UCLouvain

FACULTE DES SCIENCES

Faculté des sciences

Adsorption-facilitating enzymes in phage Deep-Blue

Author : Almudena Briceño

Supervisor : Prof. Jacques Mahillon (UCLouvain / ELI / MIAE)

Scientific advisor : Audrey Leprince

Reviewers : Prof. Claude Bragard (UCLouvain / ELI / ELIM)

Prof. Henri Batoko (UCLouvain / LIBST)

Academic year: 2020-2021

Acknowledgements

I would like to start this master thesis report by thanking all those who contributed to make this work possible in spite of the challenges posed by this particular year.

To begin with, I would like to thank my thesis advisor, Professor Jacques Mahillon, for having warmly welcomed me into its laboratory. Thank you for your availability and guidance during this year.

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

I would especially like to thank my scientific advisor, Audrey Leprince, for devoting part of her time to patiently teaching me and sharing with me her captivating work. Your advice, both within the lab and for the writing, has been essential and led me in the right direction whenever I needed it. It has been a real pleasure to work by your side.

I am also grateful for the friendly atmosphere present at all times in the MIAE lab as well as for the kind support of all of its members.

Finally, I must express my profound gratitude to my friends and family for providing me with their unfailing support throughout all these years of study and specially throughout this challenging year. This accomplishment would not have been possible without them.

Gracias por la confianza y el apoyo que me han dado constantemente, su orgullo ha sido siempre mi motivación más grande.

Almudena Briceño

Abstract

Bacteriophages or phages are bacterial viruses for whom the interest has exponentially reemerged in the last decades due to their interesting antimicrobial properties. Thus, understanding how these viruses prey on their specific hosts is essential. Phage adsorption into their bacterial host is a critical step for triggering infection and requires the specific interaction between phage receptor-binding protein(s) (RBPs) and receptors on the bacterial surface.

This adsorption process is often helped by structurally-encoded enzymes, namely depolymerases and ectolysins. Phage-encoded depolymerases are able to facilitate the adsorption process by cleaving exopolysaccharides or capsular polysaccharides that may be exhibited at the bacterial surface. Ectolysins are peptidoglycan (PG)-degrading enzymes whose main role is to facilitate cell wall crossing and tail extension without compromising host cell integrity. This work aimed to identify such antibacterial proteins involved in the adsorption of phage Deep-Blue, a tailed phage belonging to the *Herelleviridae* family and preying on the *Bacillus cereus* group.

An in-depth bioinformatic analysis of the region encoding structural proteins of Deep-Blue's tail retained four ectolysin and two depolymerase candidates. The putative ectolysins gp189, gp191, gp182 and gp195 were investigated since they contained PG-degradation associated domains such as NlpC/P60 domain for gp189 and gp182. The two putative depolymerases, gp222 and gp223, located outside the structural region, were retained for their consistent matches with sugar-binding and polysaccharide degrading associated hits like a Pectate lyase domain in gp223.

All candidates, including truncated versions of gp189 (gp189_lyt) and gp191 (gp191_1000) to facilitate purification, were cloned in *Escherichia coli.* However, production and purification were only achieved for gp189, gp191_1000, gp222 and gp223. The truncated version of gp189, gp189_lyt, turned out to be lethal for the expression strain and further investigation is needed to determine the cause.

Purified protein candidates were tested to assess their activity. For depolymerases, gp222 and gp223 were spotted on sensitive bacterial lawns and were expected to produce a degradation halo, the hallmark of depolymerase activity. Unfortunately, no clearing of the bacterial lawns was ever observed and thus, a depolymerase function could not be attributed to any of the tested candidates. Regarding ectolysins, turbidity reduction assays were conducted where only gp189 displayed consistent lytic activity on all sensitive strains tested. Up to 44.4 % of growth inhibition was detected on the *B. weihenstephanensis* Si0239 sensitive strain. These results strongly suggest gp189 to be Deep-Blue's ectolysin. Additionally, the lysis from without phenomenon, in which a high virion concentration causes bacterial lysis but no phage progeny production, was also investigated since it constitutes an indicator of ectolysin activity. Adsorption tests revealed that this ectolysin-driven phenomenon was possible both in strains equipped and lacking the RBP receptor, which indicates that specific adsorption onto the bacterial surface is not essential for ectolysin activity.

Finally, given that purification by ultracentrifugation in a CsCl gradient could not be achieved, it is possible that Deep-Blue may not be stable in CsCl or viable after ultracentrifugation as it has been the case for other phages.

Abbreviations

aa	amino acid
bp	base pairs
Cas	CRISPR associated
CBM	Carbohydrate Binding Module
CDS	Coding DNA Sequence
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
DLA	Double layer agars
ELISA	Enzyme-linked immunosorbent assay
GFP	Green fluorescent protein
gp	Gene product
ICTV	International Committee on Taxonomy of Viruses
kb	Kilobase
LB	Lysogeny Broth
LPS	Lipopolysaccharide
LTF	Long Tail Fiber
MS	Mass Spectrometry
NAM	N-acetylmuramic acid
NAG	N-acetylglucosamine
OD	Optical Density
PCR	Polymerase Chain Reaction
PFU	Plaque Forming Unit
PG	Peptidoglycan
RBP	Receptor Binding Protein
RNA	Ribonucleic acid
s.l.	*Sensu lato*
sp.	unspecified species
spp.	Several unspecified species
s.s.	*Sensu stricto*
ssDNA	Single-stranded DNA
STF	Short Tail Fiber
TAL	Tail-associated lysin
TMP	Tape Measure Protein
tRNA	Transfer ribonucleic acid
WTA	Wall Teichoic Acid

Contents

Chapter I

Introduction

1.1. Bacteriophages

Bacteriophages (or phages) are natural bacterial parasites and are estimated to be the most abundant entities on our planet with approximately 10^{31} particles (Hatfull, 2015). Before antibiotics discovery in the late 1920s, bacterial infections were a real scourge and costed many lives. For this reason, phages beneficiated from special attention in the scientific and medical fields for almost two decades. Indeed, since their discovery independently by Frederick W. Twort in 1915 and Félix d'Hérelle in 1917, these bacterial viruses were used extensively to treat various infectious diseