in other phage molecules leading to the injection of the genetic material into the host. More than one receptor and one RBP may be involved in the adsorption process. In some cases, the phage proteins and bacterial receptors involved in the reversible and the irreversible steps are different (Bertozzi Silva et al., 2016).

1.2.2 Tail proteins

The tail structures are important determinants of the phage host spectrum. Tailed phages use RBPs located at the distal part of their tail apparatus to specifically interact with the host receptors. Tail fibers, tail spikes and tail tips have been associated with RBP function (see Fig. 1.1) (Nobrega et al., 2018).

***Myoviridae* : *E. coli* phage T4**

E. coli phage T4 is used as model for the *Myoviridae* family. Its baseplate features both short (gp12) and long (gp34 to gp37) tail fibers. Presumably, the long tail fibers extend when T4 approaches the host surface and bind to bacterial surface components (*e.g.* lipopolysaccharide, LPS or outer membrane protein C, OmpC), triggering baseplate conformational change. This enables short tail fibers to irreversibly bind to their specific receptor (*e.g.* outer core region of the LPS). After further conformational change of the baseplate, the short tail fibers position phage for DNA injection. T4 baseplate attachment is illustrated in Fig. 1.5.

Notably, the T4 baseplate is more complex than the baseplates of most other *Myoviridae* (Leiman et al., 2003).

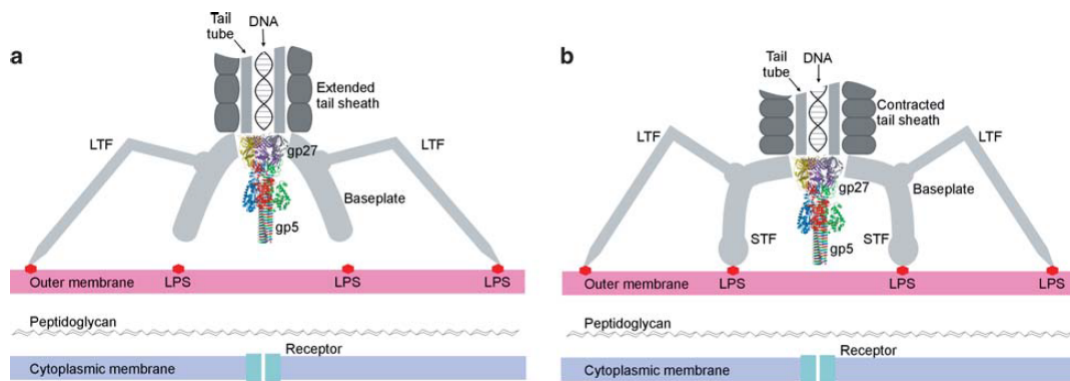


Figure 1.5: Schematic representation of T4 baseplate attachment process. a. LTFs bind to bacterial surface components, triggering change of baseplate conformation. b. T4 phage attaches its baseplate to the cell surface thanks to STFs, initiating DNA injection. Abbreviations are : LTF - long tail fibers ; STF - short tail fibers ; LPS - lipopolysaccharides. Adapted from (Leiman et al., 2003).

Siphoviridae

Siphophages that target Gram-negative bacteria do not have a baseplate-like structure and generally feature both a central straight fiber and side fibers (Nobrega et al., 2018).

Siphophages infecting Gram-positive bacteria can feature a baseplate-like structure similar to myophages (described in Section 1.2.2) but less complex. Siphophages that target protein receptors have only one straight tail fiber or spike attached to the tail or baseplate whereas those that target polysaccharide receptors usually feature side fibers with numerous saccharide-binding domains (Le et al., 2013).

The tail morphogenesis between tape measure protein (TMP, determines the tail length and facilitates DNA injection during infection) and the lysis cassette (*i.e.* gathering both holin and endolysin gene) of Gram-positive targeting *Siphoviridae* is relatively well-conserved as it can be observed in Fig. 1.6.

The first gene generally found after TMP is the distal tail protein (*dit*) gene. This well-conserved protein forms a central hexameric hub allowing the anchorage of other tail components. Recent findings suggest that some *dit* proteins termed as "evolved Dit" features two additional carbohydrate-binding modules of which one is capable of interaction with host receptors, even sometimes replacing RBP (Hayes et al., 2018).

After *dit* gene, we usually find a tail-associated lysin (*tal*) encoding gene. Those trimeric complexes found at the tip of the tail apparatus allow the PG degradation upon phage genetic material injection (Rashel et al., 2008).

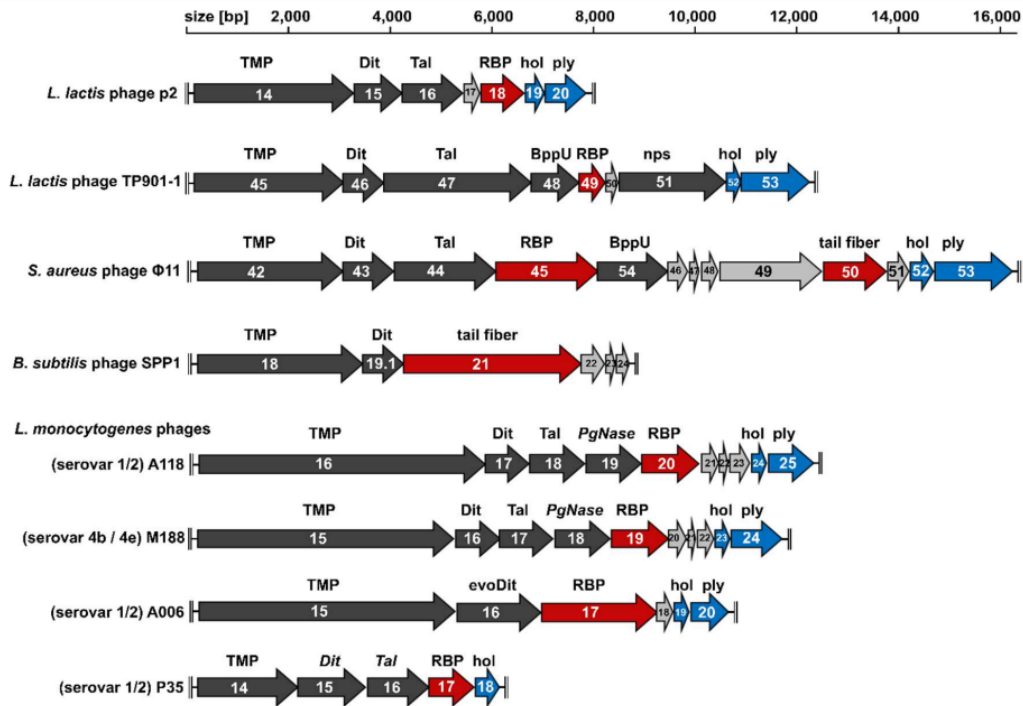


Figure 1.6: Comparison of tail proteins of representative Gram-positive targeting siphophages. Hypothetical proteins in light grey. Proteins with known or predicted functions in dark grey. RBP or tail fibers in red. Lysis cassette in blue. Abbreviations are : TMP - tail tape measure protein ; Dit - distal tail protein ; evoDit - evolved Dit ; Tal - tail-associated lysin ; BppU - baseplate upper protein ; hol - holin ; ply - endolysin (Dunne et al., 2018b).

In contrast with the previous genes, large differences between structural proteins encoded downstream of the *tal* gene can be observed, therefore reflecting the diversity of phage structures involved in host recognition. Some phages feature an additional baseplate upper protein (BppU) allowing the connection between the central Dit hub and the RBP(s) (Dunne et al., 2018b).

The RBP usually follows the Tal protein or BppU (when present). To date, the majority of described Gram-positive targeting RBPs belong to *Lactococcal Siphoviridae* family. Despite low sequence similarity, all RBPs form trimeric complexes and each monomer features a conserved modular organization featuring head (involved in host recognition), neck, stem and shoulder domains (Dunne et al., 2018b).

***Podoviridae* : *E. coli* phage T7**

E. coli phage T7 is used as model for the *Podoviridae* family. Its short non-contractile tail consists of an upper tail adaptor protein connecting its tail to its capsid and followed by a hexameric nozzle containing the tail end. Like other *Podoviridae*, T7 podophage features no baseplate structure (Nobrega et al., 2018).

Upon infection process, primary interactions with host membrane proteins facilitates interaction of gp17 tail fibers with host LPS by conformational change in tail fibers orientation (*i.e.* perpendicular to cell surface). Tail fibers attachment to host receptors triggers genetic material injection. Presumably, the ejected proteins extend the short T7 tail forming a channel for the passage of the genetic material (González-García et al., 2015).

1.2.3 Bacterial receptors

The phage-host interaction is not only governed by RBP specificity. Indeed, the host receptor also plays an important role in the adsorption process since the receptor localization and density on the cell surface also influence this crucial interaction. In addition to outer membrane presence characterizing Gram-negative bacteria, Gram-positive and Gram-negative bacteria also differ in the thickness, uniformity and lipid/lipoprotein content of their cell wall. This in turn influences the bacterial receptors availability (Nobrega et al., 2018).

Receptors of Gram-positive cell wall

PG is an important component of the Gram-positive cell wall and is a common phage receptor (*e.g.* *Myoviridae* *Listeria* phage A511) (Dunne et al., 2018b). Also, phages can attach to teichoic acid that consists of polyol phosphate polymers with either ribitol or glycerol linked by phosphodiester bonds and is associated to the muramic acid of PG. Bacterial species differ in the substituent group on the polyol chain. Teichoic acids represent up to 50% of the total cell wall mass (Siegel et al., 2016) and can also be used as phage receptors with binding either to its backbone (*e.g.* *S. aureus* phage ϕ 11) or to its substituent group (*e.g.* *S. aureus* phage ϕ K) (Nobrega et al., 2018). Furthermore, phages can bind to proteins present in the Gram-positive cell wall (*e.g.* GamR protein in *B. anthracis* cell wall used for γ phage adsorption) (Nobrega et al., 2018). Cell wall polysaccharides can also serve as receptors (*e.g.* binding to the pellicle of *Lactococcus lactis*) (Nobrega et al., 2018). In rare cases, phages preying on Gram-Positive bacteria can bind to bacterial appendages such as flagellum (*e.g.* *B. subtilis* siphophage PBS1) (Dunne et al., 2018b). Less phage receptors have been identified in Gram-positive bacteria compared to Gram-negative since Gram-negative have been historically more studied.

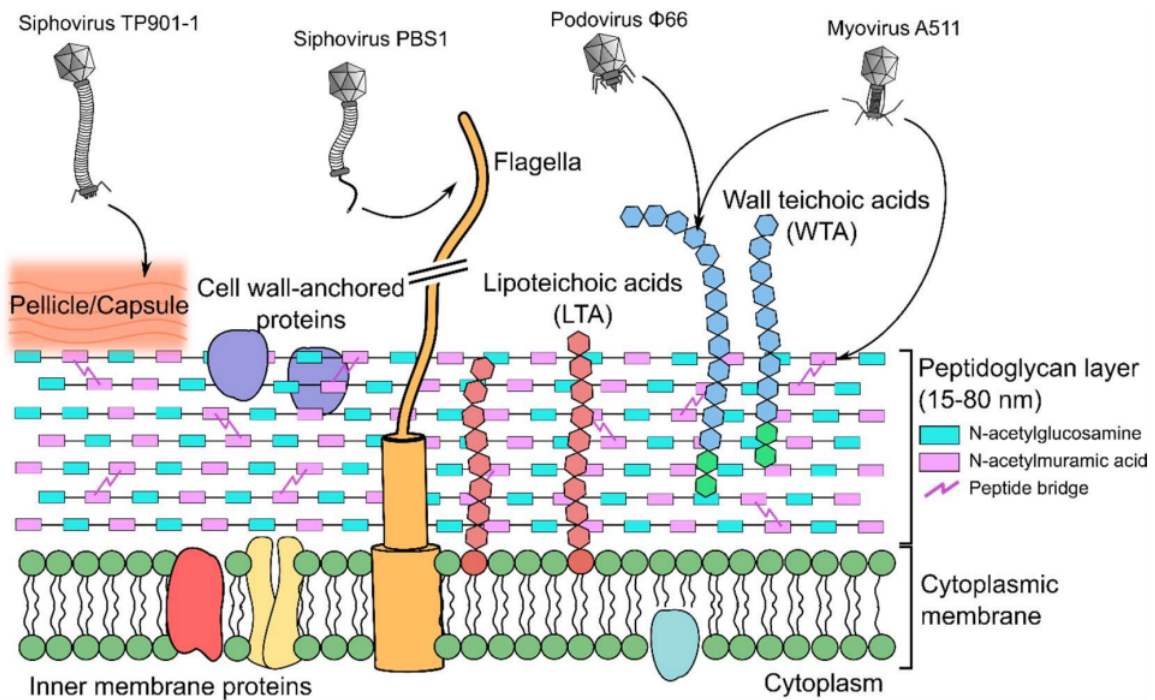


Figure 1.7: Host cell receptors of Gram-positive phages : Peptidoglycan (PG), teichoic acids (WTA) lipoteichoic acids (LTA) and polysaccharides such as pellicle layer and appendages such as flagella (Dunne et al., 2018b).

Receptors of Gram-negative cell wall

In Gram-negative bacteria, the PG layer is thinner and located inward of the outer membrane. These two layers are connected by Braun's lipoproteins. The inner layer of the outer membrane is mainly composed of phospholipids whereas the outer layer contains mostly lipopolysaccharides (Knirel and Valvano, 2011).

The main receptor in Gram-negative bacteria is LPS, which is a complex polymer composed of monosaccharides and fatty acids, in both smooth and rough forms. Although both forms possess core polysaccharide and highly polar lipid A, they differ in the presence (smooth) or absence (rough) of O-antigens. Furthermore, O-antigens are highly variable among species whereas the core of the LPS is generally more conserved. Consequently, phages targeting the core of LPS have generally a broader spectrum even in some cases, the O-antigen can prevent the binding to the core LPS (*e.g.* R-pyocin (Köhler et al., 2010)) (Rakhuba et al., 2010). Phages can recognize the O-antigen either with their tail fibers or tail spikes and they can hydrolyze the O-antigen thereby allowing the penetration of the tail. In some cases, the same phage can bind to different structures of the LPS core when interacting with different bacterial species (Nobrega et al., 2018). In addition, some phages can attach to outer membrane porins (*e.g.* myophage T4 uses either OmpC or LPS for its adsorption) (Washizaki et al., 2016). Finally, other transporter proteins can serve as receptors (*e.g.* coliphage DT57C uses vitamin B12 uptake protein BtuB98 for adsorption) (Nobrega et al., 2018).

Appendages such as flagella and pili can also function as phage receptors therefore increasing the chance of phage attachment and infection (*e.g.* *Caulobacter crescentus* phage phiCbK) (Guerrero-Ferreira et al., 2011). Some phages use their capsid and tail to attach to the flagellum, therefore leaving the distal part of the tail free for the adsorption to the cell surface (*e.g.* *Asticcacaulis biprosthecium* siphophages AcM4 and AcS2) (Nobrega et al., 2018).

Many Gram-negative bacteria produce external protective layers made of exopolysaccharides. Even if these layers may block access to cell wall receptors, they also might contain primary receptors for phage adsorption (*e.g.* Vi antigen present on the capsules of *Salmonella* spp. bound by myophage Vi01 (Pickard et al., 2010)) ; (Nobrega et al., 2018). This may be associated with the reversible depolymerization of the main chain of polysaccharides (*e.g.* endoglucosidase activity in *E. coli* phage K29) followed by the irreversible attachment of the phage to the bacterial receptors (Rakhuba et al., 2010).

1.3 Host cell lysis

The bacterial PG is made of glycan strands cross-linked by short peptides. This structure surrounds the cytoplasmic membrane and defines the bacterial shape. Its main function is to preserve cell integrity by withstanding the internal turgor pressure (5 atm for Gram-negative organisms; up to 50 atm for Gram-positive organisms (Seltmann and Holst, 2013)) (Vollmer et al., 2008). PG also serves as a scaffold for the anchorage of other cell envelope component such as proteins (Dramsı et al., 2008) or teichoic acids (Neuhaus and Baddiley, 2003). Any inhibition of PG synthesis (*e.g* by mutation, use of antibiotic) or its specific degradation results in cell lysis.

PG hydrolases family gathers numerous lytic enzymes acting upon bacterial PG and can be classified based on their origin. Autolysins are produced by bacteria for cell growth, division, maintenance and repair of the PG. Bacterial exolysins are secreted to kill bacteria of a different strain belonging to the same ecological niche (*e.g.* lysostaphin secreted by *Staphylococcus simulans* to lyse *S. aureus* (Kumar, 2008)). PG hydrolases are also used by most phages during the infection process (exceptions are filamentous phages which extrude from the host and ssDNA and RNA phages that lyse bacteria without any muralytic enzyme activity (Young et al., 2000)). For classical dsDNA phages, the degradation of the PG at the end of the replication cycle is made by endolysins that work in collaboration with holins. Both late genes are located in a lysis cassette in the late region of the phage genome (Nelson et al., 2012).

1.3.1 Overall process

At the end of the infection process, endolysins accumulate in a fully folded and active state in the bacterial cytosol. In most cases, endolysins do not possess any secretory signal, therefore lysis requires the activity of a small membrane protein termed holin (Wang et al., 2000). This protein perforates the plasma membrane and therefore allows the endolysin to gain access to the PG for its degradation (Fig. 1.8) (Dewey et al., 2010). Therefore, holins define the time of bacterial lysis (Cisek et al., 2017). For Gram-negative bacteria, lysis requires the additional action of the spanin complex allowing the disruption of the outer membrane (Berry et al., 2012).

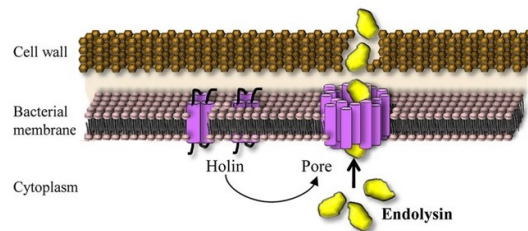


Figure 1.8: Endolysins and holins are produced at the end of the phage infection cycle. Holins form a pore in the bacterial membrane, allowing endolysins to reach the PG for its degradation (Gutiérrez et al., 2018).

1.3.2 Endolysins

Endolysins targeting Gram-positive bacteria feature a modular organization, generally separating enzymatic activity (N-terminal enzymatically active domain, EAD) and cell wall recognition (C-terminal cell wall binding domain, CBD) into distinct domain, connected by flexible linkers (Fenton et al., 2010). Some endolysins can feature multiple EADs and/or CBDs (Kashani et al., 2018).

Endolysins targeting Gram-negative bacteria are generally small single-domain globular proteins (Walmagh et al., 2013).

Enzymatically active domain

Regarding the bond they cleave in the PG, endolysins are classified into six categories : N-acetylglucosaminidase (cleaves the glycan bond located on the reducing side of the GlcNAc), N-acetylmuramidase (cleaves on the reducing side of the MurNAc sugar), transglycosylase (degrades the PG through the formation of a non-reducing N-acetyl-1,6-anhydromuramic acid and N-acetylglucosamine at the two ends), N-acetylmuramoyl-L-alanine amidase (cleavage of the amide bond between the MurNAc and the L-alanine moiety on the stem peptide), endopeptidase (hydrolysis of the peptide bonds in both the stem peptide and the interpeptide bridge) and Cysteine, histidine-dependent amidohydrolase/peptidase (CHAP, cleaves the PG amide and/or peptide bonds) (Broendum et al., 2018). These enzymatic activities are described in Fig. 1.9.

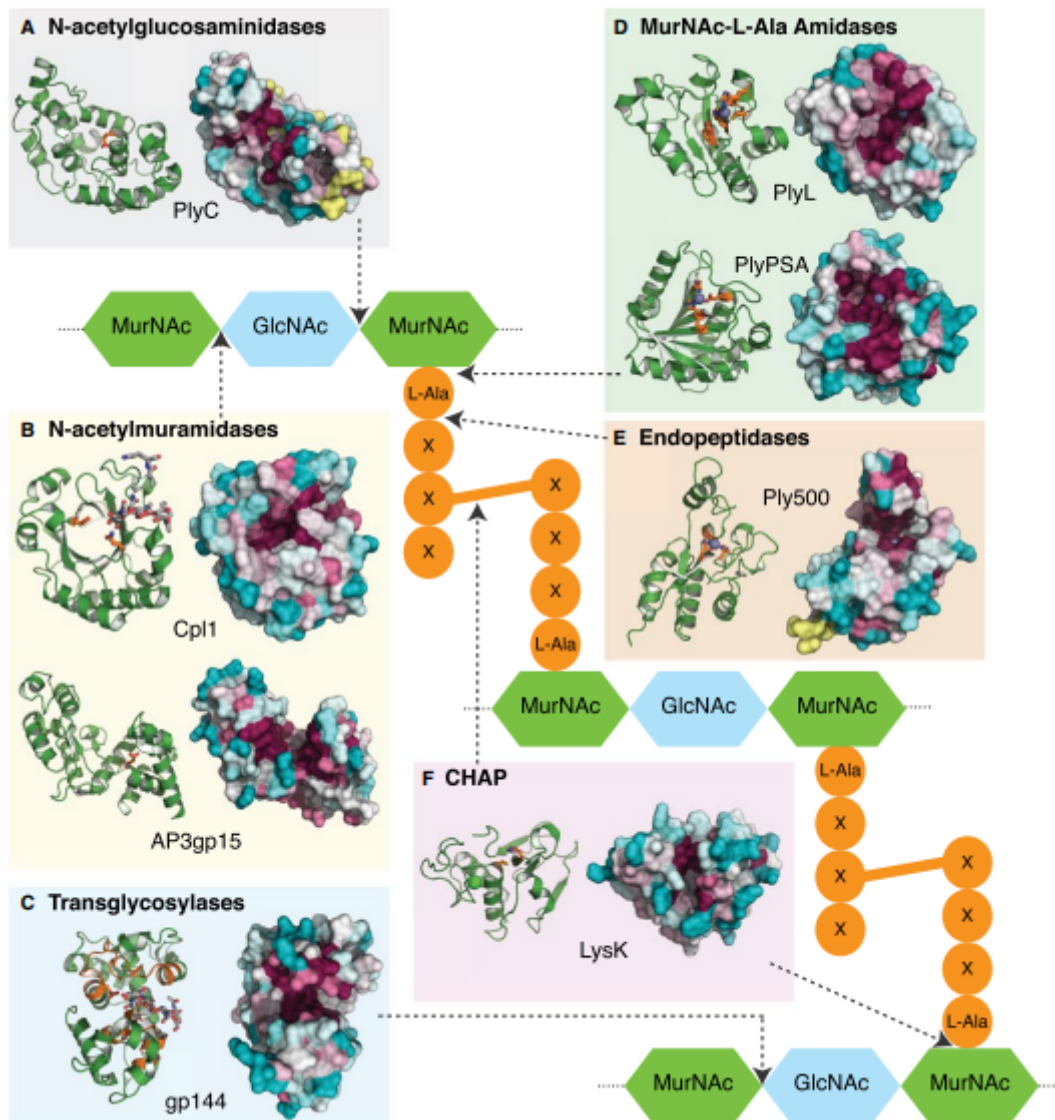


Figure 1.9: Representation of catalytic activity of EADs and examples for each type of EAD. A. N-acetylglucosaminidase B. N-acetylmuramidase C. Transglycosylase D. N-acetylmuramoyl-L-alanine amidase E. Endopeptidase F. CHAP. Abbreviations are : MurNAc - N-acetylmuramic acid ; GlcNAc - N-acetylglucosamine ; CHAP - cysteine, histidine-dependent amidohydrolase/peptidase ; L-alanine (Broendum et al., 2018).

Cell wall Binding Domain

The CBD confers specificity to endolysins by recognizing and binding to ligands within the cell wall (Fischetti, 2005). CBD can be specific to one or multiple bacterial genera or be restricted to the strain level (Broendum et al., 2018). Broad spectrum endolysins share similar CBDs as they recognize similar ligands within PG layer (López and García, 2004) while narrow spectrum endolysins likely bind specific PG ligands (Eugster et al., 2011). Many CBDs share similar sequence or structure. Therefore, they can be classified into four main classes that can be seen in Fig. 1.10. They are all characterized by varying degrees of repetition. How is this achieved and how does it improve activity remains unclear (Broendum et al., 2018).

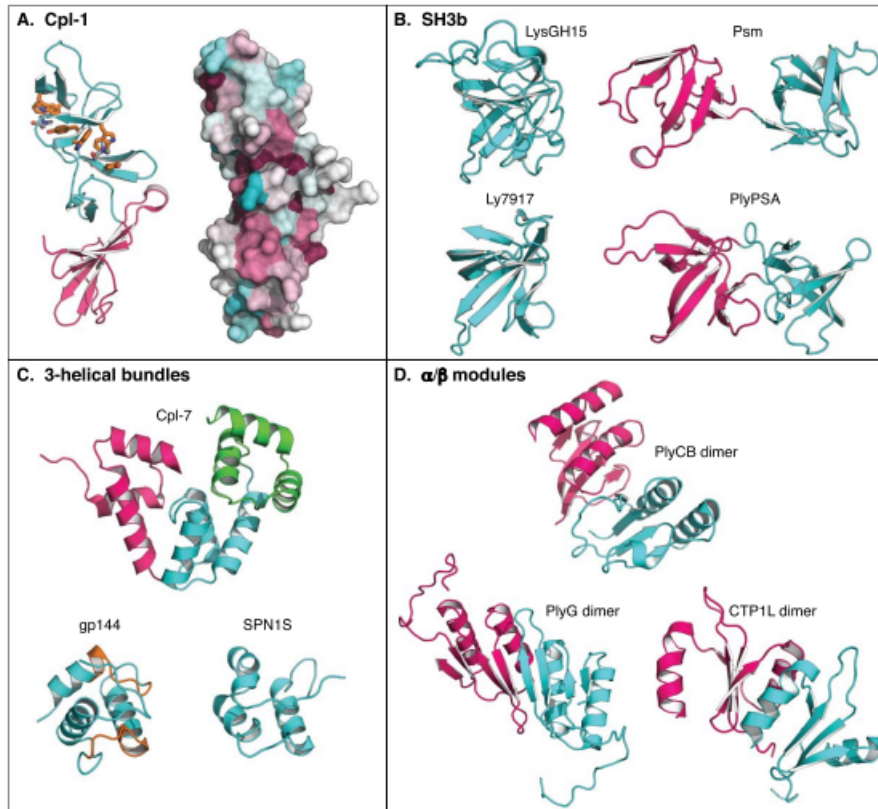


Figure 1.10: Various structural repeats used by CBDs and examples. A. Choline binding module B. SH3 domain C. Three-helical bundle D. α/β module (Broendum et al., 2018).

Choline binding module (ChBr) allows the binding to the choline-containing teichoic acid. It features an N-terminal super-helical moiety and a C-terminal β -sheet interacting with the EAD. SH3b domains generally form a characteristic SH3 β -barrel fold, which generally features between five and seven β -strands, often forming two antiparallel β -sheets, connected by linkers. This fold recognizes conserved ligands within a bacterial genus. The three-helix bundle is a widely used fold in CBDs but it can target diverse ligands. α/β multimers feature a secondary structure consisting of a central parallel or antiparallel four-stranded β -sheet flanked by two helices (Broendum et al., 2018).

Some endolysins have the ability to work without CBD, sometimes even EAD alone shows a higher lytic activity than full-length endolysin. This ability has been suggested to be related to the net charge of the EAD as positively charged EADs can bind the PG independently of the CBD (Low et al., 2011) ; (Kashani et al., 2018).

1.3.3 Holins

Holins are proteins that perforate the cytoplasmic membrane at the end of the infection cycle, allowing endolysins to get access to the PG.

The primary structure of holins is highly variable. However, differences in amino acid sequences among holins are not linked to differences in their function. As transmembrane proteins, all holins feature at least one hydrophobic transmembrane domain and a highly charged C-terminal hydrophilic domain. Also, holin encoding genes are mainly located in close proximity to the endolysin encoding genes and are relatively short proteins (ranging from 34 to 210 aa (Reddy and Saier Jr, 2013)) (Wang et al., 2000).

Holins can be classified into three classes regarding the number of transmembrane domains (TMDs) that they feature. Class I includes proteins featuring three TMDs and have more than 95 amino acid residues in length (*e.g.* *E. coli* phage λ S105 protein). Class II gathers holins having 65 to 95 aa and two TMDs (*e.g.* Lambdoid phage 21 S protein). Holins with only one TMD in the central region belong to class III (*e.g.* phage ϕ CP26F holin) (Wang et al., 2000).

The timing of lysis by holins is regulated through a dual-start model. In this model, the holin gene encodes both holin and anti-holin with opposite function differing by their terminal extension. Therefore, the time of bacterial lysis is determined by holin-antiholin ratio. Fig. 1.11 shows the S105/S107 example, studied in λ phage (White et al., 2011) ; (White et al., 2010).

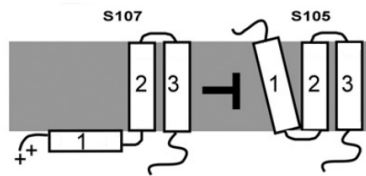


Figure 1.11: TMD1 of S107 is prevented from entering the membrane as long as the membrane is energized because of its two positive charges at its N terminus, which confers anti-holin character to S107 inhibiting S105. Adapted from (White et al., 2010).

The holin-anti-holin ratio is regulated at the translational level. When this ratio reaches a specific threshold, the holin nucleates to form rafts (close helical packing of TMDs that excludes lipid). At an allele-specific size, the rafts lose the ability to support the proton motive force, therefore inducing local depolarization. This triggers conformational change of holins that will finally allow endolysins release (White et al., 2010) ; (Tran et al., 2005). This process can be seen in Fig. 1.12.

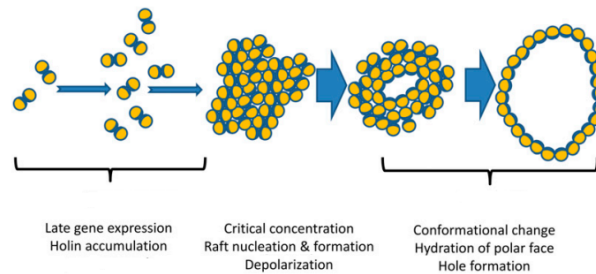


Figure 1.12: Holins, in orange (hydrophilic face in blue), accumulate as mobile homodimers, without perturbation of membrane integrity. When a critical concentration is reached the holins form RAFT, inducing local depolarization, triggering a conformational change in the holins, consisting in the hydration of the more hydrophilic surfaces of the holin and therefore forming a hole in the plasma membrane. Adapted from (White et al., 2011).

1.4 Deep-Purple, the Siphovirus, infecting *B. cereus*

1.4.1 The *B. cereus* group

B. cereus sensu lato (*s.l.*) bacteria are Gram-positive aerobic-to-facultative anaerobic, motile, spore-forming, rod-shaped bacteria (Bottone, 2010). The *B. cereus* group is composed of closely related bacteria with divergent pathogenic properties. Many of them are widely distributed in the environment with the main reservoir being soil where they are present as spores that can germinate and grow when brought in contact with organic matter, insect or animal hosts (Priest et al., 2004).

In the past, bacteria of this group were differentiated based on 16S rRNA sequences and characteristics such as the presence of virulence plasmid, colonial morphology, optimal temperature for cell growth and ill-related features. The most widely used classification is based on plasmid-borne determinants but the emergence of exceptional strains and the easy transfer or loss of plasmids is the main drawback for this classification and new classification methods are still controversial (Liu et al., 2015).

B. anthracis, the causative agent of anthrax, is an obligate pathogen that represents a threat to humans and herbivores, due to the presence of two large plasmids, pXO1 and pXO2 carrying major virulence factors. They consist of a tripartite toxin and an antiphagocytic capsule made of poly-D-glutamic acid. The loss of one of those plasmids results in an important decrease of virulence (Jang et al., 2011). Recently, *B. anthracis* has raised concern because of its potential as biological weapon for its ability to form heat resistant spores easily produced and infecting via the aerosol route (Rivera et al., 2010).

B. thuringiensis differentiates from the other members of the *B. cereus* group by the production of cytolytic toxins or δ -endotoxins. These toxins form intracellular proteins crystals that are toxic to a broad range of insect larvae, making *B. thuringiensis* the most widely used biological pesticide worldwide for agricultural purposes as well as for the control of disease-related insect vectors. These toxins are encoded by the *cry* or *cyt* genes, located on large plasmids. (Bravo et al., 2007)

B. cereus have been reported in an increasing number of foodborne intoxications, representing ca. 6% of all outbreaks reported within EU in 2017 (Authority et al., 2018). This is partly due the production of highly adhesive endospores able to spread to food matrices. Though, in most cases, the disease is self-limiting and of short duration even if some fatal cases have been reported in recent years (Stenfors Arnesen et al., 2008) ; (Dierick et al., 2005). *B. cereus sensu stricto* (*s.s.*) is associated with two forms of food poisoning : emetic and diarrhoeal. The emetic disease is an intoxication caused by the mainly plasmid-encoded cereulide toxin, a small ring-formed dodecadepsipeptide, inducing vomiting. The diarrhoeal syndrome is an infection caused by the ingestion of vegetative cells producing enterotoxins in the small intestine, putatively inducing diarrhoea by disruption of the integrity of the plasma membrane of intestinal epithelial cells (Stenfors Arnesen et al., 2008). Three toxins have been suggested as etiological agents of this syndrome : the pore-forming cytotoxins haemolysin BL (Hbl), the non hemolytic enterotoxin (Nhe) and the cytotoxin K (CytK) (Fagerlund et al., 2010).

While *B. cereus* bacteria were originally described as mesophilic organisms, with an optimal growth temperature ranging from 35°C to 40°C, *B. weihenstephanensis* are psychrotolerant strains having the ability to grow at 7°C but not at 43°C (Stenfors Arnesen et al., 2008). Refrigerated foods are therefore a potential niche for those bacteria. Consequently, the food industry is worrying because of the possible threat of emetic strains as foodborne pathogens (Thorsen et al., 2009).

The *B. cereus* group also gathers other bacteria including mostly non-pathogenic *B. mycoides* and *B. pseudomycoides* that distinguished from other members of the group because of their

rhizoidal colony shape and their fatty acid composition (Di Franco et al., 2002). *B. cytotoxicus* is a thermotolerant member of the *B. cereus* group that has so far only been isolated from vegetable products such as mashed potatoes, therefore potentially representing a threat for consumers (Lapidus et al., 2008).

New members of the *B. cereus* have been recently proposed, including *B. toyonensis*, *B. gaemokensis*, *B. manliponensis* and *B. bingmayongensis* among others (Liu et al., 2015).

1.4.2 Phages preying on *B. cereus*

Phages infecting *B. cereus* belong to the *Myoviridae*, *Siphoviridae*, *Podoviridae* and *Tectiviridae* families, the latter family gathering tailless icosahedral phages possessing an internal lipid vesicle that acts as a tail upon adsorption (Saren et al., 2005). To date, no filamentous phages, ssDNA or RNA phages have been reported to infect *B. cereus* (King et al., 2011). The main phage categories infecting *B. cereus* are transducing phages, phages with a chromosomal or plasmidial prophage state, γ -like phages and jumbo-phages (Gillis and Mahillon, 2014).

1.4.3 Phage Deep-Purple

Deep-Purple is a new member of the *Siphoviridae* family, isolated from a Belgian agricultural soil, that possesses an icosahedral head of 63 nm and a non-contractile tail of 165 nm (Hock et al., 2018) (see Fig. 1.13).

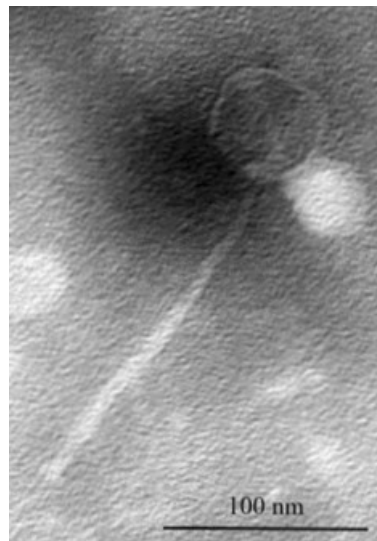


Figure 1.13: Transmission electron microscopy image of phage Deep-Purple negatively stained with 2% potassium phosphotungstate (Hock et al., 2018).

Deep-Purple shows lytic activity against 55% of emetic strains of the *B. cereus* group and its preferential host is the *Bacillus weihenstephanensis* LH002. Deep-Purple is able to perform a lytic cycle within 45 minutes and is stable over a broad range of pH and temperatures (Hock et al., 2018).

Deep-Purple genome

Deep-Purple dsDNA genome features 36,278 base pairs with a GC content of 38,36% and a coding density of 88% (Hock et al., 2018). It displays 40 putative coding DNA sequences (CDSs) with most of them being unique to Deep-Purple. 24 of those CDSs have no predicted function. Predicted CDSs has been categorized into four functional groups : structural related proteins (capsid, tail and minor structural proteins), DNA replication and transcription (DNA binding protein and transcriptional regulators), DNA packaging (terminase and portal proteins) and host lysis (putative endolysin) as shown in Fig. 1.14. No potential antimicrobial resistance gene and no tRNA has been identified for Deep-Purple, indicating that the phage relies completely on host machinery after infection.

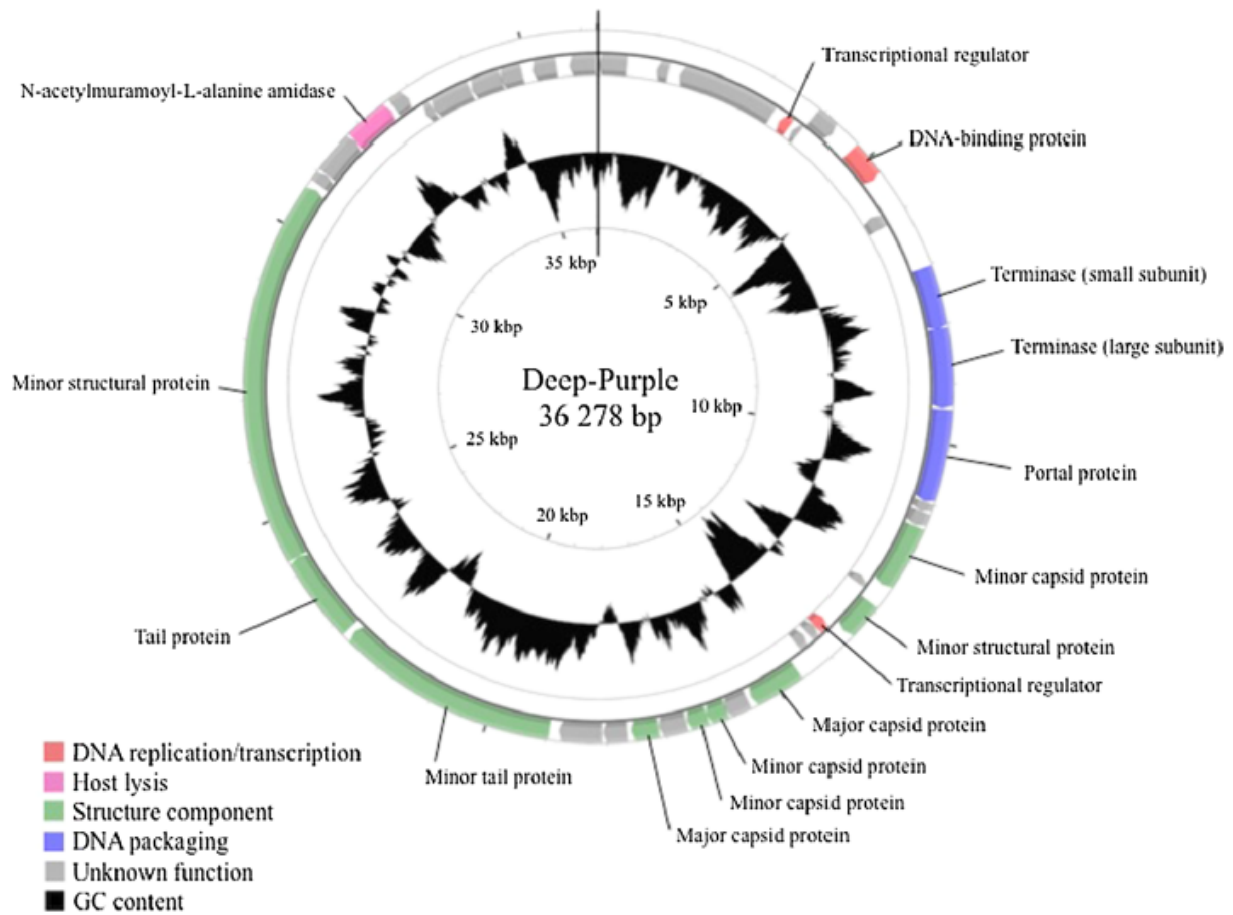


Figure 1.14: Circular map of the phage Deep-Purple. Circles display (1) CDSs on the forward strand, (2) CDSs on the reverse strand, (3) GC content. Putative functions were assigned to CDSs and are represented with a color code (Hock et al., 2018).

No lysogeny-associated genes could be identified in Deep-Purple genome. However, it cannot be excluded that Deep-Purple is a clear-plaque mutant of a plasmidial (extrachromosomal) prophage (Hock et al., 2018).

Objectives

For all phages, the infection process is initiated by the adsorption step, which is mediated by the irreversible attachment of Receptor Binding Proteins (RBP) to specific bacterial receptors. This specific interaction is responsible for the phage host range. At the end of phage infection cycle, phage endolysin degrades peptidoglycan (PG) in collaboration with holin that perforates the cytoplasmic membrane allowing the endolysin to reach PG. Deep-Purple phage is a member of the *Siphoviridae* family that comprises dsDNA phages with long non-contractile tails. Deep-Purple shows lytic activity against members of the *B. cereus* group, including emetic strains of *B. cereus* and *B. weihenstephanensis*.

This Master thesis aims to better understand the lifestyle of Deep-Purple with a particular focus on the proteins involved in its adsorption (*i.e.* RBP) and lysis (*i.e.* endolysin and holin)

Deep-Purple was isolated through an enrichment of the soil sample with bacteria that potentially contain prophages (dormant form of phage). As Deep-Purple does not have a typical lytic behavior and considering the method used for its isolation, it is thought that Deep-Purple could be a lytic variant of a prophage initially present in the strains used to enrich the sample. To test this hypothesis, a PCR screening will be performed to detect the presence of Deep-Purple in a representative set of strains of the *B. cereus* group.

Potential RBP, endolysin and holin candidates will be selected by bioinformatics. Potential RBP candidate(s) will be cloned with fusion to GFP and expressed in *E. coli*. After evaluation of the proteins solubility, the proteins will be purified in native conditions. The proteins will finally be tested in a cell wall decoration assay in order to evaluate if the proteins are able to adsorb to the bacterial host.

Endolysin candidate(s) will also be cloned in *E. coli* and purified in native conditions. The endolysin activity will then be tested in a turbidity reduction assay. Once identified, the endolysin will be characterized regarding its biochemical properties (pH, salt concentration and temperature stability) and host range. The complete endolysin and its CBD, which is the part of the endolysin responsible for the attachment to the bacteria, will be fused to a GFP to assess its binding capacity and its potential as detection tool. Finally potential holin candidate(s) will be cloned and expressed in *E. coli*. Holin activity will be characterized by monitoring the decrease of optical density during protein induction and by fusing to GFP to assess the protein localization during induction.

Chapter 2

Material and methods

2.1 Strains and plasmids

Bacterial strains used in this work are listed in Table 2.1.

All clonings were done with the pET30 vector. It contains a T7 promoter which is recognized by *E.coli* T7 RNA polymerase and is tightly regulated by the presence of LacI that binds to LacO and inhibits transcription. This inhibition can be relieved by the addition of IPTG. pET30 vector also features His tags that allow recombinant proteins purification and a Kanamycin resistance gene for selection after transformation. The genetic map of pET30 vector is represented in Fig. 2.1.

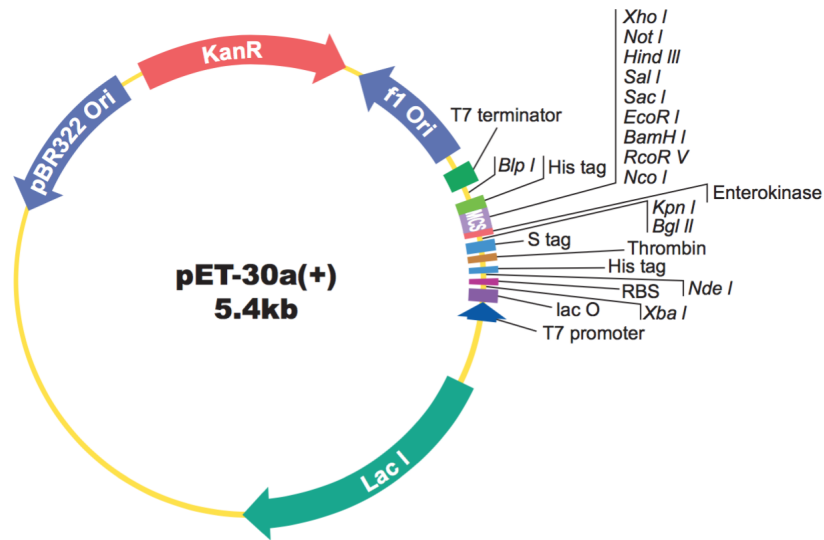


Figure 2.1: pET30 vector. Abbreviations are : KanR - Kanamycin resistance ; pBR322 Ori - origin of replication (originally presents on pBR322 vector) ; LacI - operon lactose inhibitor ; T7 promoter - promoter for phage T7 RNA polymerase ; LacO - operon lactose operator ; RBS - ribosome binding site ; His tag - histidine tag ; thrombin - thrombin recognition and cleavage site ; enterokinase - enterokinase recognition and cleavage site ; MCS - multiple cloning site ; T7 terminator - transcription terminator of phage T7 RNA polymerase ; f1 ori - f1 phage origin of replication. Adapted from Genscript.

Table 2.1: Bacterial strains used in this work. ○ strains are used for PCR screening for Deep-Purple presence in *B.cereus* strains ● strains are used for the endolysin characterization.

Bacterial strain	Relevant features	Origin
<i>E. coli</i> 10β	Competent cells	Biolabs
<i>E. coli</i> BL21(DE3)	Competent cells	Biolabs
<i>E. coli</i> 10β[pUC18]	Containing GFP gene	BBGM
<i>E. coli</i> BL21(DE3)[pET30::gp32]	Containing the recombinant plasmid pET30::gp32	This work
<i>E. coli</i> BL21(DE3)[pET30::gp32_lytic]	Containing the recombinant plasmid pET30::gp32_lytic	This work
<i>E. coli</i> BL21(DE3)[pET30::GFP(link)::gp32_CBD(149)]	Containing the recombinant plasmid pET30::GFP(link)::gp32_CBD(149)	This work
<i>E. coli</i> BL21(DE3)[pET30::GFP(link)::gp32_CBD(180)]	Containing the recombinant plasmid pET30::GFP(link)::gp32_CBD(180)	This work
<i>E. coli</i> BL21(DE3)[pET30::GFP(link)::gp32_CBD(196)]	Containing the recombinant plasmid pET30::GFP(link)::gp32_CBD(196)	This work
<i>E. coli</i> BL21(DE3)[pET30::GFP(link)::gp32]	Containing the recombinant plasmid pET30::GFP(link)::gp32	This work
<i>E. coli</i> BL21(DE3)[pET30::gp28::(link)GFP]	Containing the recombinant plasmid pET30::gp28::(link)GFP	This work
<i>E. coli</i> BL21(DE3)[pET30::GFP(link)::gp28_mid]	Containing the recombinant plasmid pET30::GFP(link)::gp28_mid	This work
<i>E. coli</i> BL21(DE3)[pET30::GFP(link)::gp26]	Containing the recombinant plasmid pET30::GFP(link)::gp26	This work
<i>E. coli</i> BL21(DE3)[pET30::GFP(link)::gp29_1200]	Containing the recombinant plasmid pET30::GFP(link)::gp29_1200	This work
<i>E. coli</i> BL21(DE3)[pET30::GFP(link)::gp29_1400]	Containing the recombinant plasmid pET30::GFP(link)::gp29_1400	This work
<i>E. coli</i> BL21(DE3)[pET30::gp30::(link)GFP]	Containing the recombinant plasmid pET30::gp30::(link)GFP	This work
<i>E. coli</i> BL21(DE3)[pET30::gp30]	Containing the recombinant plasmid pET30::gp30	This work
MC67 ○	Emetic <i>B. weihenstephanensis</i>	Thorsen
BtB2-4 ○	Emetic <i>B. weihenstephanensis</i>	MIAE
LH002 ○ ●	Emetic <i>B. weihenstephanensis</i>	MIAE
Cer074 ○	Emetic <i>B. weihenstephanensis</i>	MIAE
Cer057 ○	Emetic <i>B. weihenstephanensis</i>	MIAE
TIAC303 ○	Emetic <i>B. cereus</i>	Sciensano
Cer075 ○	Emetic <i>B. cereus</i>	Mahler
5969c ○	Emetic <i>B. cereus</i>	Sciensano
Cer039 ○	Emetic <i>B. cereus</i>	Sciensano
TIAC181 ○	Emetic <i>B. cereus</i>	Sciensano
F4810/72 ○ ●	Emetic <i>B. cereus</i>	Turnbull
AFRH5 ○	Emetic <i>B. cereus</i>	MIAE
AND1284 ○	Emetic <i>B. cereus</i>	A.Wicks
ISO75 ○	Emetic <i>B. cereus</i>	I. Swiecicka
5958b ○	Emetic <i>B. cereus</i>	Sciensano
5975c ○	Emetic <i>B. cereus</i>	Sciensano
MHI1699 ○	Emetic <i>B. cereus</i>	Moravek
ISP3971 ○	Emetic <i>B. cereus</i>	Sciensano
H3081.97 ○ ●	Emetic <i>B. cereus</i>	Hoffmaster
Cer072 ○	Emetic <i>B. cereus</i>	MIAE

NVH1230-88 ◦	<i>B. cereus</i>	MIAE
NVH883-00 ◦	<i>B. cereus</i>	MIAE
MHI1497 ◦	<i>B. cereus</i>	Ludwig-Maximilians University
B16 ◦	<i>B. cereus</i>	MIAE
Si0270 ◦	<i>B. weihenstephanensis</i>	MIAE
Si0239 ◦ ●	<i>B. weihenstephanensis</i>	MIAE
Si0120 ◦	<i>B. thuringiensis</i>	MIAE
Si0188 ◦	<i>B. thuringiensis</i>	MIAE
T50 001 ◦	<i>B. thuringiensis</i>	MIAE
T65 001 ◦	<i>B. thuringiensis</i>	MIAE
VKM-B446 ◦	<i>B. thuringiensis</i>	All-Russian collection
TIAC468 ◦	<i>B. cereus</i>	Sciensano
TIAC896 ◦	<i>B. cereus</i>	Sciensano
VD014 ◦	<i>B. cereus</i>	MIAE
VD021 ◦ ●	<i>B. cereus</i>	MIAE
VD078 ◦	<i>B. cereus</i>	MIAE
Bt5 ◦	<i>B. thuringiensis</i>	MIAE
DBT090 ◦	<i>B. thuringiensis</i>	MIAE
DBT242 ◦	<i>B. thuringiensis</i>	MIAE
TIAC76 ◦	<i>B. cereus</i>	Sciensano
ATCC10987 ●	<i>B. cereus</i>	MIAE
WSBC10202 ●	<i>B. weihenstephanensis</i>	MIAE
WSBC10204 ●	<i>B. weihenstephanensis</i>	Lechner S.
HD4 ●	<i>B. thuringiensis</i>	<i>Bacillus</i> Genetic Stock Center (USA)
GBJ001 ●	<i>B. thuringiensis</i>	MIAE
NVH391-98 ●	<i>B. cytotoxicus</i>	INRA
BcA20 086 ●	<i>B. cytotoxicus</i>	MIAE
KBS 2-12 ●	<i>B. mycoides</i>	MIAE
KNC2-18 ●	<i>B. mycoides</i>	MIAE
Bs168 ●	<i>Bacillus subtilis</i>	MIAE
Si0100 ●	<i>Bacillus amyloliquefaciens</i>	MIAE
ATCC10716 ●	<i>Bacillus licheniformis</i>	BCCM
Si0279 ●	<i>Bacillus pumilus</i>	MIAE
Liv001 ●	<i>Listeria ivanovii</i>	MIAE
ATCC12228 ●	<i>Staphylococcus epidermis</i>	BCCM
Si003 ●	<i>Bacillus megaterium</i>	MIAE
ATCC8739 ●	<i>E. coli</i>	BCCM

2.1.1 Competent cells preparation

Competent cells (NEB) stocked at -80°C were spread on LB plate and incubated O/N at 37°C . A single colony was transferred to 10 mL of LB medium and grown O/N at 37°C and 180 rpm. Then 1 L of LB was inoculated with 10 mL of the O/N culture and grown at 37°C with shaking. When the OD_{600} reached 0.35-0.4, cultures were put on ice. After chilling 20 to 30 min, the culture was split in 4 by pouring about 250 mL into ice cold centrifuge bottles. Cells were harvested by centrifugation at 3000 g for 15 min at 4°C . After discarding supernatant, pellets were resuspended in 100 mL of ice cold MgCl_2 . All fractions were then combined into one centrifuge bottle to harvest cells by centrifugation at 2000 g for 15 min at 4°C . Pellet were resuspended in 200 mL of ice cold CaCl_2 . This suspension was kept on ice for at least 20 min before harvesting cells by centrifugation at 2000 g for 15 min at 4°C . Cells were then resuspended in 50 mL of ice cold 85 mM CaCl_2 supplemented with glycerol (15%). The suspension was transferred into an ice cold Falcon and cells were harvested by centrifugation at 1000 g for 15 min at 4°C . The pellet was after resuspended in 2 mL of ice cold 85 mM CaCl_2 , 15% glycerol. 50 μl aliquots were finally prepared and stored at -80°C .

2.2 Culture media and conditions

All cell cultures were grown in LB medium (Biorad) with the exception of *Listeria ivanovii* Liv001 that was grown in BHI medium (Biorad). Media compositions are detailed in Table 2.2. Recovering of cells after heat shock during transformation was done in SOC medium of which composition is detailed in Table 2.3. Top agar was prepared with 5 mL LB medium supplemented with 6% agar and ions necessary for phage lytic activity (0.361 g/L of MgSO_4 and 0.333 g/L of CaCl_2). Kanamycin (Sigma Aldrich) stock solutions were prepared and sterilized by filtration through 0,22 μm filters and stored at 20°C . When required for selection, Kanamycin was added to LB medium at final concentration of 50 $\mu\text{g}/\text{mL}$. *Bacillus* spp. were grown at 30°C and 120 rpm whereas *L. ivanovii*, *E. coli* and *Staphylococcus epidermis* were grown at 37°C and 180 rpm.

Table 2.2: Growth media used in this work.

Component	Concentration (g/L)	
	LB	BHI
Yeast Extract	5	-
Tryptone	10	-
NaCl	5	5
Glucose	-	2
Agar (if solid)	15	15
Calf Brain, infusion of	-	200
Beef Heart, infusion of	-	250
NaHPO_4	-	0.5
Distilled water	up to 1L	

Table 2.3: Soc medium composition.

Component	Origin	Concentration (g/L)
Yeast Extract	Sigma	5
Tryptone	Sigma	20
NaCl	VWR Chemicals	0.584
MgSO_4	VWR Chemicals	2.467
Glucose	VWR Chemicals	3.603
KCl	VWR Chemicals	0.186
MgCl_2	VWR Chemicals	0.952
Distilled water		up to 1L

2.3 Molecular methods

2.3.1 Cloning

Polymerase chain reaction

All DNA amplification were done by PCR. Primers used in this work are listed in Table 2.7. The primers were solubilized in sterile water after centrifugation at 10 000 rpm during 10 min to obtain a concentration of 100 pmol/ μ l. After 10 min of incubation at RT, aliquots were prepared by diluting 10 times in sterile water.

All routine PCR were done with the *OneTaq*[®] polymerase (NEB) whereas the high-fidelity *Q5*[®] polymerase (NEB) was used for DNA amplification.

Table 2.4: PCR programs used with *Q5*[®] polymerase and *OneTaq*[®] polymerase. T_m were calculated with NEB T_m Calculator.

	Step	<i>Q5</i> [®] polymerase		<i>OneTaq</i> [®]	
35x	Initial denaturation	98°C	30 s	95°C	5 min
	Denaturation	98°C	10 s	95°C	30 s
	Hybridization	Primer-dependent	30 s	Primer-dependent	30 s
	Extension	72°C	30 s/kb	68°C	1 min/kb min
	Final extension	72°C	2 min	68°C	5 min
	Hold	10°C	∞	10°C	∞

Agarose gel electrophoresis

Agarose gel and TAE buffer composition are detailed in Table 2.5 and Table 2.6. Electrophoresis was run for 30 min at 100V and the DNA fragments were revealed under the UV light thanks to the *Quantity One*[®] software (Bio-Rad). While *OneTaq*[®] buffer contains an inert tracking dye, 2 μ l of Purple loading dye (NEB) was added to 5 μ l of PCR product when using *Q5*[®] polymerase.

Table 2.5: Agarose gel composition.

Component	Quantity	Origin
TAE Buffer (0.5x)	100 mL	Lab preparation
Agarose	0.8 g	ChemCruz
EtBr	5 μ L	ICN biomedical

Table 2.6: Composition of the TAE Buffer stock solution.

Component	Quantity	Origin
Tris-(hydroxymethyl)-aminomethane	242 g	Merck
NaOAc.3H ₂ O	57 mL	Merck
EDTA (0.5M; pH8)	100 mL	Sigma-Aldrich
Distilled water	up to 1 L	

Table 2.7: Primers used in this work. Direction and restriction enzymes are included in primer names. F = forward primer. R = Reverse primer. Restriction sites in bold type.

Primer	Primer Sequence (5' → 3')	DNA target	Construction	Product size (bp)
gp32_NcoI_F	TATCCATGGACAAAAAGATTTTAGATATTCACATCAC	gp32	pET30::gp32	786
gp32(149)_EagI_F	CCGGGCGGCTTATCCTTGTGACATATG	gp32	pET30::gp32_lytic	540
gp32(180)_EagI_F	AACGGCCGGTTAGCTTTCTATATGGAACCTCTT	gp32	pET30::GFP(link)::gp32_CBD(149)	260
gp32(196)_EagI_F	AACGGCCGGTACAGTAGTTGAAAGGGC	gp32	pET30::GFP(link)::gp32_CBD(180)	249
gp32[Nier]_EagI_F	AACGGCCGAAAAAGATTTTAGATATTTACATCAC	gp32	pET30::GFP(link)::gp32_CBD(196)	201
gp32_XhoI_R	ATTACTCGAGTTAAATGATTTTACGTTATCGCCA	gp32	pET30::GFP(link)::gp32	787
gp32_XhoI_R'	GGGGCTCGAGTTAAATGATTTTACGTTATCGCCA	gp32	pET30::GFP(link)::gp32_CBD(180)	786
gp32(180)_XhoI_R	CCTTCTCGAGTTAAACACCTTTTGTACCTAATA	gp32	pET30::GFP(link)::gp32_CBD(196)	201
gp28_EagI_F	TTGGGCGGTCAACATTCAAATCAACG	gp28	pET30::GFP(link)::gp32	787
gp28_mid_EagI_F	TACGGCGGAAACAACATGAGACAACTAGGA	gp28	pET30::GFP(link)::gp32_CBD(149)	260
gp28_BstBI_F	GGTTTTCGGAATCAACATTCAAATCAACG	gp28	pET30::GFP(link)::gp32	540
gp28_XhoI_R	TATACTCGAGTTACTTATAAGGCTCCTTGT	gp28	pET30::GFP(link)::gp28_mid	1488
gp28_KpnI_R	AAGGTACGCTTATAACGCTCCTTGTAAAG	gp28	pET30::GFP(link)::gp28	672
gp26_EagI_F	AACGGCGGTTAAACTGACAGATACTAATTTAG	gp26	pET30::GFP(link)::gp26	1482
gp26_XhoI_R	AACCTCGAGTTAATCTAAAGCGTTAAAGAGC	gp26	pET30::GFP(link)::gp26	718
gp29_1200_EagI_F	ATCGGCGGATTACTCAAAATTAAGGCTCAAGG	gp29	pET30::GFP(link)::gp26	718
gp29_1400_EagI_F	TACGGCGGAGCATGACAGACTCGAG	gp29	pET30::GFP(link)::gp29_1200	2892
gp29_XhoI_F	TATACTCGAGTTAATTTGCGCTAAAGAACCTC	gp29	pET30::GFP(link)::gp29_1400	2292
gp30_NcoI_F	AAACCATGGAAATTTCAAAAAGAACTACTAC	gp30	pET30::gp30	210
gp30_EcoRI_R	AAAAGAATCTTATCTCTCTGCTTATCTT	gp30	pET30::gp30::(link)GFP	210
gp30_EcoRI_R'	AAAAGAATCTTCTCTGCTTATCTTCTT	gp30	pET30::gp30::(link)GFP	210
GFP_KpnI_F	TTGGGTACCAAGGAGAGAACTTTTCACTGGAG	GFP	pET30::GFP(link)::gp28	747
GFP_EcoRI_F	TTCCGAATTCAAAGGAGAACTTTTCACTGGAG	GFP	pET30::GFP(link)::gp28_mid	747
			pET30::GFP(link)::gp32_CBD(149)	
			pET30::GFP(link)::gp32_CBD(180)	
			pET30::GFP(link)::gp32_CBD(196)	
			pET30::GFP(link)::gp32	
			pET30::GFP(link)::gp26	
			pET30::GFP(link)::gp28	
			pET30::GFP(link)::gp29_1200	
			pET30::GFP(link)::gp29_1400	
GFP_EagI_link_R	TTGGGCGGTCCACTACCTGTATCCACTACCTTTGTAGAGCTCATCCATGCC	GFP	pET30::GFP(link)::gp28_mid	747
			pET30::GFP(link)::gp32_CBD(149)	
			pET30::GFP(link)::gp32_CBD(180)	
			pET30::GFP(link)::gp32_CBD(196)	
			pET30::GFP(link)::gp32	
			pET30::GFP(link)::gp26	
			pET30::GFP(link)::gp28	
			pET30::GFP(link)::gp29_1200	
			pET30::GFP(link)::gp29_1400	
GFP_EcoRI_link_F	TATAGAATTCGGGTAGTGGATCAGGTAGTGGAAGAGGAGAACTTTTCACTGGAG	GFP	pET30::gp30::(link)GFP	750
GFP_KpnI_link_F	AAGGTACGGGTAGTGGATGATGAGGTAGTGGAAGAGGAGAACTTTTCACTGGAG	GFP	pET30::gp30::(link)GFP	750
GFP_EagI_Stop_R	TACGGCGGTATTTGTAGAGCTCATCCATGCC	GFP	pET30::gp28::(link)GFP	750
Seq_pET30_F	ATACCCACGCCGAAACAA	MCS	-	insert + 359
Seq_pET30_R	CCGAGATAGTCTCCTCTTCAG	MCS	-	insert + 359

Plasmid extraction and DNA purification

The GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich) or QIAquick® gel extraction kit (Qia-gen) were used for plasmid extraction.

The GenElute™ PCR Clean-up kit (Sigma-Aldrich) was used for all DNA purification after PCR and restriction.

Restriction

Restrictions were carried out using the restriction enzymes (NEB) listed in Table 2.8. A standard restriction mix is detailed in Table 2.9. DNA was then purified after 4 h of incubation and DNA concentration was finally measured with a Nanodrop ND-1000 spectrophotometer (Isogen Life Sciences).

Table 2.8: Restriction enzymes used in this work.

Component	Restriction site	Buffer	Temperature of incubation
<i>EagI</i> -HF	CGGCCG	Cutsmart® Buffer	37°C
<i>NcoI</i> -HF	CCATGG	Cutsmart® Buffer	37°C
<i>XhoI</i>	CTCGAG	Cutsmart® Buffer	37°C
<i>BstBI</i>	TTCGAA	Cutsmart® Buffer	65°C
<i>KpnI</i> -HF	GGTACC	Cutsmart® Buffer	37°C
<i>EcoRI</i> -HF	GAATTC	Cutsmart® Buffer	37°C

Table 2.9: Composition of standard restriction mix.

Component	Quantity
DNA	~ 1 µg
Cutsmart® Buffer 10x	5 µl
Restriction enzyme 1	1 µl
Restriction enzyme 2	1 µl
Distilled water	to 50 µl

Ligation

Ligation of inserts into pET30 vector was done with the T4 DNA ligase (Promega). The composition of a standard ligation mix is detailed in Table 2.10. The ligation reactions were performed O/N at 4°C. The amounts of vector and insert for the ligation mix were calculated using NEBioCalculator™ in order to obtain a vector:insert molar ratio of 1:7.

Table 2.10: Composition of standard ligation mix.

Component	Origin	Quantity
T4 DNA ligase	Promega	0.5 µl
T4 DNA ligase Buffer 10x	Promega	1 µl
DNA (vector + insert)		8.5 µl

Transformation

2 μl of ligation product were added to 50 μl of competent 10 β cells. After 30 min of incubation on ice, cells were thermally shocked at 42°C for 30 sec. Then, tubes were transferred to ice and incubated for 5 min before adding 950 μl of SOC medium. Tubes were after placed horizontally in the incubator and incubated at 37°C and 180 rpm for 1 h. Thereafter, 100 μl of our cell cultures were spread at 10⁰ and 10⁻¹ concentration on agar plates containing Kanamycin for the selection of clones that have taken the recombinant plasmid. The following day, a PCR screening was performed to check if clones have acquired the recombinant plasmid. The same protocol was used for transformation into the expression strain BL21(DE3).

Sequencing

The positive clones revealed by the PCR screening were extracted following the protocol detailed earlier to be sequenced. To do so, 1.5 mL Eppendorf filled with 7.5 μl of the extracted plasmid and 2.5 μl of primer were sent to Macrogen Europe B.V.

2.3.2 PCR screening for Deep-Purple presence in *Bacillus* strains

To detect the presence of Deep-Purple as prophage in *B. cereus s.l.* strains, PCR were done with the *OneTaq*[®] polymerase (NEB) with DNA extracted with the Wizard[®] genomic DNA extraction kit (Promega) and primers listed on Table 2.11, designed to hybridize on different Deep-Purple genes. After running PCR, agarose gel electrophoresis was performed as previously explained.

Table 2.11: Primers used for PCR Screening for Deep-Purple presence in *Bacillus* strains.

Primer	Primer Sequence (5' → 3')	DNA target	Product size (bp)
Purple_screen_1_F	GGAGCAGGTGAGCAATGTTTA	gp26	857
Purple_screen_1_R	GCCCCTATCATCTAATTAAAGACCTC	gp26	857
Purple_screen_2_F	ATGAAAATGGAAGGGGATGA	gp02	1017
Purple_screen_2_R	CTTGAAAGAGGAAAAGAACCCTT	gp02	1017
Purple_screen_3_F	CATGAAAAAAGACGGTTCTCCT	gp16-17-18	943
Purple_screen_3_R	CCCCGCCTTACACATTAATGTA	gp16-17-18	943
Purple_screen_4_F	GGGGATTTTCAATATCACCAC	gp36-37	1225
Purple_screen_4_R	CTCATGGTTTTCTTCCCCTTATC	gp36-37	1225

2.4 Recombinant protein expression

2.4.1 Induction

5 mL of O/N culture of *E. coli* BL21(DE3)[pET30::protein_of_interest] were inoculated in 50 mL of LB supplemented with Kanamycin and the culture was grown at 37°C with vigorous shaking (180 rpm) until the OD₆₀₀ reaches 0.5-0.8. Just prior induction, 1 mL of sample was taken for the solubility assessment as non induced control and kept at 4°C. IPTG (Omega Biotech) was then added at a final concentration of 1 mM and the culture was grown at 28°C and 180 rpm. After 5 h of incubation, 2 samples of 1 mL were taken for the solubility assessment as induced control. The rest of the culture was collected into 50 mL Falcons. All the samples were centrifuged at 4000 g for 15 min at 4°C. All pellets were stored at -20°C after discarding supernatant.

2.4.2 Solubility assay

1 mL samples were thawed on ice for 30 min. Pellets were then resuspended in 80 μ L of NPI10 buffer of which the composition is detailed in Table 2.12 and 1 mg/mL lysozyme and 1x protease inhibitor cocktail (Sigma) were added. Samples were then incubated at 10°C with shaking for 30 min. 3 unit/mL Benzonase[®] nuclease (Sigma-Aldrich) was after added and samples were incubated at 10°C with shaking for 15 min. Thereafter, one of the induced control was centrifuged at maximum speed for 30 min at 4°C to separate the supernatant (*i.e.* the soluble fraction) from the pellet (*i.e.* the insoluble fraction). All fractions (*i.e.* the non-induced control, the induced control, the soluble fraction and the insoluble fraction) were kept for SDS-PAGE analysis.

Table 2.12: Composition of NPI10 buffer. Sterilized by filtration on 0.22 μ m filters.

Component	Origin	Quantity
NaH ₂ PO ₄	Merck	690 mg
NaCl	VWR Chemicals	1.75 g
Imidazole	Sigma-Aldrich	68 mg
NaOH/HCl	VWR Chemicals	Adjust to pH 8
Distilled water		to 100 mL

2.4.3 Protein gel electrophoresis

The samples were loaded onto a Mini-PROTEAN[®] TGX[™] Precast protein gel (Bio-Rad) according to the following protocol :

1. Add DTT (Roche) to Laemmli buffer (Bio-Rad) to obtain a final concentration of 100 mM.
2. Add 20 μ L of DTT-Laemmli buffer to 20 μ L of sample and heat at 95°C for 5 min.
3. Prepare 700 mL of running buffer by diluting 10 times the 10x TGS Buffer (Bio-rad).
4. Remove carefully the comb from the gel. Rinse with distilled water and remove the green tape at the bottom.
5. Place the gel in the cassette with the short side inside.
6. Fill the inner chamber with the buffer up to the top of the wells.
7. Load 20 μ L of samples and 2 μ L of Ladder (Color Prestained Protein Standard, Broad Range (10-250 kDa), NEB)
8. Insert the cassette in the tank and fill it with the rest of running buffer.
9. Run the electrophoresis for 30 min at 200V.
10. Remove the gel and transfer in the square petri dish. Rinse the tank with distilled water.
11. Fix the gel during at least 3 h in the fixation solution (composition detailed in Table 2.13).
12. Wash the gel 3 times 30 min with MilliQ[®] water.
13. Add the coloration solution (composition detailed in Table 2.13) and shake for 1 h.
14. Let color for 1 to 3 days in blue solution (composition detailed in Table 2.13).
15. Discolor with MilliQ[®] water.

Table 2.13: Composition of fixation, coloration and blue solutions.

Component	Origin	Fixation	Coloration	Blue
		Percentage	Percentage	Percentage
Ethanol	CWR Chemicals	50% (v/v)	-	-
Methanol	Merck	-	34% (v/v)	34% (v/v)
(NH ₄) ₂ SO ₄	VWR Chemicals	-	17% (w/v)	17% (w/v)
H ₃ PO ₃	Riedel-de Haën	2% (v/v)	3% (v/v)	3% (v/v)
Bleu de Coomassie R-250	Merck	-	-	40% (w/v)

2.4.4 Purification under native conditions

Section Soluble recombinant proteins were purified by affinity chromatography on a nickel column, thanks to the six-histidine tag located at the N-terminal end of the protein, using the Ni-NTA Spin Kit from Qiagen.

Pellets were treated as described above in 2.4.2 with 10x concentration factor. The resulting lysate was then filtrated (0.45 μ m filter).

The column was equilibrated with 700 μ L of NPI10 buffer (composition detailed in Table 2.12) by centrifugation 2 min at 3100 rpm with open lid, the eluate was discarded. 700 μ L of cleared lysate was then added to the column and centrifugation followed at 2000 rpm for 5 min with open lid. Clear lysate was loaded 2 times. Flow-through was kept for SDS-PAGE analysis. Then column was washed 2 times with 700 μ L of NPI20 buffer (Composition detailed in Table 2.14) and once with 700 μ L of NPI40 buffer (Composition detailed in Table 2.14) by centrifuging at 3100rpm for 2 min with open lid. The wash fractions were also kept for SDS-PAGE analysis. Finally, the proteins were eluted twice after transfer to clean Eppendorf with 200 μ L of NPI500 buffer (Composition detailed in Table 2.14) by centrifugation at 3100 rpm for 5 min. The concentration of protein in elution fractions was measured thanks to the Bradford method. All fractions were subsequently loaded onto a Mini-PROTEAN® TGX™ Precast protein gel (Bio-Rad).

Table 2.14: Composition of NPI20, NPI40 and NPI500 buffer. Sterilized by filtration on 0.22 μ m filters

Component	Origin	Quantity		
		NPI20	NPI40	NPI500
NaH ₂ PO ₄	Merck	690 mg	690 mg	69 0mg
NaCl	VWR Chemicals	1.75 g	1.75 g	1.75 g
Imidazole	Sigma-Aldrich	136 mg	272 mg	3.4 g
NaOH/HCl	VWR Chemicals	Adjust to pH 8 to 100mL		
Distilled water				

2.5 Cell wall extraction

First, 1 L of LB was inoculated with fresh colony and incubation followed at 37°C and 180 rpm until reaching a OD₆₀₀ of around 1. Then the culture was split in 250 mL fraction and cells were harvested at 4°C, 4000 rpm and 20 min. Cells were after resuspended in 40 mL of cold PBS (Oxoid) and disrupted by sonication by 5 cycles of 1 min with Bandelin Sonopuls HD220 Sonicator. Centrifugation at 1500 g during 10 min allowed to remove unbroken cells by taking the supernatant and discarding the pellet. Thereafter, cells were centrifuged at maximum speed during 5 min at 4°C and pellet was resuspended in 20 mL of SDS (Merck) and the solution was boiled during 20 min. Finally, three wash steps with 5 mL of warm deionized water, two wash steps with 5 mL of 1 M NaCl by centrifugation at room temperature, for 5 min, at maximum speed and 4 wash steps with 5 mL water sequentially centrifuged at 1500 g for 5 min

to remove the unbroken cell by keeping the supernatant and at maximum speed for 5 min and by resuspending the pellet. After wash steps, the pellet was eventually resuspended in 5 to 10 mL of MilliQ[®] water.

2.6 Endolysin characterization

2.6.1 Turbidity reduction assay

A decrease in light scattering (*i.e.* turbidity reduction) of a suspension of live cells or cell wall extract can be used to assess the activity of PG hydrolases. The reduction of optical density over time of the cells treated with the potential endolysin is compared to a buffer only control, both preparations have to be treated identically for the same period of time.

4-h culture of bacteria (mid-log) were first pelleted at 5000 rpm for 5 min at RT, cells were then resuspended in sterile buffer. Starting OD₆₀₀ has to be between 0.8 and 1.0. The assay was performed in a standard 96-well titration plate (CellStar[®] Greiner Bio-one). In each well, 20 μ L of purified protein or NPI500 buffer (if control well) was added to 180 μ L of bacterial suspension. A change in OD₆₀₀ was measured every min over the duration of the assay (30 min) on a Multiskan[™] FC Microplate spectrophotometer (Thermofisher scientific).

Percentage of inhibition is measured by the following formula :

$$\% \text{ inhibition} = \frac{OD_{600, t_0} - OD_{600, t_{30}}}{OD_{600, t_0}} \cdot 100 \quad (2.1)$$

Percentage of inhibition of the corresponding control treatment was subtracted from every data. All experiments were performed in triplicate.

Host range

Turbidity reduction assay was performed on different strains to assess the host range of the endolysin activity. All bacterial strains were resuspended in sterile Tris-HCl buffer (pH = 8) and tested with 100 μ L/mL of endolysin.

Biochemical characterization

All biochemical characterization was performed with *B. weihenstephanensis* LH002 strain with 100 μ L/mL of endolysin.

Turbidity reduction assay was performed with buffers ranging from pH 4 to 10 to assess the protein stability in different pH conditions.

Buffers with NaCl concentrations ranging from 50mM to 1000mM were used to evaluate the salt concentration effect on the endolysin.

To assess the temperature stability of the endolysin, proteins were incubated respectively at 4°C, 30°C, 37°C, 45°C, 60°C for 30 min and subsequently incubated 5 min on ice.

2.7 Cell wall decoration assay

The cell wall decoration assay was performed in order to investigate whether the recombinant proteins containing GFP interact with the cell surface of *B. weihenstephanensis* LH002 (or other if specified). *E. coli* was used as negative control. If proteins fused to GFP recognize the bacterial surface, the bacteria have to appear green under fluorescent microscope.

500 μ L of a 4-h exponential culture was centrifuged at 10000 g and 4°C for 2 min. Then, pellet was washed two times with 100 μ L of cold PBS. Thereafter, pellet was resuspended in 100 μ L of cold PBS and 10 $\sim$ μ g of the protein (or 10 μ L of NPI500 buffer if negative control) was added to the bacterial suspension and the samples were incubated at RT with agitation during

15 min. Then, samples were washed two times with 1 mL of cold PBS to remove excess of proteins. Finally the samples were resuspended in 100 μ L of cold PBS and 10 μ L is taken to be observed with the Leica Microsystems fluorescent microscope, using the LAS AFR[®] software.

2.8 Top agar

5 mL Top agar was contaminated with 100 μ L of O/N bacterial culture. Top agar were after spread on LB plates and dried for 20 min. Phage suspension was 10x diluted in sterile water to obtain dilutions ranging from 10^0 to 10^{-6} . 10 μ L droplets of phage dilutions were then spotted on the plates that were incubated O/N at 30°C or 37°C.

2.9 Bioinformatics

Basic local alignment search tool (BLAST) algorithm was used for amino acid sequence analysis and pairwise sequence. BLASTp, the BLAST program used in this work, searches protein databases using a protein query. Conserved domain searches were conducted using Interpro and HHpred was used to determine structural similarities.

Chapter 3

Results and discussion

3.1 Bioinformatic analyses of Deep-Purple genome

Deep-Purple has a genome of 36,278 bp [GenBank accession number - MF176161] featuring 40 putative CDSs, of which 24 have no predicted functions (Table S1). Numerous Deep-Purple CDSs feature BLAST hits with bacterial chromosomal proteins, suggesting that Deep-Purple could be a prophage.

As this work aimed to shed light on proteins involved in the adsorption and lysis processes, bioinformatic analyses focused on late genes to determine potential RBP and endolysin candidates.

3.1.1 Focus on tail proteins

In dsDNA phage genomes, structural proteins are often encoded in the late gene region in successive order of the virion assembly process (Habann et al., 2014). For siphoviruses, tail proteins are comprised between the TMP and the lysis cassette, gathering the holin and endolysin. As described in Section 1.2.2, the tail morphogenesis of *Siphoviridae* has generally the same organization and features, in order, a TMP that dictates the tail length and facilitates DNA transit to the host cytoplasm ; a Dit protein that forms a central hub for the anchorage of other tail components ; a Tal protein with a hydrolase activity for PG degradation during DNA injection and finally a RBP for the recognition of bacterial receptor (Dunne et al., 2018b). The tail region may also feature proteins with unknown functions. This general organization can be observed in the *L. lactis* phage p2 (Fig. 3.1).

Comparison of Deep-Purple putative tail proteins with other tail proteins was performed using both BLAST (Basic Local Alignment Search Tool, finds regions of local similarity between sequences) and HHpred (protein homology detection and structure prediction (Söding et al., 2005)).

Deep-Purple gp27 is a 1245 amino acid (aa)-protein that shows homology with some phage TMPs. Interestingly, some gp27 top BLAST hits were associated with carbamoyl-phosphate synthase large subunit (Coverage of 100% and percentage of identity of 70%). HHpred analysis indicates a structural homology between the N-terminal part of gp27 and an inner ring of nuclear pore complex (from aa 15 to 226).

Deep-Purple gp28 (495 aa) features aa homology with five Dit proteins, two BLAST hits cover the entire gp28 protein with a percentage of identity ca. 40% whereas the other hits only match to 24% of gp28 sequence (from aa 1 to 138) with around 30% of identity. BLAST analyses also showed 21 hits with putative tail fibers. Additionally gp28 features a conserved protease domain (PA_VSR, accession number : cd02125). BLAST alignments revealed a more conserved N-terminal part of the protein (Fig. S1). Interestingly, HHpred revealed homology

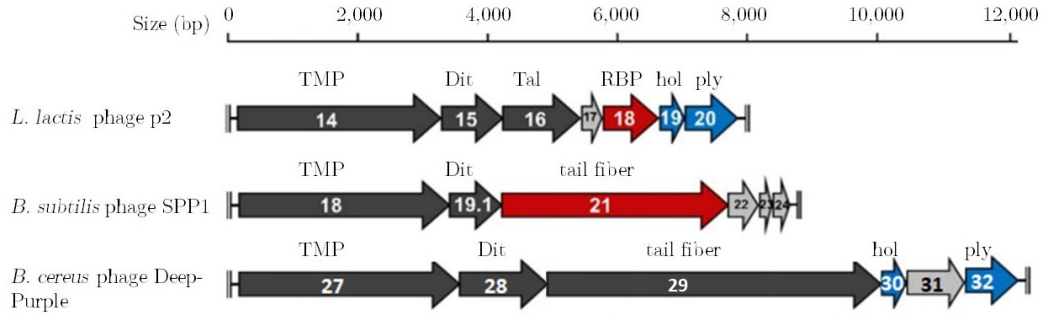


Figure 3.1: *L. lactis* phage p2, *B. subtilis* phage SPP1, *B. cereus* phage Deep-Purple tail morphology. Light grey arrows represent hypothetical proteins whereas dark grey arrows indicate proteins of known or predicted functions. Abbreviations are : TMP - Tape measure protein ; Dit - Distail tail protein ; Tal - tail associated lysin ; hol - holin ; ply - endolysin. Adapted from (Dunne et al., 2018b).

of the central region of gp28 (from aa 182 to 435) with *L. casei* phage J-1 Dit protein that is defined as "evolved Dit". Dieterle et al. (2017) analyzed J-1 evolved Dit and revealed that it features two extra modules when compared to classical Dit. Both domains were associated with carbohydrate binding domain and tested for their binding capacity to the cell wall. Only the second domain showed ability to bind to host cell. However, it cannot be excluded that the first module could be functional for other phages. In J-1 phage, the evolved Dit was identified as unique RBP. This suggests that evolved Dit should be investigated as active proteins toward the host instead of passive hubs. They can act as RBP alone or be associated with recognition modules during adsorption (Dieterle et al., 2017).

Deep-Purple gp29 (2156 aa) showed some BLAST hits with tail fibers and some minor structural proteins. BLAST alignments revealed a less conserved N-terminal part (Fig. S2). In addition, gp29 features a conserved C-terminal intramolecular chaperone domain (IMC) (from aa 2016 to 2156, Peptidase_S74, accession number : pfam13884). This catalytic dyad allows the C-terminal portion of the protein to be cleaved and folded properly. In *Salmonella* phage S16, the last 112 residues of gp37 (part of the LTF with gp38) form an IMC assisting with folding of gp37 before autoproteolysis (Dunne et al., 2018a). HHpred analysis indicated structure similarity between those two IMCs. Additionally, gp29 features three other conserved domains :

- The prophage_tail domain (aa 75 to 352), probably associated with endopeptidase functions
- The DUF4795 (aa 480 to 596) of unknown function
- The gp58 domain (aa 1402 to 1469) associated with minor structural protein.

No Deep-Purple gene product showed similarity with any Tal protein. This either suggests that Deep-Purple has the ability to work without any Tal protein, as in the case of *B. subtilis* phage SPP1 (Godinho et al., 2018) (Fig. 3.1) or that gp29 featuring a putative endopeptidase domain could act as Tal function.

Deep-Purple general tail organization is summarized in Fig. 3.1. The comparison of the *Bacillus* phage vB_BtS_B83, Deep-Purple and *Bacillus* phage PfNC7401 tail protein regions can be seen in Fig. 3.2. It shows that the three phages display a similar organization.

In regards with these bioinformatic analyses, gp28 and gp29 were chosen as RBP candidates. While gp26, which is a hypothetical protein outside of the region coding for tail apparatus (*i.e.* before the TMP) was chosen as negative control as it is expected that this protein is not able to adsorb to the bacterial host.

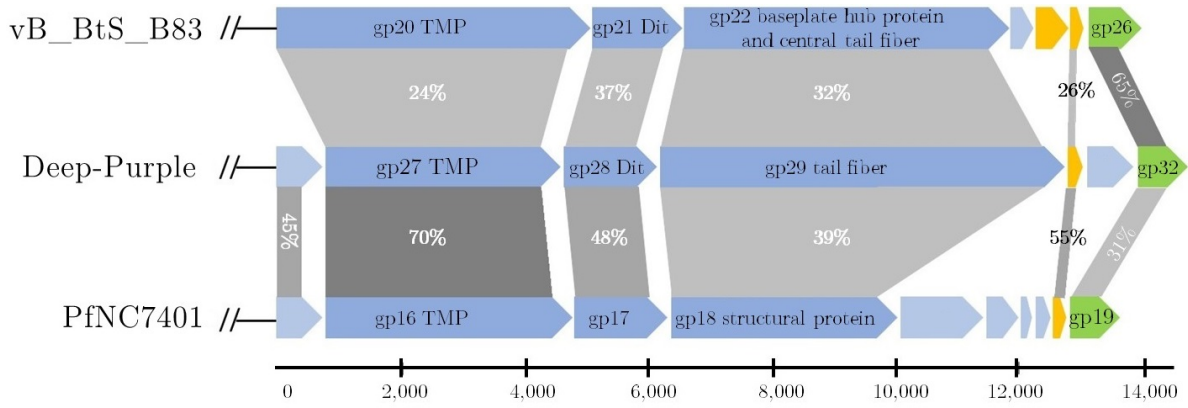


Figure 3.2: Genome alignment of phages vB_BtS_B83, Deep-Purple and PfNC7401. Hypothetical proteins in light blue, proteins with predicted functions in dark blue, holin in yellow and endolysin in green. The degree of amino acid sequence identity is indicated in the grey portions. Abbreviations are : gp - gene product ; TMP - Tape measure protein ; Dit - distal tail protein

3.1.2 Lysis cassette

Deep-Purple gp30 (70 aa) shows three BLAST hits with holins. It features a single predicted transmembrane domain identified with TMHMM server (aa 12 to 34). Furthermore, gp30 features two start codons (aa 1 and 12) in agreement with the dual-start model that states that most holins code for holin and anti-holin.

gp32 (262 aa) is the putative endolysin because of numerous BLAST hits with endolysins including endolysin of *B. cereus* PBC5 phage (Lee et al., 2019) and Bcp1 phage (Kong and Ryu, 2015). As expected, BLAST alignments showed a more conserved C-terminal region (Fig. S3). Interpro analyses allowed to define two domains : a N-terminal glycosyl hydrolase family 25 (GH25) catalytic domain (aa 1 to 180, IPR002053) and a C-terminal SH3 domain (aa 196 to 262, G3DSA:2.30.30.40) for host recognition (Fig. 3.3). The GH25 family is associated with 1,4-N-acetylmuramidase activity that hydrolyses the (1-4)- β -linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues in the peptidoglycan layer. SH3 domains putatively recognize conserved ligands within a bacterial genus. Additionally, HHpred analysis showed structural similarities for the entire gp32 sequence with *Clostridium perfringens* phage phiSM101 endolysin (22% of identity).

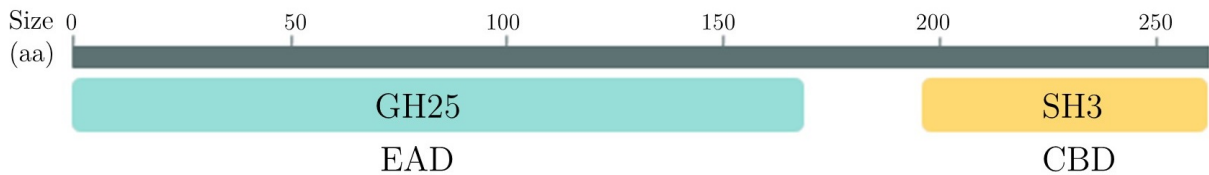


Figure 3.3: Endolysin schematic representation. Putative N-terminal EAD domain ranging from aa 1 to 180 and putative C-terminal CBD domain ranging from aa 196 to 262.

3.2 Deep-Purple lifestyle

Deep-Purple was isolated through enrichment of an agricultural soil sample with bacteria. This consisted in homogenization of solid samples in broth, buffer or water followed by addition of liquid growth medium and a suspension of target bacteria for which we want to isolate phages. After incubation, the mix was filtered to remove the bacteria and the presence of phage was

revealed by top agar (Ackermann and Węgrzyn, 2014). The main disadvantage of this technique is that prophages originally present in the bacteria used for the enrichment step can be isolated.

Moreover, BLAST analyses of Deep-Purple proteins revealed many hits with bacterial chromosomal proteins (Table S1). This led to the hypothesis that Deep-Purple could be a lytic variant of a plasmidial prophage present in the strain that was originally used to enrich the sample. In order to verify this hypothesis, a PCR screening was performed to detect the presence of Deep-Purple in a representative collection of strains of the *B. cereus* group (listed in Table 2.1).

The 20 emetic strains used for the enrichment step as well as 20 random non-emetic *B. cereus* strains were screened by PCR with four pairs of primers specific to Deep-Purple genome. None of the tested strain displayed Deep-Purple genome amplicon in their DNA. A representative result of PCR screening is shown in Fig. 3.4

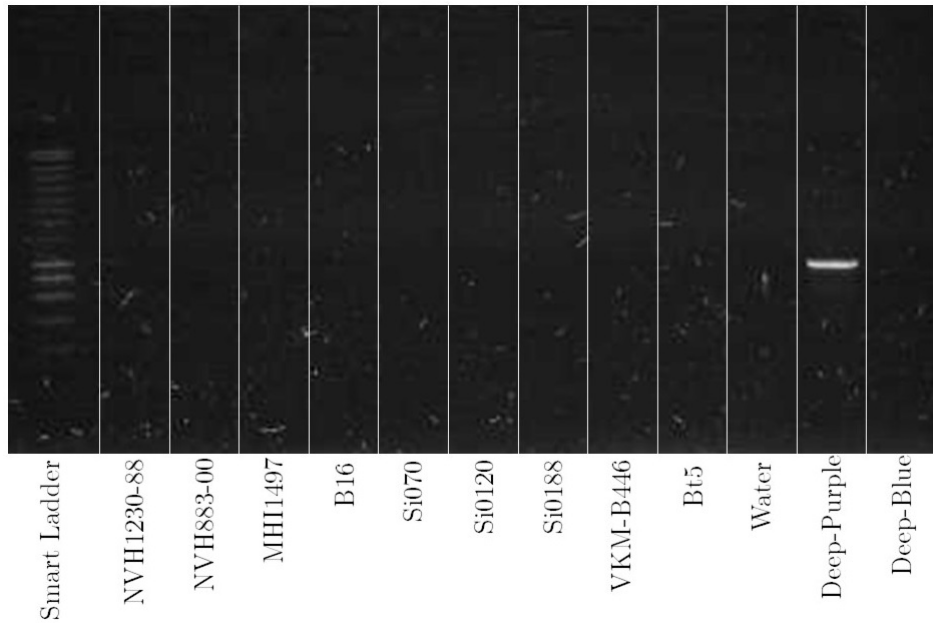


Figure 3.4: Representative result for PCR screening for Deep-Purple presence in bacterial strains where no amplification of Deep-Purple DNA in the tested strain was observed. Primers used : Purple_screen_2_F and Purple_screen_2_R (Table 2.11). Positive control with Deep-Purple DNA and negative controls with both water and myophage Deep-Blue DNA.

Although, it cannot be concluded that Deep-Purple is not a prophage. Its prevalence, if any, is very low among the *B. cereus* strains. It is also possible that Deep-Purple is not a prophage and that it is only a virulent phage with an abnormal behaviour.

3.3 RBP identification

As discussed in Section 3.1.1, RBP candidates were gp28 and gp29, firstly because of their position in the genome between the putative TMP protein and the lysis cassette. Additionally, gp28 showed a putative evolved Dit function and therefore potentially features two extra modules compared to classical Dit of which one could be involved in bacterial recognition. Gp29 showed numerous BLAST hits with tail fibers and contains a putative chaperone domain as some phages need chaperone activity for the good folding of tail components (Dunne et al., 2018a). The different constructions made for RBP investigation are summarized in Fig. 3.5. All genes of interest were amplified by PCR with primers listed in Table 2.7. All clonings steps were realized

by restriction/ligation of the gene of interest into pET30 vector followed by GFP insertion in the resulting vector.

To assess the binding capacity of RBP candidates, they were expressed as recombinant proteins fused with a GFP. Since in most RBPs, the C-terminal part comprises the binding site for the bacterial receptors, the GFP was cloned on the N-terminal part, except for gp28 where the GFP cloning was done on the C-terminal part (pET30::gp28::(link)GFP). In parallel, another gp28 cloning was performed with only the C-terminal part of the protein (pET30::GFP(link)::gp28_mid) that is believed to be responsible for the binding. To investigate gp29 function, two clonings of the C-terminal part of the protein were performed (pET30::GFP(link)::gp29_1200, starting at aa 1200 and pET30::GFP(link)::gp29_1400, starting at aa 1400).

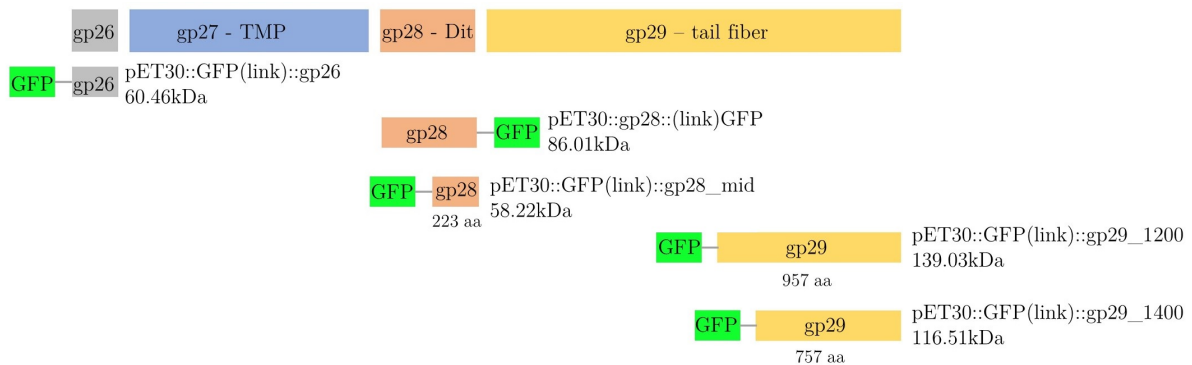


Figure 3.5: Graphic summary of the cloning constructions made for RBP identification. Abbreviations are : GFP - Green Fluorescent Protein ; gp - gene product ; TMP - Tape measure protein ; Dit - distal tail protein. Final protein molecular weights are indicated.

The recombinant plasmids were subsequently transformed in competent *E. coli* 10 β and transformants were screened by PCR. Sequencing confirmed that all constructions were correct. Finally, plasmids were transformed into the expression strain *E. coli* BL21(DE3).

The expression of six-histidine-tagged recombinant proteins was induced by the addition of IPTG to cultures of *E. coli* BL21(DE3)[pET30::protein_of_interest]. For each protein, the concentration of IPTG, the duration and the temperature of incubation of the induced culture were optimized in order to obtain soluble proteins.

All recombinant proteins except GFP(link)::gp28_mid were found in the soluble fraction when induction was performed at 28°C for 5 h with 1 mM of IPTG. Induction of GFP(link)::gp28_mid at 22°C for 5 h with 1 mM still did not yield soluble proteins (data not shown).

Purification of soluble recombinant proteins GFP(link)::gp26, gp28::(link)GFP, GFP::gp29_1200 and GFP::gp29_1200 was performed in native conditions by affinity chromatography on Ni-NTA columns. After several washing steps, proteins were eluted by addition of imidazole. The obtained concentration for gp29 protein was smaller than for gp26 and gp28. Representative results for gp28 and gp26 solubility assessment and purification are shown in Fig. 3.6.

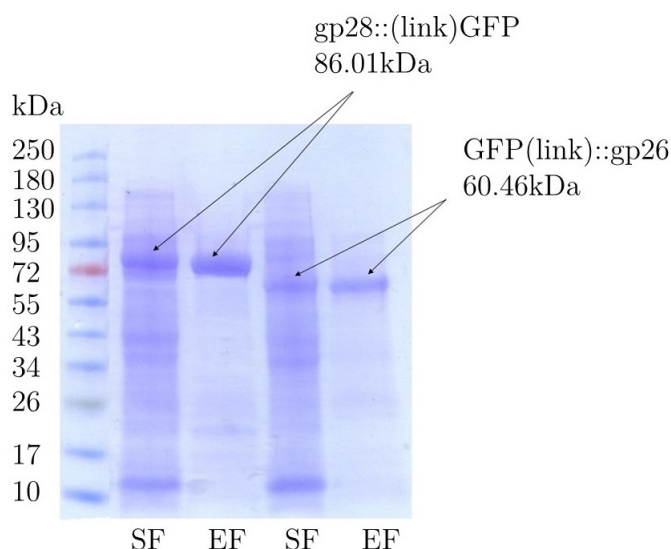


Figure 3.6: SDS-PAGE gel for gp28 and gp26. Abbreviations are : SF - soluble fraction ; EF - elution fraction.

The binding capacity of the RBP candidates was then assessed using the cell wall decoration assay (Fig. 3.7).

As shown in Fig. 3.7B, *B. cereus* ATCC10987 cell wall decoration assay for gp28 displayed obvious fluorescence indicating its capacity to bind to *B. cereus*. Also, gp28 successfully bound to *B. weihenstephanensis* LH002 (data not shown). These results strongly suggested that gp28 is a RBP.

Very faint fluorescence was also seen when *B. cereus* ATCC10987 was subjected to gp 29_1200 (Fig. 3.7, C) or gp29_1400 (data not shown). The same results were obtained when *B. weihenstephanensis* LH002 was subjected to gp29 proteins (data not shown).

Faint fluorescence could also be observed when *B. cereus* ATCC10987 was subjected to gp26 protein (Fig. 3.7, A). However, it is unlikely that gp26 really adsorb to the host and fluorescence could be the result of unbound protein in the environment.

E. coli 10 β cell wall decoration assay for gp26, gp28 and gp 29_1200 (Fig. 3.7, D-F) and gp29_1400 (data not shown) served as negative control. No fluorescence was observed for all the tested proteins, as Deep-Purple RBP is not expected to adsorb to *E. coli*.

To conclude, the strong fluorescence observed for cells treated with gp28 strongly suggests that gp28 is a RBP. The very faint fluorescence observed for cells treated with gp29 indicates that it could possibly also have a role to play in the adsorption process. Indeed, some phages including *E. coli* phage phi92, feature two or more RBPs that have the ability to recognize different bacterial receptors, therefore contributing to their wide host range (Schwarzer et al., 2012), which seems not to be the case for Deep-Purple as it does not feature a broad spectrum. Also, *E. coli* phage T4 and other phages, possess two RBPs and each is involved in one step of the recognition process (*i.e.* reversible and irreversible binding) (Leiman et al., 2003). Finally, some phages including *L. lactis* phage Tuc2009 feature two different carbohydrate binding modules : one is the *bona fide* RBP and one in an accessory protein BppA that enhances adsorption to cells (Legrand et al., 2016).

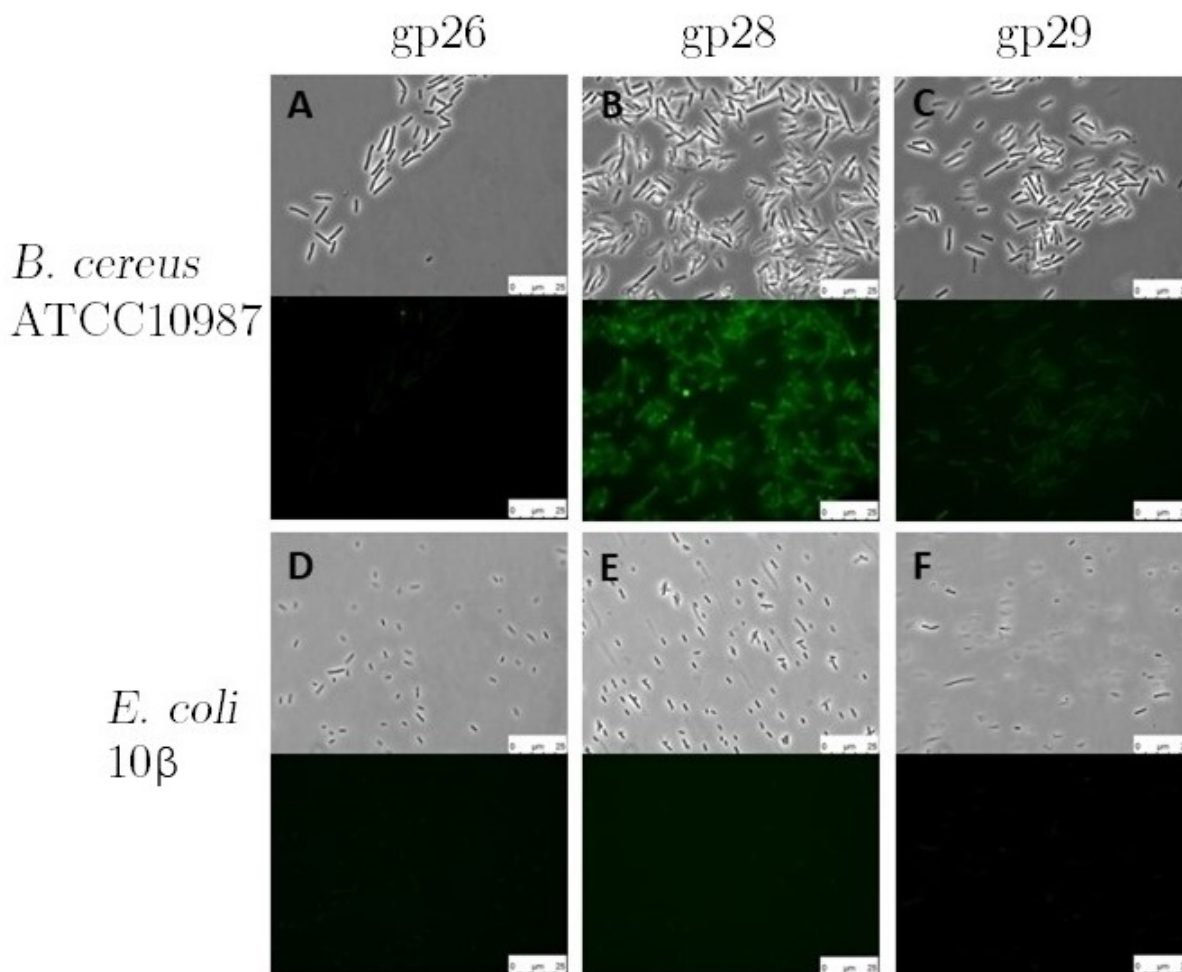


Figure 3.7: *B. cereus* ATCC10987 (A-C) and *E. coli* 10 β (D-F) cell wall decoration assay for GFP(link)::gp26 (A, D), gp28::(link)GFP (B, E) and GFP(link)::gp29_1200 (C, F). Line 1 and 3 : Bright field microscopy. Line 2 and 4 : fluorescence micrograph. Also experiments were performed with ca. 10 μ g of proteins.

3.4 Endolysin

3.4.1 Molecular cloning and protein induction

As discussed in Section 3.1.2, gp32 was selected as Deep-Purple putative endolysin. Cloning steps performed for the endolysin identification and characterization are summarized in Fig. 3.8. The same strategy as for RBP identification was used.

To assess gp32 lytic activity, cloning of the full-length gene was realized (pET30::gp32). To evaluate the binding capacity of gp32, three cloning experiments were performed, one with the CBD domain (pET30::GFP(link)::gp32_CBD(196)), one featuring both CBD and linker region (pET30::GFP(link)::gp32_CBD(180)) and one with the full-length protein, all fused to GFP.

All recombinant proteins listed in Fig. 3.8 were found in the soluble fraction when induction was performed at 28°C for 5 h with 1 mM of IPTG. A representative result for gp32 solubility assessment is shown in Fig. 3.9. Purification of soluble recombinant proteins was performed in native conditions by affinity chromatography on Ni-NTA columns. All purifications gave protein concentrations of around 1 to 2 mg/mL. Protein purity was verified by SDS-PAGE (Fig. 3.10).

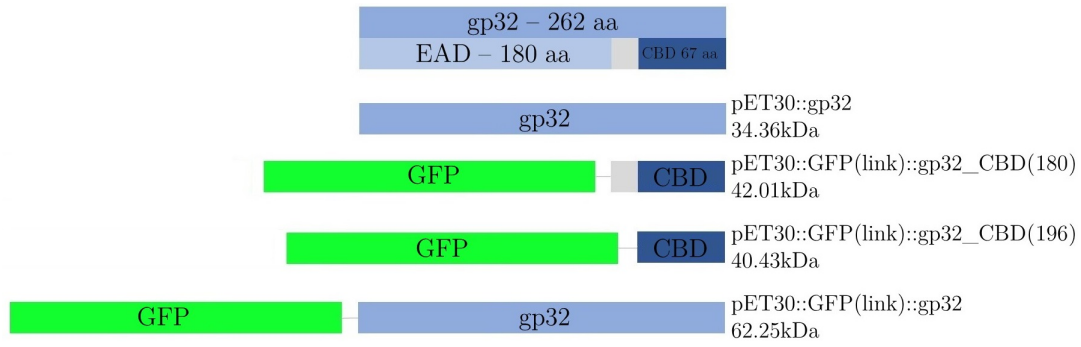


Figure 3.8: Summary of the cloning steps performed for endolysin identification and characterization. Abbreviations are : GFP - Green Fluorescent Protein ; gp - gene product ; EAD - Enzymatically Active Domain ; CBD - Cell wall Binding Domain. Final protein weights are indicated.

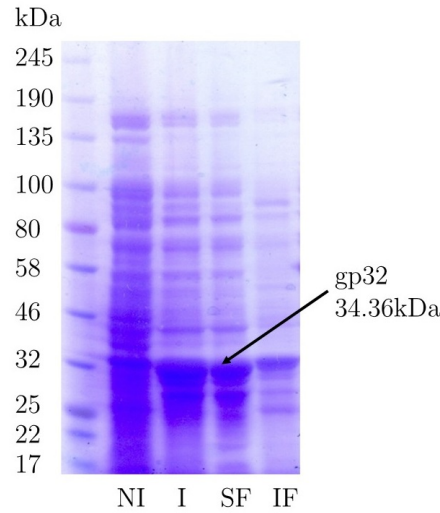


Figure 3.9: Solubility assessment of gp32 protein using SDS-PAGE. Abbreviations are : NI - Non-Induced control ; I - Induced Control ; SF - Soluble Fraction ; IF - Insoluble Fraction.

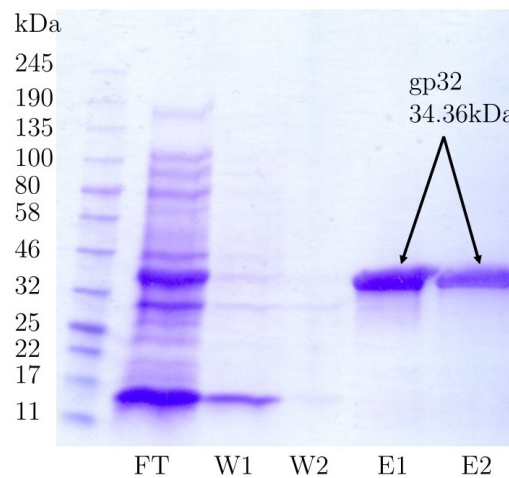


Figure 3.10: SDS-PAGE gp32 purification steps. Abbreviations are : FT - flow-through ; W1 - wash 1 ; W2 - wash 2 ; E1 - elution 1 ; E2 - elution 2.

3.4.2 Gp32 lytic activity

Gp32 lytic activity was assessed thanks to the turbidity reduction assay that relies on the decrease in light scattering (measured by the variation of optical density) of a suspension of live cells or cell wall extract when subjected to a PG hydrolase.

To assess the effect of gp32 protein concentration on cells, *B. weihenstephanensis* LH002 were treated with different concentrations of gp32. Optical density (OD) was measured every min during 40 min. Results are shown in Fig. 3.11. This graph shows a dose-response phenomenon where higher protein fractions led to a higher percentage of inhibition. Around 42% of inhibition was reached after 15 min when treated with 2 mg/mL of endolysin whereas 22 min were needed to obtain the same percentage of inhibition when subjected to 1 mg/mL of endolysin. In comparison, *B. cereus* cells subjected to 2.5 mg/mL of phage PBC5 endolysin led to 50% of inhibition after 15 min (Lee et al., 2019).

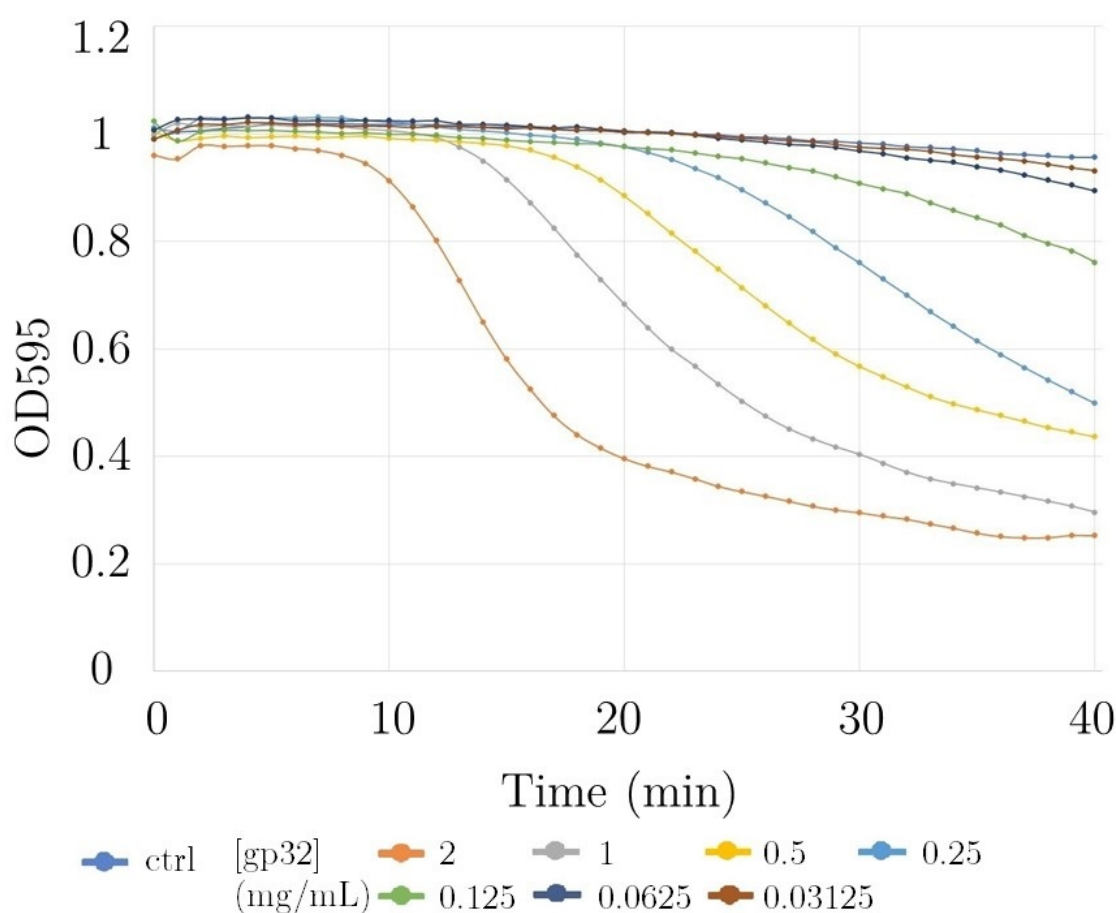


Figure 3.11: *B. weihenstephanensis* LH002 resuspended in TrisHCl buffer (pH 8) subjected to decreasing amounts of gp32, ranging from 2 to 0.03125 mg/mL. OD was measured every min for 40 min at 30°C.

B. cereus ATCC10987 cell wall was extracted and subjected to gp32 to assess its activity on PG. As shown in Fig. 3.12, a clear dose-response was also observed where only 10 min was needed to obtain ca. 43% of inhibition for the treatment with 1 mg/mL of gp32. Twenty-six min were needed to obtain the same percentage of inhibition with 0.5 mg/mL of gp32.

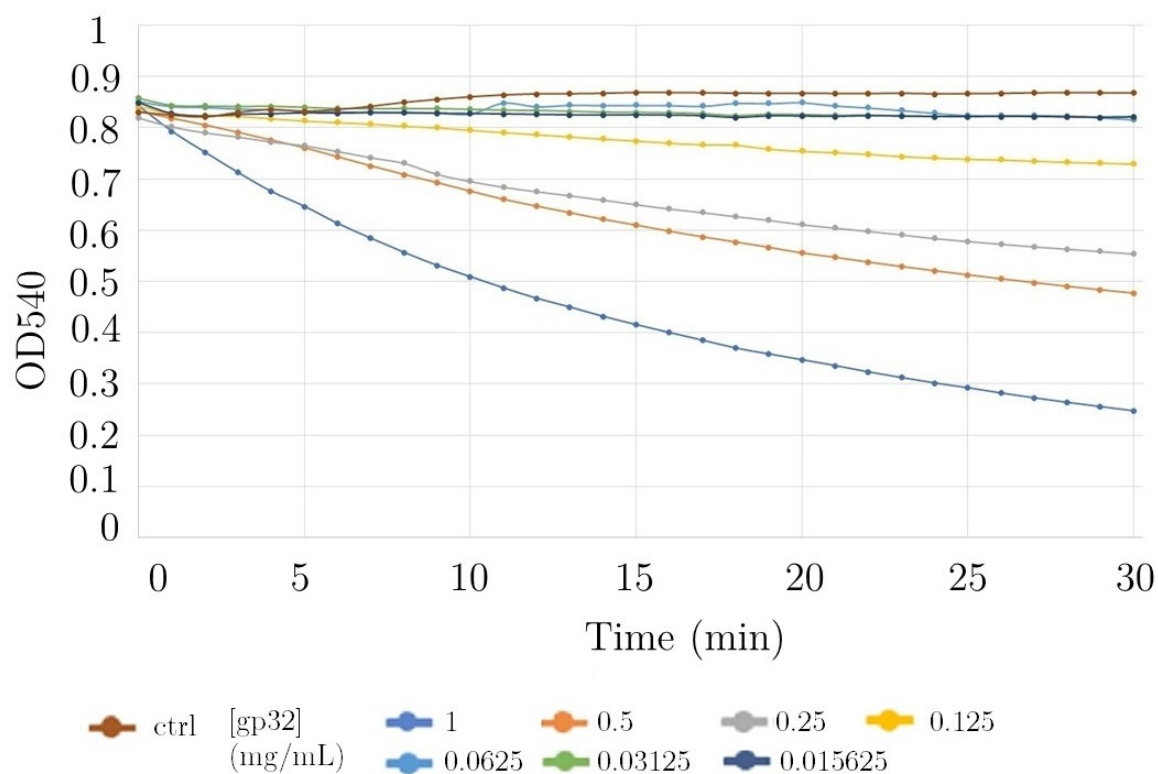


Figure 3.12: *B. cereus* ATCC10987 cell wall extract subjected to decreasing amounts of gp32 ranging from 1 to 0.015625 mg/mL. OD was measured every min for 30 min at 30°C.

Activity spectrum

The activity spectrum of endolysin was assessed via turbidity reduction assay on a representative collection of strains (Fig. 3.13, summarized in Table 3.1). The endolysin shows lytic activity against all the tested members of the *B. cereus* group and some strains out of the *B. cereus* group were also sensitive to gp32.

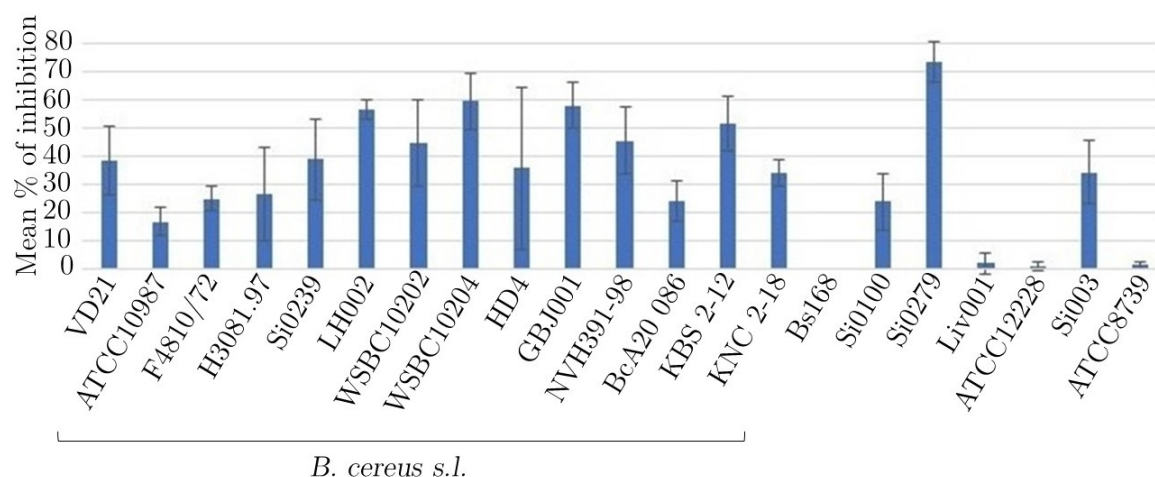


Figure 3.13: gp32 activity spectrum assessed by turbidity reduction assay. Mean percentage of inhibition and standard errors are indicated. This relies on the OD decrease of a bacterial suspension submitted to gp32. The protein was tested at 100µg/mL for 30 min. Experiments were performed in triplicate. These results are also reported in summary Table 3.1 in section 3.4.4.

Biochemical characterization

As seen in previous studies, the endolysin activity can be influenced by salt concentration, pH and temperature (Nelson et al., 2012) ; (Etobayeva et al., 2018). Therefore, the lytic activity of gp32 was assessed over a broad range of pH (4 to 10), NaCl concentrations (0 to 1000 mM) and exposure to different temperatures (4 to 60°C). Results can be seen in Fig. 3.14. Firstly, gp32 seems to be active at a broad range of pH from 4 to 8, with maximum activity at pH 6. A similar pattern was observed for the endolysin of *B. cereus* phage PlyTB40 (Etobayeva et al., 2018). Gp32 seems to be highly sensitive to NaCl concentration, with a drop to less than 15% for all treatment with NaCl, whereas 69% of inhibition is observed for the control treatment without NaCl. A similar NaCl sensitivity was observed for the endolysin of *S. aureus* phage LysCSA13 (Cha et al., 2019). Finally, gp32 was stable at 30 and 37°C and still displayed activity at 45°C. The same temperature stability pattern has been observed for PlyB endolysin of *B. cereus* phage vB_BanS_Bcp1 (Schuch et al., 2019).

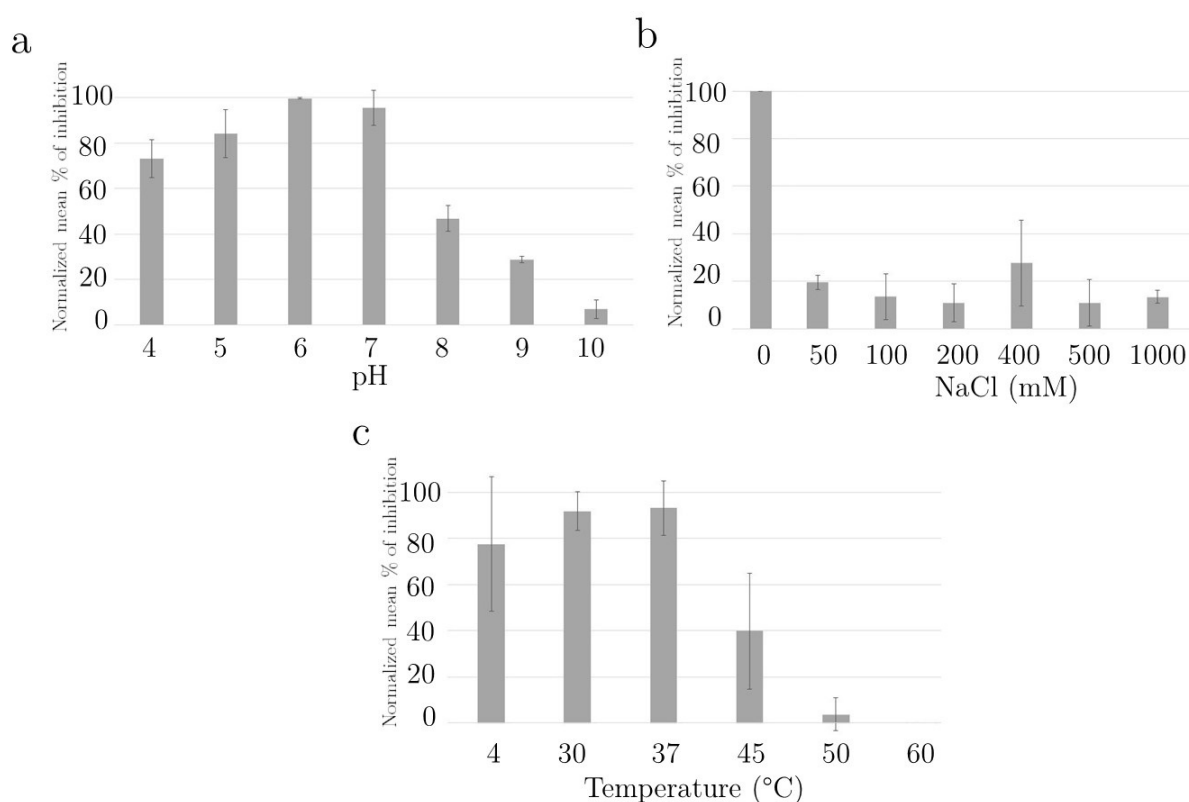


Figure 3.14: Biochemical characterization of gp32 activity. The effects of pH (a), NaCl dependence (b), and temperature (c) was evaluated for gp32. The protein was tested at 100 μ g/mL against exponential phase *B. weihenstephanensis* LH002 for 30 min. Values are represented as normalized mean percentage of inhibition and standard errors are indicated. This relies on the OD decrease of a bacterial suspension submitted to gp32. The experiments were performed in triplicate.

3.4.3 Cell wall decoration assay

Binding capacity of gp32 full-length and CBD

As seen in Fig. 3.3, gp32 features an SH3 domain that is thought to function as CBD. The binding capacity of the recombinant proteins was assessed using a cell wall decoration assay. As shown in 3.15, all three constructions tightly bound to *B. weihenstephanensis* LH002. *E. coli* cells were also subjected to the recombinant proteins and showed no fluorescence as endolysin is not expected to bind to Gram-negative bacteria due to the presence of the outer membrane surrounding the PG (data not shown). Additional control experiment with only GFP was also performed (data not shown).

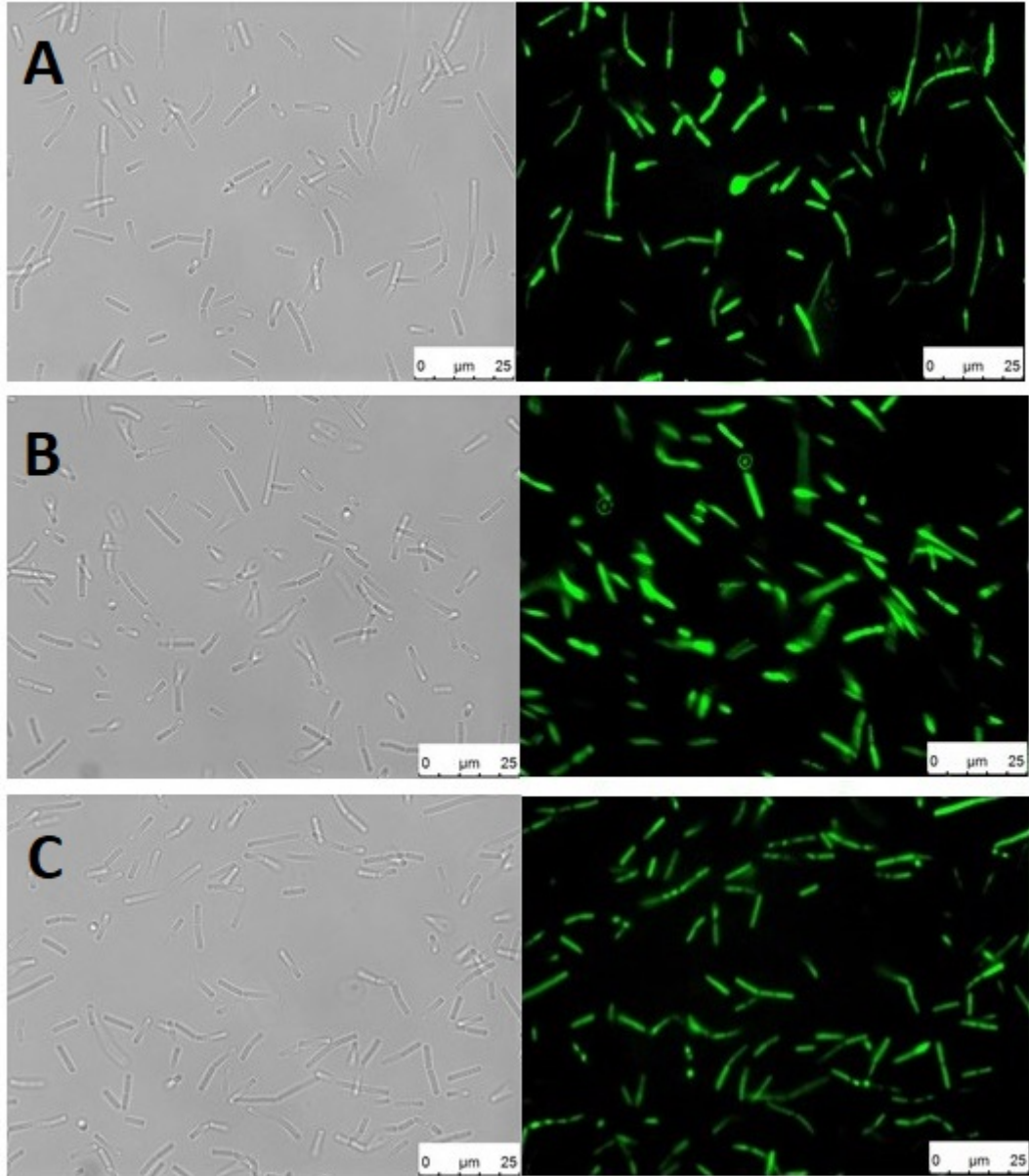


Figure 3.15: Cell wall decoration on *B. weihenstephanensis* LH002 strain with GFP(link)::gp32 (A) ; GFP(link)::gp32_CBD(180) (B) and GFP(link)::gp32_CBD(196) (C). Left : Bright field microscopy. Right : fluorescence micrograph. Cells were treated with around 10 μ g of recombinant protein

CBD activity spectrum

All strains tested for the assessment of endolysin activity were also subjected to cell wall decoration assay with ca. 10 μ g of GFP(link)::gp32_CBD(196) in order to evaluate the binding capacity spectrum. Representative results are shown in Fig. 3.16. All results are also reported in the summary Table 3.1. Those results correlate with the endolysin activity assessed in Section 3.4.2.

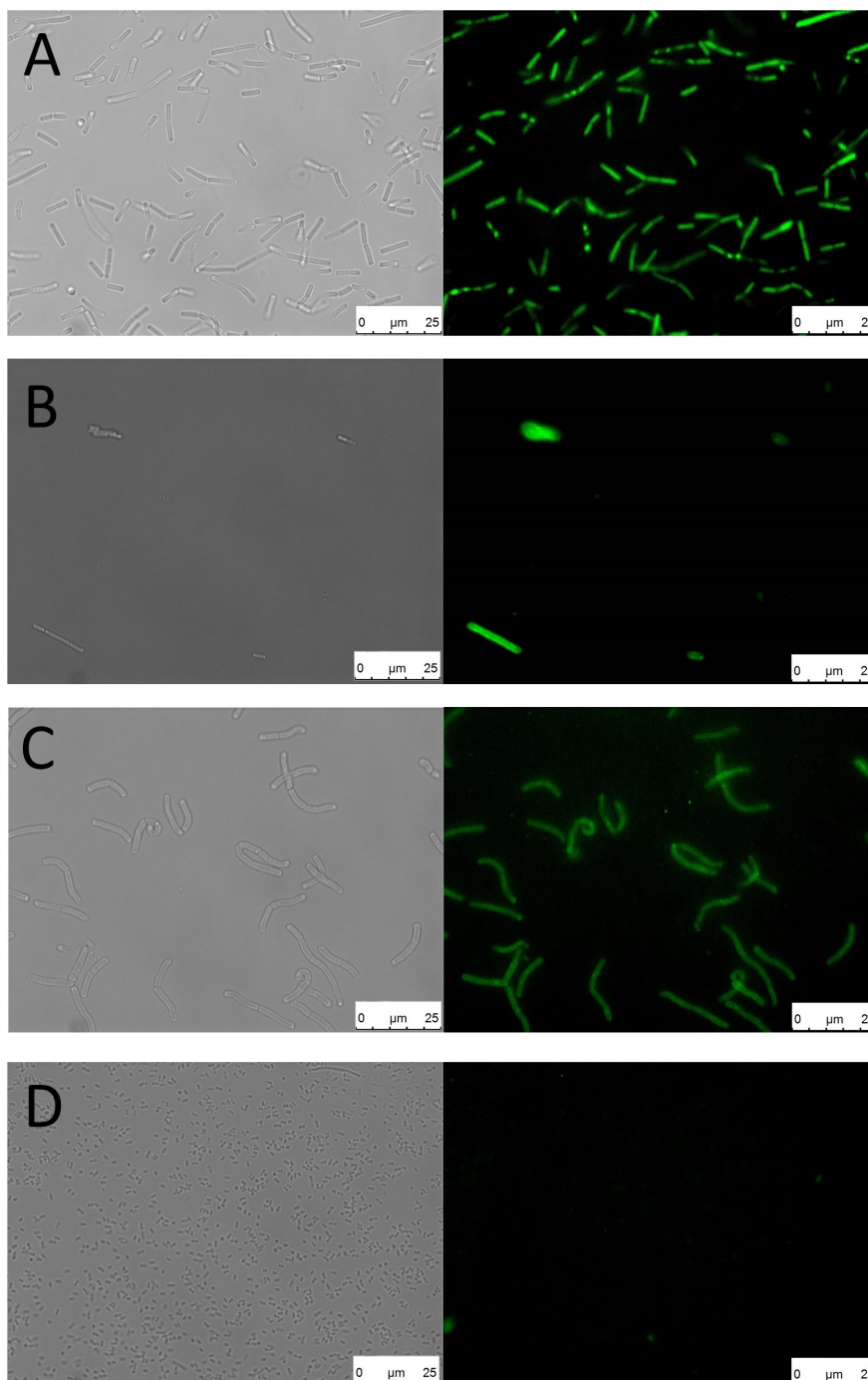


Figure 3.16: Cell wall decoration on (A) *B. weihenstephanensis* LH002 strain, (B) *B. thuringiensis* GBJ001, (C) *B. megaterium* Si003 and (D) *E. coli* ATCC8739 with 10 μ g of GFP(link)::gp32_CBD(196). Left : Bright field microscopy. Right : fluorescence micrograph.

3.4.4 Gp32 activity summary

Endolysin activity and CBD binding capacity assessed on a representative collection of strains were compared to Deep-Purple host spectrum. Deep-Purple host range was evaluated thanks to top agar experiments. These results are compiled in Table 3.1.

Globally, three types of results were encountered : *sensitive bacteria*, when phages were able to enter the bacteria and accomplish a replication cycle leading to individual plaques ; *insensitive bacteria*, when no lysis were observed and *bacterial lysis*, when phages only induced lysis at the highest concentration where numerous phages are trying to inject their DNA, therefore making little holes and killing the bacteria. This process is called "lysis from without" (Abedon, 2011).

It shows that all members of the *B. cereus* group are sensitive to the endolysin whereas only a few are sensitive to the phage itself. The broader range of activity for the endolysin is explained by the fact that endolysin target conserved bound in the PG.

Table 3.1: Summary Table for gp32 activity and CBD binding compared to the phage spectrum. Phage spectrum assessed with top agar : S : sensitive (phages capable to enter the bacteria, resulting in countable lysis plaques) ; I : insensitive (no lysis plaques at all) ; L : lysis (lysis from without). Endolysin activity : - : <10% of inhibition ; + : 10-40% inhibition ; ++ : 40 - 70% inhibition ; +++ > 70% inhibition. Strains in bold type belong to the *B. cereus* group

Strains	Bacterial species	Phage spectrum	Endolysin activity	Binding of CBD
VD21	<i>B. cereus</i>	S	+	+
ATCC10987	<i>B. cereus</i>	I	+	+
F4810/72	<i>B. cereus</i>	L	+	+
H3081.97	<i>B. cereus</i>	I	+	+
Si0239	<i>B. weihenstephanensis</i>	L	+	+
LH002	<i>B. weihenstephanensis</i>	S	++	+
WSBC10202	<i>B. weihenstephanensis</i>	L	++	+
WSBC10204	<i>B. weihenstephanensis</i>	L	++	+
HD4	<i>B. thuringiensis</i>	S	+	+
GBJ001	<i>B. thuringiensis</i>	L	++	+
NVH391-98	<i>B. cytotoxicus</i>	L	++	+
BcA20	<i>B. cytotoxicus</i>	L	+	+
KBS 2-12	<i>B. mycoides</i>	S	++	+
KNC 2-18	<i>B. mycoides</i>	I	+	+
Bs168	<i>B. subtilis</i>	I	-	-
Si0100	<i>B. amyloliquefaciens</i>	I	+	+
Si0279	<i>B. pumilus</i>	I	+++	+
Liv001	<i>L. ivanovii</i>	I	-	-
ATCC12228	<i>S. epidermis</i>	I	-	-
Si003	<i>B. megaterium</i>	I	+	+
ATCC8739	<i>E. coli</i>	I	-	-

Chapter 4

Conclusions and perspectives

As natural bacterial parasites, phages and their proteins offer numerous properties for the detection and the biocontrol of pathogens. Understanding how they enter (*i.e.* adsorption) and how they exit (*i.e.* lysis) bacteria is crucial for the development of phage technologies. The siphophage Deep-Purple was the object of this master thesis and interest for this phage resides on its lytic activity against members of the *B. cereus* group. This work aimed to shed light on Deep-Purple RBP(s) involved in host recognition and endolysin that degrades the bacterial peptidoglycan at the end of the infection cycle.

Beside the main objectives of this master thesis, it was tested whether Deep-Purple could be a lytic variant of a plasmidial prophage that was originally present in strains used to enrich the sample during its isolation. To do so, 40 emetic and non-emetic *B. cereus* strains were screened for Deep-Purple DNA presence with four pairs of primers specific to Deep-Purple genome. None of the tested strains displayed Deep-Purple genome amplicon in their DNA. Although one cannot totally exclude the possibility that Deep-Purple is a plasmidial prophage, its prevalence among *B. cereus* strains would be very low.

One of the main emphases of this work was the adsorption process, which results from the interaction between a phage RBP and a receptor on the bacterial surface. This highly specific interaction influences the host range.

Bioinformatic analyses of Deep-Purple region coding for tail apparatus (*i.e.* between TMP and lysis cassette for *Siphoviridae*) retained two potential RBP candidates. Gp28 is associated with a so-called "evolved Dit" function, putatively featuring two additional modules compared to classical Dit, of which one was reported to bind to bacteria and function as RBP in some phages including the *L. casei* phage J-1 (Dieterle et al., 2017). Gp29 was also considered as RBP candidate because BLAST analysis showed relevant matches with putative tail fibers and because it features a C-terminal IMC domain with a structure similar to the IMC of the RBP (*i.e.* gp37) of *Salmonella* phage S16 (part of LTF) that assists with folding of gp37 after autoproteolysis (Dunne et al., 2018a).

Gp28 and the N-terminal end of gp29 were tagged with a GFP (on the C-terminal end for gp28 and on the N-terminal end for gp29). In parallel, cloning of the C-terminal part of the gp28 protein (called gp28_mid) N-terminally fused to GFP was performed. Tagged RBP were then transformed into the expression strain BL21(DE3). Inductions with IPTG yielded soluble proteins except for GFP::gp28_mid. Lowering the temperature of induction still did not yield soluble GFP::gp28_mid proteins. Paths that could have been explored to remedy this solubility problem include changes in inducer concentration and duration of induction. Moreover, it would have been interesting to co-express molecular chaperones or to use a different expression vector and/or another host. As last resort, insoluble proteins could be solubilized with denaturing agents and then refolded after extraction. Finally, all proteins of interest were

successfully purified after cell lysis under native conditions by affinity chromatography using Ni-NTA columns.

To assess the ability of RBP candidates to interact with bacterial hosts, cell wall binding assays were performed. To do so, purified recombinant proteins were put into contact with bacterial cells. Fluorescence observed for gp28 on *B. cereus* indicates that gp28 is able to interact with bacterial host and act as probable RBP. Additionally, very faint fluorescence was also observed for gp29, suggesting a weak binding to *B. cereus*. However, as the tested amount of gp29 protein was rather low (data not shown), this decoration assays should be repeated with more appropriate amounts of protein.

Taking into account the problems discussed above, we have to be careful when concluding and at this point, we can only say that both gp28 and gp29 seem to interact with the bacterial host. After repetition of the cell wall decoration assay, the functions of both proteins can be further investigated through competition test in which proteins without GFP tag and Deep-Purple are mixed with bacterial host. A decrease in the phage adsorption rate should be observed if those proteins play a role in the adsorption step as their interaction with the bacterial receptor should prevent phage adsorption. In addition, the protein location at the phage baseplate can be investigated by antibody cross-linking and transmission electron microscopy. It could also be interesting to explore whether carbohydrate binding modules identified in gp28 are responsible for host recognition alone and if this activity is restricted to one domain only. To do so, carbohydrate binding modules can be fused separately to GFP, assessed in cell wall decoration assay and compared to full-length gp28 fluorescence. In the same way, the cloning of only the IMC domain of gp29 can be performed to determine whether this domain is sufficient for host binding capacity. Finally, the host spectrum of RBP(s) can be compared to that of Deep-Purple.

The second focus of this work was the lysis process and the identification and characterization of the endolysin. Gp32 was the only endolysin candidate as it features numerous BLAST hits with endolysins. Interpro analyses allowed to define a N-terminal glycosyl hydrolase family 25 catalytic domain associated with 1,4-N-acetylmuramidase activity (putative EAD) and a C-terminal SH3 domain involved in host recognition (putative CBD).

To assess gp32 lytic activity, cloning of the full-length gene into pET30 vector was performed. In parallel, clonings of the CBD (gp32_CBD(196)), both the CBD and linker region (gp32_CBD(180)) and the full-length gene fused to GFP on the N-terminal were performed in order to assess gp32 host binding capacity. All constructions were then transformed into the expression strain BL21(DE3). Inductions yielded soluble proteins for all recombinant proteins and all proteins were successfully purified in native conditions.

Turbidity reduction assays confirmed gp32 endolysin activity on both *B. weihenstephanensis* LH002 cells and *B. cereus* ATCC10987 peptidoglycan. The activity spectrum was then assessed via turbidity reduction assay on a representative collection of strains. It showed that all tested members within the *B. cereus* group were sensitive to gp32 and some members outside of the *B. cereus* group also showed endolysin susceptibility. Then the lytic activity of gp32 was assessed over a broad range of pH (4 to 10), NaCl concentrations (0 to 1000 mM) and exposure to different temperatures (4 to 60°C). Gp32 remained active in conditions ranging from pH 4 to 8 and under 37°C. However, the endolysin revealed to be highly sensitive to NaCl concentrations. The effect of ions on gp32 activity remains to be investigated. However, I-TASSER analyses, aiming to predict protein function based on structural homology, showed gene ontology terms associated with potassium and magnesium binding capacity (respectively GO:0030955 and GO:0000287 terms) which suggests that the presence of these ions will certainly have an impact on gp32 activity. The ability of the enzymatic domain to show lytic activity without CBD was also briefly investigated. This has been reported to be the case for positively charged EAD (Low

et al., 2011) ; (Kashani et al., 2018). In our case, gp32_{lytic} has a isoelectric point evaluated to 5.75 (5.77 for the full-length endolysin). Preliminary results showed that gp32_{lytic} seemed to be able to lyse bacteria, at a lesser extend than gp32 full-length (Fig. S4) but this could not be confirmed when repeating the experiment. Those tests should therefore be repeated for a longer period and at a higher concentration as suggested in Lee et al. (2019) and at a pH inferior to its isoelectric point in order to ensure that the protein is positively charged. Eventually, zymogram could be performed to confirm gp32 enzymatic activity.

The binding capacity of gp32 and its CBD was assessed via cell wall decoration assay. Firstly, cell wall decoration of recombinant proteins on *B. weihenstephanensis* LH002 allowed to confirm that gp32 and its CBD tightly bind to Deep-Purple host cell. Therefore, the CBD ability to bind to a representative collection of bacteria was assessed and allowed to confirm the endolysin activity spectrum evaluated before. Deep-Purple host spectrum was also assessed by top agar experiments. The results showed that the endolysin activity spectrum was broader than that of Deep-Purple phage (*i.e.* 81% of the tested strains sensitive to the endolysin *versus* 19% sensitive to the phage).

It should also be noted that preliminary experiments on the putative Deep-Purple holin function have been performed. Induction of GFP-tagged gp30 showed high fluorescence spots restricted to some part of the bacteria which seems to be different from induction of other recombinant proteins containing GFP (Fig. S5). However, these subjective visual observations have to be taken very carefully and this analysis has to be completed with functional analyses including measurement of intracellular components released in the extracellular environment, resulting from bacterial lysis when cells are subjected to holin.

This master thesis has mainly investigated both adsorption and lysis processes of Deep-Purple phage infecting the *B. cereus* group. Our research has identified two proteins capable of interacting with *B. cereus* cells upon adsorption. Further work will determine whether both are acting as RBP or not. Our study also provides a substantial amount of information regarding the endolysin activity and gives the first information about the holin identification, leaving the ground for further investigations.

Supplementary Figures and Tables

Table S1: Predicted CDSs in bacteriophage Deep-Purple. Structural proteins in bold type. Adapted from (Hock et al., 2018).

Gene product	Protein size (aa)	Str- and	Putative function	Interesting features	Conserved domains (E-value)
gp01	73	-	Hypothetical protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	
gp02	605	-	DNA primase/helicase	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	
gp03	67	-	Transcriptional regulator	Unique to Deep-Purple	HTH XRE (1.34e-06)
gp04	41	-	Hypothetical protein		
gp05	140	+	Hypothetical protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	
gp06	206	+	DNA binding protein		MerR_1 (4.37e-04); OMS28_porin (2.97e-03); DUF3967 (2.54e-08)
gp07	88	-	Hypothetical protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	
gp08	351	+	Terminase small subunit	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	Terminase_2 (3.2e-22); PRK07353 (5.57e-03)
gp09	448	+	Terminase large subunit	Match with 100% coverage and 100% identity with <i>Streptococcus pneumoniae</i> protein	Terminase_3 (9.94e-43)
gp10	520	+	Portal protein		Phage_prot_Gp6 (3.05e-26)
gp11	55	+	Hypothetical protein		
gp12	83	+	Hypothetical protein	Match with 100% coverage and 100% identity with <i>Streptococcus pneumoniae</i> protein; No match with other phages	
gp13	377	+	Minor capsid protein		Phage_min_cap2 (7.01e-101)
gp14	52	-	Hypothetical protein	No match with other phages	
gp15	217	+	Scaffolding protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	Phage_GP20 (1.07e-28)
gp16	70	-	DNA binding protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	
gp17	59	-	Hypothetical protein	No match with other phages	
gp18	69	-	Hypothetical protein	No match with other phages	
gp19	300	+	Major capsid protein		
gp20	136	+	Hypothetical protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	Yqbg (7.37e-08)
gp21	114	+	Minor capsid protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	Minor_capsid_1 (3.49e-09)
gp22	116	+	Minor capsid protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	Minor_capsid_2 (1.25e-12)
gp23	153	+	Hypothetical protein	No match with other phages	
gp24	154	+	Capsid protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	
gp25	141	+	Hypothetical protein		Phage_TAC_8 (2.90e-03)
gp26	239	+	Hypothetical protein		Phage_Gp15 (4.15e-29)
gp27	1245	+	Tape measure protein	Match with carbamoyl-phosphate synthase large subunit	COG5412 (1.66e-13)
gp28	495	+	Dit protein	Match with distal tail protein	Sipho_tail (1.11e-25); PA_VSR (5.46e-03)
gp29	2156	+	Tail fiber	Match with tail fiber and C-ter part less conserved	Prophage_tail (2.06e-08); DUF4795 (2.03e-03); gp58 (4.71e-03); Peptidase_S74 (2.33e-06)
gp30	70	+	Holin		1 TM domain
gp31	240	+	Hypothetical protein	No match with other phages	Smc (3.39e-03)
gp32	262	+	N-acetylmutaroyl-L-alanine	Nearly matches only with <i>B. cereus</i>	GH25-PlyB-like (1.39e-80)
gp33	98	+	Hypothetical protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	
gp34	57	-	Hypothetical protein		
gp35	251	-	Hypothetical protein		
gp36	188	-	Hypothetical protein		
gp37	136	-	Hypothetical protein	No match with other phages	
gp38	146	-	Hypothetical protein		
gp39	161	-	Hypothetical protein	Match with 100% coverage and 100% identity with <i>B. cereus</i> protein	
gp40	178	-	Hypothetical protein	No match with other phages	

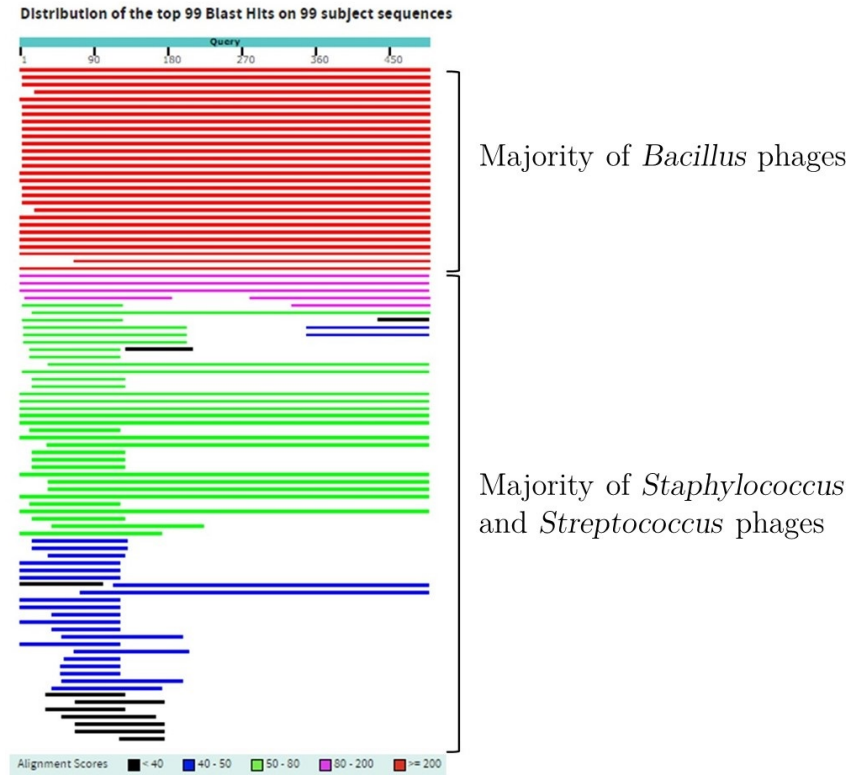


Figure S1: Graphic summary of BLASTp hits restricted to tailed phages for Deep-Purple gp28.

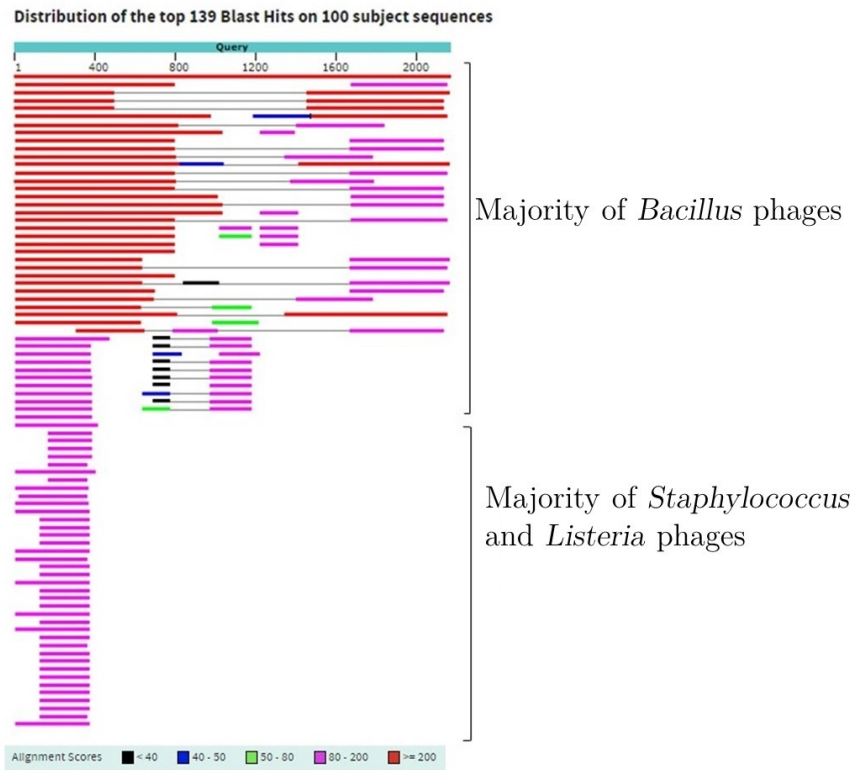


Figure S2: Graphic summary of BLASTp hits restricted to tailed phages for Deep-Purple gp29.

PBC5 MGTEI I D I S K W N A G S I N W D V L A P Q L G L C I C R V Q Y G S N K V D E W Y K N Y V E N L E A R N V P H A A Y A Y G C F V S I A D A K V E A E D F H N R
 DP - M K K I L D I S H - H - N T I N W D K V A G N V E F I I A R V Q D G S N T V D T K Y K S F V A N M K Q R G I P F G N Y A F C R F V S V N D A R V E A R D F W A R
 Bcu4 - M G Y I V D I S K W N - G D I N W D I A A G Q L D L A I A R V Q D G S N Y V D P M Y K S Y A A S M K A R N V P F G N Y A F C R F V S V E D A K V E A R D F W A R
 Bcp1 - M G Y I V D M S K W N - G S P D W D T A K G Q L D L V I A R V Q D G S N Y V D P V Y K D Y V A A M K A R N I P F G S Y A F C R F V S V E D A K V E A R D F W N R

PBC5 T S K N A R F L V L D V E D D T V A A M S A A G N L K D L A A S Q A F I D H L K A K G - W R V G F Y V S H H M Y G S Y N L Q S V K A D F Y W I P R Y S T S A P S
 DP G D K D A K F W V A D V E V K T M S - - - - - D M R A G T Q A F I D E L R A L G A K K V G L Y V G H H T Y E E F N A S Q I K A D F V W I P R Y G G K N P A
 Bcu4 G D K E A L F W V A D V E V Q T M G - - - - - D M Q A G T Q A F I D E L Y R L G A K K V G L Y V G H H T Y V S F G A K N I R C D F T W I P R Y G G S Q P D
 Bcp1 G D K D S L F W V A D V E V T T M S - - - - - D M R A G T Q A F I D E L Y R L G A K K V G L Y V G H H K Y E E F G A A Q I K C D F T W I P R Y G A - K P A

PBC5 Y Y C D L W Q Y A D G E T G G Y M N G V G K V D R N R L M N G R T L Q W F L G E T A P V S E P E G I G V A L S I Y P D G Y G V N L Y E R P - - - S D P I Y A G N
 DP Y P C D I W Q H S E - - - D E W I D G I G E C D S N T L L G S K G V S F Y I G N S S A P S I S N - - - - - G S T V V E R G V I R T T D R A N V R T G P S L S D P
 Bcu4 F P C D L W Q Y T E - - - T G N V P G I G K C D L N R L N S D K T L D W F T G K D G T V A P P P Q G A Y D S S W F T K Q T G V F T L D R T I N L R T A P F P N A P
 Bcp1 Y P C D L W Q Y D E - - - Y G Q V P G I G K C D L N R L N G D K S L D W F T G K G E E A V Q P P Q G N Y D S S W F T K Q N G V F T L D R T I N L R T A P F P T A P

PBC5 I T K K I P Y K V F A G Y W G G G D K D M I C L G G E K Q W A Y N K H F T I D W Y K V R S K Y P V G W G V N F Y D G P S G N F L G N I D G S E V Y N A H N R V G
 DP V V K T L P I N T D W E Y W S I E L - K - - - - - G D Y N W F N V G A G - - - - - Y V R G D N V K I - - - I -
 Bcu4 L I A Q L N A G D N V N Y E A Y G Y E K - - - - - D G Y V W L R Q N R G - - - - - N G N F G Y I A S G E T K N G Q - R I S
 Bcp1 L I A Q L N A G D N V T D G Y G Y E K - - - - - D G Y V W L R Q N R G - - - - - N G N Y G Y I A S G E T K N G Q - R I S

Figure S3: gp32 alignments with *B. cereus* phage PBC5 endolysin (Lee et al., 2019), Deep-Purple (DP) gp32, *B. cereus* phage Bcu4 putative endolysin and *B. cereus* phage Bcp1 endolysin (Kong and Ryu, 2015).

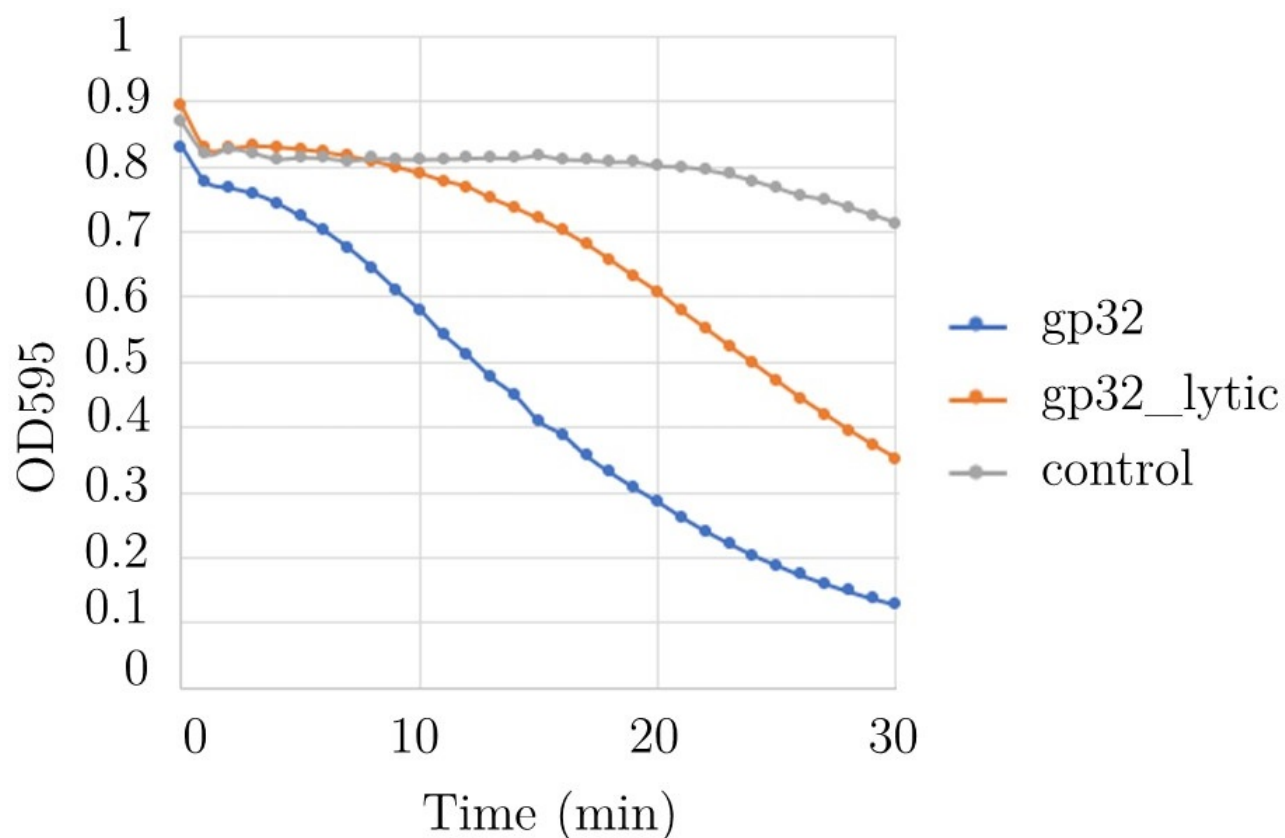


Figure S4: Turbidity reduction assay of gp32_lytic protein on *B. thuringiensis* GBJ001 cells. Cells treated with ca. 10 μ g of proteins.

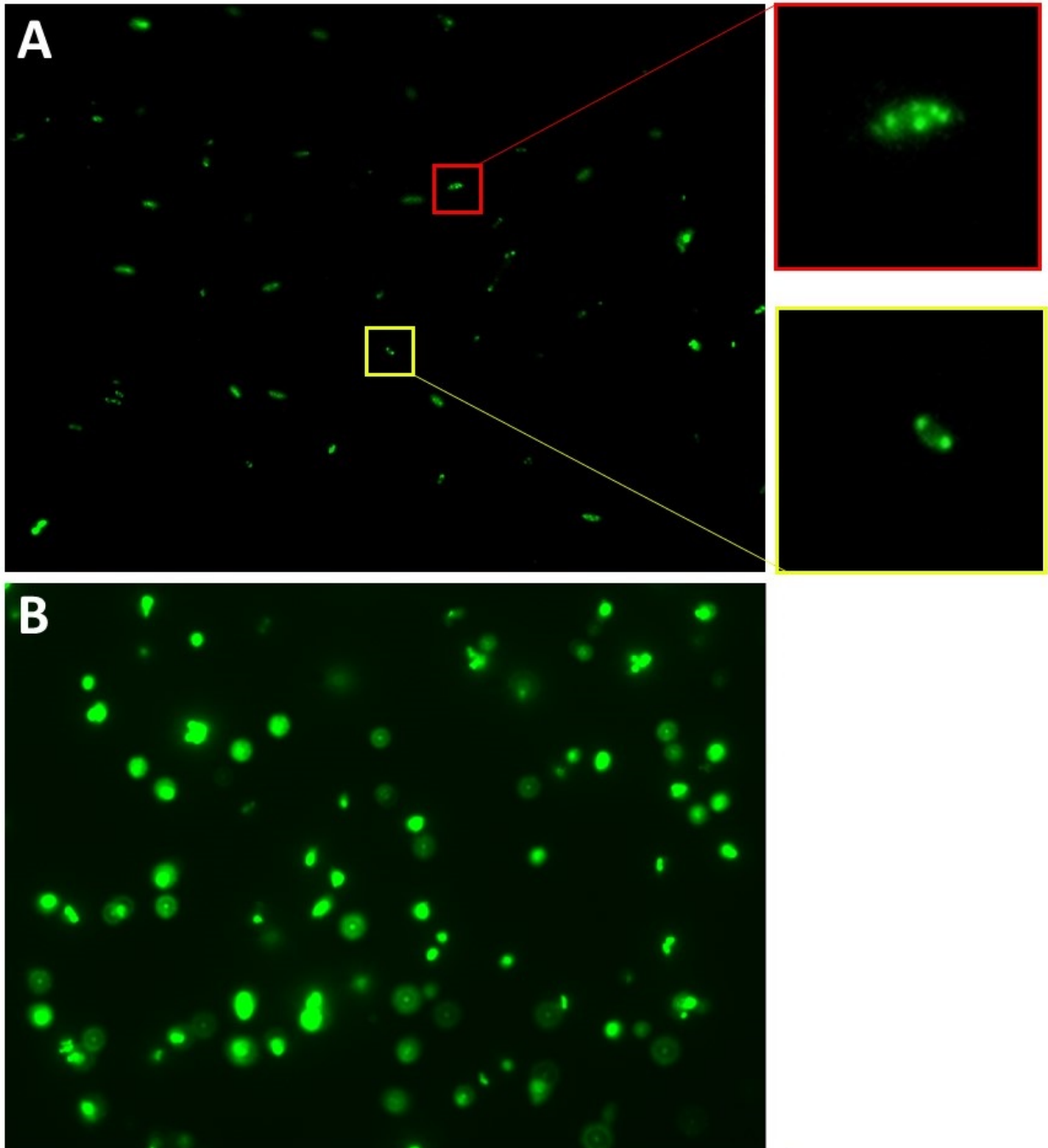


Figure S5: Induction of gp30::(link)GFP (A) and gp28::(link)GFP (B). *E. coli* cells observed after 5h of incubation at 28°C with 0.1 mM of IPTG.

References

- Abedon, S. T.
2011. Lysis from without. *Bacteriophage*, 1(1):46–49.
- Ackermann, H. W.
2006. Classification of bacteriophages. *The Bacteriophages*, 2:8–16.
- Ackermann, H.-W.
2007. 5500 phages examined in the electron microscope. *Archives of virology*, 152(2):227–243.
- Ackermann, H. W.
2011. Bacteriophage taxonomy. *Microbiol. Aust*, 32(2):90–94.
- Ackermann, H.-W. and G. Węgrzyn
2014. General characteristics of bacteriophages. *Phage Therapy*, Pp. 3–22.
- Authority, E. F. S., E. C. for Disease Prevention, C. (EFSA, and ECDC)
2018. The European union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. *EFSA Journal*, 16(12):e05500.
- Berry, J., M. Rajaure, T. Pang, and R. Young
2012. The spanin complex is essential for lambda lysis. *Journal of bacteriology*, 194(20):5667–5674.
- Bertozi Silva, J., Z. Storms, and D. Sauvageau
2016. Host receptors for bacteriophage adsorption. *FEMS microbiology letters*, 363(4):fnw002.
- Bottone, E. J.
2010. *Bacillus cereus*, a volatile human pathogen. *Clinical microbiology reviews*, 23(2):382–398.
- Bravo, A., S. S. Gill, and M. Soberon
2007. Mode of action of *Bacillus thuringiensis* cry and cyt toxins and their potential for insect control. *Toxicon*, 49(4):423–435.
- Broendum, S. S., A. M. Buckle, and S. McGowan
2018. Catalytic diversity and cell wall binding repeats in the phage-encoded endolysins. *Molecular microbiology*, 110(6):879–896.
- Brussow, H. and R. W. Hendrix
2002. Phage genomics: small is beautiful. *Cell*, 108(1):13–16.
- Cha, Y., B. Son, and S. Ryu
2019. Effective removal of staphylococcal biofilms on various food contact surfaces by *Staphylococcus aureus* phage endolysin LysCSA13. *Food microbiology*, 84:1–7.

- Chibeu, A., L. Agius, A. Gao, P. M. Sabour, A. M. Kropinski, and S. Balamurugan
2013. Efficacy of bacteriophage LISTEXTM p100 combined with chemical antimicrobials in reducing *Listeria monocytogenes* in cooked turkey and roast beef. *International journal of food microbiology*, 167(2):208–214.
- Cisek, A. A., I. Dąbrowska, K. P. Gregorczyk, and Z. Wyżewski
2017. Phage therapy in bacterial infections treatment: one hundred years after the discovery of bacteriophages. *Current microbiology*, 74(2):277–283.
- Dewey, J. S., C. G. Savva, R. L. White, S. Vitha, A. Holzenburg, and R. Young
2010. Micron-scale holes terminate the phage infection cycle. *Proceedings of the National Academy of Sciences*, 107(5):2219–2223.
- d’Herelle, M.
1917. Sur un microbe invisible antagoniste des bacilles dysentériques. *Comptes Rendus Acad. Sci. (Paris)*, 165:373–375.
- Di Franco, C., E. Beccari, T. Santini, G. Pisaneschi, and G. Tecce
2002. Colony shape as a genetic trait in the pattern-forming *Bacillus mycoides*. *BMC microbiology*, 2(1):33.
- Dierick, K., E. Van Coillie, I. Swiecicka, G. Meyfroidt, H. Devlieger, A. Meulemans, G. Hoedemaekers, L. Fourie, M. Heyndrickx, and J. Mahillon
2005. Fatal family outbreak of *Bacillus cereus*-associated food poisoning. *Journal of clinical microbiology*, 43(8):4277–4279.
- Dieterle, M.-E., S. Spinelli, I. Sadvskaya, M. Piuri, and C. Cambillau
2017. Evolved distal tail carbohydrate binding modules of *Lactobacillus* phage j-1: a novel type of anti-receptor widespread among lactic acid bacteria phages. *Molecular microbiology*, 104(4):608–620.
- Dramsi, S., S. Davison, S. Magnet, and M. Arthur
2008. Surface proteins covalently attached to peptidoglycan: examples from both Gram-positive and Gram-negative bacteria. *FEMS Microbiol Rev*, 32:307–320.
- Dunne, M., J. M. Denyes, H. Arndt, M. J. Loessner, P. G. Leiman, and J. Klumpp
2018a. *Salmonella* phage S16 tail fiber adhesin features a rare polyglycine rich domain for host recognition. *Structure*, 26(12):1573–1582.
- Dunne, M., M. Hupfeld, J. Klumpp, and M. J. Loessner
2018b. Molecular basis of bacterial host interactions by Gram-positive targeting bacteriophages. *Viruses*, 10(8):397.
- Etobayeva, I., S. Linden, F. Alem, L. Harb, L. Rizkalla, P. Mosier, A. Johnson, L. Temple, R. Hakami, and D. Nelson
2018. Discovery and biochemical characterization of Plyp56, Plyn74, and Plytb40—*Bacillus* specific endolysins. *Viruses*, 10(5):276.
- Eugster, M. R., M. C. Haug, S. G. Huwiler, and M. J. Loessner
2011. The cell wall binding domain of *Listeria* bacteriophage endolysin PlyP35 recognizes terminal GlcNAc residues in cell wall teichoic acid. *Molecular microbiology*, 81(6):1419–1432.
- Fagerlund, A., T. Lindbäck, and P. E. Granum
2010. *Bacillus cereus* cytotoxins Hbl, Nhe and CytK are secreted via the Sec translocation pathway. *BMC microbiology*, 10(1):304.

- Fenton, M., O. McAuliffe, J. O'Mahony, and A. Coffey
2010. Recombinant bacteriophage lysins as antibacterials. *Bioengineered bugs*, 1(1):9–16.
- Fischetti, V. A.
2005. Bacteriophage lytic enzymes: novel anti-infectives. *Trends in microbiology*, 13(10):491–496.
- Fruciano, D. E. and S. Bourne
2007. Phage as an antimicrobial agent: d'Herelle's heretical theories and their role in the decline of phage prophylaxis in the west. *Canadian journal of infectious diseases and medical microbiology*, 18(1):19–26.
- Gillis, A. and J. Mahillon
2014. Phages preying on *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*: past, present and future. *Viruses* 6: 2623–2672.
- Godinho, L. M., M. El Sadek Fadel, C. Monniot, L. Jakutyte, I. Auzat, A. Labarde, K. Djacem, L. Oliveira, R. Carballido-Lopez, S. Ayora, et al.
2018. The revisited genome of *Bacillus subtilis* bacteriophage spp1. *Viruses*, 10(12):705.
- González-García, V. A., M. Pulido-Cid, C. Garcia-Doval, R. Bocanegra, M. J. van Raaij, J. Martín-Benito, A. Cuervo, and J. L. Carrascosa
2015. Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *Journal of biological chemistry*, 290(16):10038–10044.
- Guerrero-Ferreira, R. C., P. H. Viollier, B. Ely, J. S. Poindexter, M. Georgieva, G. J. Jensen, and E. R. Wright
2011. Alternative mechanism for bacteriophage adsorption to the motile bacterium *Caulobacter crescentus*. *Proceedings of the national academy of sciences*, 108(24):9963–9968.
- Gutiérrez, D., L. Fernández, A. Rodríguez, and P. García
2018. Are phage lytic proteins the secret weapon to kill *Staphylococcus aureus*? *MBio*, 9(1):e01923–17.
- Habann, M., P. G. Leiman, K. Vandersteegen, A. Van den Bossche, R. Lavigne, M. M. Shneider, R. Biemann, M. R. Eugster, M. J. Loessner, and J. Klumpp
2014. *Listeria* phage A511, a model for the contractile tail machineries of SPO1-related bacteriophages. *Molecular microbiology*, 92(1):84–99.
- Hagens, S. and M. J. Loessner
2007. Application of bacteriophages for detection and control of foodborne pathogens. *Applied microbiology and biotechnology*, 76(3):513–519.
- Hatfull, G. F.
2015. Dark matter of the biosphere: the amazing world of bacteriophage diversity. *Journal of virology*, 89(16):8107–8110.
- Hayes, S., R. Vincentelli, J. Mahony, A. Nauta, L. Ramond, G. A. Lugli, M. Ventura, D. Van Sinderen, and C. Cambillau
2018. Functional carbohydrate binding modules identified in evolved dits from siphophages infecting various Gram-positive bacteria. *Molecular microbiology*, 110(5):777–795.
- Hock, L., A. Gillis, and J. Mahillon
2018. Complete genome sequence of bacteriophage deep-purple, a novel member of the family siphoviridae infecting *Bacillus cereus*. *Archives of virology*, 163(9):2555–2559.

- Howard-Varona, C., K. R. Hargreaves, S. T. Abedon, and M. B. Sullivan
2017. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *The ISME journal*, 11(7):1511.
- Huys, I., J.-P. Pirnay, R. Lavigne, S. Jennes, D. De Vos, M. Casteels, and G. Verbeken
2013. Paving a regulatory pathway for phage therapy. *EMBO reports*, 14(11):951–954.
- Hyman, P. and S. T. Abedon
2010. Bacteriophage host range and bacterial resistance. In *Advances in applied microbiology*, volume 70, Pp. 217–248. Elsevier.
- Hyman, P., P. Hyman, and S. T. Abedon
2012. *Bacteriophages in Health and Disease: Bacteriophages in Health and Disease*. Cabi.
- Jang, J., M. Cho, J.-H. Chun, M.-H. Cho, J. Park, H.-B. Oh, C.-K. Yoo, and G.-e. Rhie
2011. The poly- γ -D-glutamic acid capsule of *Bacillus anthracis* enhances lethal toxin activity. *Infection and immunity*, 79(9):3846–3854.
- Kashani, H. H., M. Schmelcher, H. Sabzalipoor, E. S. Hosseini, and R. Moniri
2018. Recombinant endolysins as potential therapeutics against antibiotic-resistant *Staphylococcus aureus*: current status of research and novel delivery strategies. *Clinical microbiology reviews*, 31(1):e00071–17.
- King, A. M., E. Lefkowitz, M. J. Adams, and E. B. Carstens
2011. *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*, volume 9. Elsevier.
- Knirel, Y. A. and M. A. Valvano
2011. *Bacterial lipopolysaccharides: structure, chemical synthesis, biogenesis and interaction with host cells*. Springer Science & Business Media.
- Köhler, T., V. Donner, and C. Van Delden
2010. Lipopolysaccharide as shield and receptor for R-pyocin-mediated killing in *Pseudomonas aeruginosa*. *Journal of bacteriology*, 192(7):1921–1928.
- Kong, M. and S. Ryu
2015. Bacteriophage PBC1 and its endolysin as an antimicrobial agent against *Bacillus cereus*. *Applied environmental microbiology*, 81(7):2274–2283.
- Kretzer, J. W., R. Lehmann, M. Schmelcher, M. Banz, K.-P. Kim, C. Korn, and M. J. Loessner
2007. Use of high-affinity cell wall-binding domains of bacteriophage endolysins for immobilization and separation of bacterial cells. *Applied environmental microbiology*, 73(6):1992–2000.
- Kumar, J. K.
2008. Lysostaphin: an antistaphylococcal agent. *Applied microbiology and biotechnology*, 80(4):555–561.
- Labrie, S. J., J. E. Samson, and S. Moineau
2010. Bacteriophage resistance mechanisms. *Nature Reviews Microbiology*, 8(5):317.
- Lapidus, A., E. Goltsman, S. Auger, N. Galleron, B. Ségurens, C. Dossat, M. L. Land, V. Brousolle, J. Brillard, M.-H. Guinebretiere, et al.
2008. Extending the *Bacillus cereus* group genomics to putative food-borne pathogens of different toxicity. *Chemico-biological interactions*, 171(2):236–249.

- Le, S., X. He, Y. Tan, G. Huang, L. Zhang, R. Lux, W. Shi, and F. Hu
2013. Mapping the tail fiber as the receptor binding protein responsible for differential host specificity of *Pseudomonas aeruginosa* bacteriophages PaP1 and JG004. *PloS ONE*, 8(7):e68562.
- Lee, K. O., M. Kong, I. Kim, J. Bai, S. Cha, B. Kim, K.-S. Ryu, S. Ryu, and J.-Y. Suh
2019. Structural basis for cell-wall recognition by bacteriophage PBC5 endolysin. *Structure*, 27(9):1355–1365.
- Legrand, P., B. Collins, S. Blangy, J. Murphy, S. Spinelli, C. Gutierrez, N. Richet, C. Kellenberger, A. Desmyter, J. Mahony, et al.
2016. The atomic structure of the phage Tuc2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules. *MBio*, 7(1):e01781–15.
- Leiman, P., S. Kanamaru, V. Mesyanzhinov, F. Arisaka, and M. Rossmann
2003. Structure and morphogenesis of bacteriophage t4. *Cellular and molecular life sciences CMLS*, 60(11):2356–2370.
- Liu, Y., Q. Lai, M. Göker, J. P. Meier-Kolthoff, M. Wang, Y. Sun, L. Wang, and Z. Shao
2015. Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Scientific reports*, 5:14082.
- López, R. and E. García
2004. Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. *FEMS microbiology reviews*, 28(5):553–580.
- Low, L. Y., C. Yang, M. Perego, A. Osterman, and R. Liddington
2011. Role of net charge on catalytic domain and influence of cell wall binding domain on bactericidal activity, specificity, and host range of phage lysins. *Journal of biological chemistry*, 286(39):34391–34403.
- Mojica, F. J., C. Díez-Villaseñor, J. García-Martínez, and C. Almendros
2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, 155(3):733–740.
- Nelson, D. C., M. Schmelcher, L. Rodriguez-Rubio, J. Klumpp, D. G. Pritchard, S. Dong, and D. M. Donovan
2012. Endolysins as antimicrobials. In *Advances in virus research*, volume 83, Pp. 299–365. Elsevier.
- Neufeld, T., A. Schwartz-Mittelmann, D. Biran, E. Ron, and J. Rishpon
2003. Combined phage typing and amperometric detection of released enzymatic activity for the specific identification and quantification of bacteria. *Analytical chemistry*, 75(3):580–585.
- Neuhaus, F. C. and J. Baddiley
2003. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiology and molecular biology reviews*, 67(4):686–723.
- Nicastro, J., S. Wong, Z. Khazaei, P. Lam, J. Blay, and R. A. Slavcev
2016. *Bacteriophage Applications-Historical Perspective and Future Potential*. Springer.
- Nobrega, F. L., M. Vlot, P. A. de Jonge, L. L. Dreesens, H. J. Beaumont, R. Lavigne, B. E. Dutilh, and S. J. Brouns
2018. Targeting mechanisms of tailed bacteriophages. *Nature reviews microbiology*, 16(12):760–773.

- Pickard, D., A. L. Toribio, N. K. Petty, A. van Tonder, L. Yu, D. Goulding, B. Barrell, R. Rance, D. Harris, M. Wetter, et al.
2010. A conserved acetyl esterase domain targets diverse bacteriophages to the vi capsular receptor of *Salmonella enterica* serovar typhi. *Journal of bacteriology*, 192(21):5746–5754.
- Priest, F. G., M. Barker, L. W. Baillie, E. C. Holmes, and M. C. Maiden
2004. Population structure and evolution of the *Bacillus cereus* group. *Journal of bacteriology*, 186(23):7959–7970.
- Rakhuba, D., E. Kolomiets, E. S. Dey, and G. Novik
2010. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. *Polish journal of microbiology*, 59(3):145–155.
- Rashel, M., J. Uchiyama, I. Takemura, H. Hoshiba, T. Ujihara, H. Takatsuji, K. Honke, and S. Matsuzaki
2008. Tail-associated structural protein gp61 of *Staphylococcus aureus* phage ϕ MR11 has bifunctional lytic activity. *FEMS microbiology letters*, 284(1):9–16.
- Reddy, B. L. and M. H. Saier Jr
2013. Topological and phylogenetic analyses of bacterial holin families and superfamilies. *Biochemical and biophysical acta (BBA)-Biomembranes*, 1828(11):2654–2671.
- Rivera, J., R. J. Cordero, A. S. Nakouzi, S. Frases, A. Nicola, and A. Casadevall
2010. *Bacillus anthracis* produces membrane-derived vesicles containing biologically active toxins. *Proceedings of the national academy of sciences*, 107(44):19002–19007.
- Salmond, G. P. and P. C. Fineran
2015. A century of the phage: past, present and future. *Nature reviews microbiology*, 13(12):777–786.
- Samson, J. E., A. H. Magadán, M. Sabri, and S. Moineau
2013. Revenge of the phages: defeating bacterial defences. *Nature reviews microbiology*, 11(10):675–687.
- Saren, A.-M., J. J. Ravantti, S. D. Benson, R. M. Burnett, L. Paulin, D. H. Bamford, and J. K. Bamford
2005. A snapshot of viral evolution from genome analysis of the Tectiviridae family. *Journal of molecular biology*, 350(3):427–440.
- Schmelcher, M., D. M. Donovan, and M. J. Loessner
2012. Bacteriophage endolysins as novel antimicrobials. *Future microbiology*, 7(10):1147–1171.
- Schmelcher, M. and M. J. Loessner
2014. Application of bacteriophages for detection of foodborne pathogens. *Bacteriophage*, 4(2):e28137.
- Schmelcher, M., T. Shabarova, M. R. Eugster, F. Eichenseher, V. S. Tchang, M. Banz, and M. J. Loessner
2010. Rapid multiplex detection and differentiation of *Listeria* cells by use of fluorescent phage endolysin cell wall binding domains. *Applied environmental microbiology*, 76(17):5745–5756.
- Schuch, R., A. J. Pelzek, D. C. Nelson, and V. A. Fischetti
2019. The PlyB endolysin of bacteriophage vB_BanS_Bcp1 exhibits broad-spectrum bactericidal activity against *Bacillus cereus Sensu Lato* isolates. *Applied and environmental microbiology*, 85(9):e00003–19.

- Schwarzer, D., F. F. Buettner, C. Browning, S. Nazarov, W. Rabsch, A. Bethe, A. Oberbeck, V. D. Bowman, K. Stummeyer, M. Mühlenhoff, et al.
2012. A multivalent adsorption apparatus explains the broad host range of phage phi92: a comprehensive genomic and structural analysis. *Journal of virology*, 86(19):10384–10398.
- Seltmann, G. and O. Holst
2013. *The bacterial cell wall*. Springer Science & Business Media.
- Servick, K.
2016. Beleaguered phage therapy trial presses on. *Sciences*, 352:1506.
- Sharma, S., S. Chatterjee, S. Datta, R. Prasad, D. Dubey, R. K. Prasad, and M. G. Vairale
2017. Bacteriophages and its applications: an overview. *Folia microbiologica*, 62(1):17–55.
- Siegel, S. D., J. Liu, and H. Ton-That
2016. Biogenesis of the Gram-positive bacterial cell envelope. *Current opinion in microbiology*, 34:31–37.
- Singh, A., D. Arutyunov, M. T. McDermott, C. M. Szymanski, and S. Evoy
2011. Specific detection of *Campylobacter jejuni* using the bacteriophage NCTC 12673 receptor binding protein as a probe. *Analyst*, 136(22):4780–4786.
- Söding, J., A. Biegert, and A. N. Lupas
2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic acids research*, 33(suppl_2):W244–W248.
- Squirrell, D., R. Price, and M. Murphy
2002. Rapid and specific detection of bacteria using bioluminescence. *Analytica chimica acta*, 457(1):109–114.
- Stenfors Arnesen, L. P., A. Fagerlund, and P. E. Granum
2008. From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS microbiology reviews*, 32(4):579–606.
- Thorsen, L., B. B. Budde, L. Henrichsen, T. Martinussen, and M. Jakobsen
2009. Cereulide formation by *Bacillus weihenstephanensis* and mesophilic emetic *Bacillus cereus* at temperature abuse depends on pre-incubation conditions. *International journal of food microbiology*, 134(1-2):133–139.
- Tran, T. A. T., D. K. Struck, and R. Young
2005. Periplasmic domains define holin-antiholin interactions in T4 lysis inhibition. *Journal of bacteriology*, 187(19):6631–6640.
- Twort, F. W.
1915. An investigation on the nature of ultra-microscopic viruses. *Lancet*, 2:1241–1243.
- Vollmer, W., D. Blanot, and M. A. De Pedro
2008. Peptidoglycan structure and architecture. *FEMS microbiology reviews*, 32(2):149–167.
- Walcher, G., B. Stessl, M. Wagner, F. Eichenseher, M. J. Loessner, and I. Hein
2010. Evaluation of paramagnetic beads coated with recombinant *Listeria* phage endolysin-derived cell-wall-binding domain proteins for separation of *Listeria monocytogenes* from raw milk in combination with culture-based and real-time polymerase chain reaction-based quantification. *Foodborne pathogens and disease*, 7(9):1019–1024.

- Walmagh, M., B. Boczkowska, B. Grymonprez, Y. Briers, Z. Drulis-Kawa, and R. Lavigne
2013. Characterization of five novel endolysins from Gram-negative infecting bacteriophages. *Applied microbiology and biotechnology*, 97(10):4369–4375.
- Wang, I.-N., D. L. Smith, and R. Young
2000. Holins: the protein clocks of bacteriophage infections. *Annual reviews in microbiology*, 54(1):799–825.
- Washizaki, A., T. Yonesaki, and Y. Otsuka
2016. Characterization of the interactions between *Escherichia coli* receptors, LPS and OmpC, and bacteriophage T4 long tail fibers. *MicrobiologyOpen*, 5(6):1003–1015.
- White, R., S. Chiba, T. Pang, J. S. Dewey, C. G. Savva, A. Holzenburg, K. Pogliano, and R. Young
2011. Holin triggering in real time. *Proceedings of the national academy of sciences*, 108(2):798–803.
- White, R., T. A. T. Tran, C. A. Dankenbring, J. Deaton, and R. Young
2010. The N-terminal transmembrane domain of λ s is required for holin but not antiholin function. *Journal of bacteriology*, 192(3):725–733.
- Yang, S., M. Sadekuzzaman, and S.-D. Ha
2017. Reduction of *Listeria monocytogenes* on chicken breasts by combined treatment with UV-C light and bacteriophage ListShield. *Food science and technology*, 86:193–200.
- Young, R.
2014. Phage lysis: three steps, three choices, one outcome. *Journal of microbiology*, 52(3):243–258.
- Young, R., N. Wang, and W. D. Roof
2000. Phages will out: strategies of host cell lysis. *Trends in microbiology*, 8(3):120–128.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des sciences

Place des sciences, 2 bte L6.06.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/sc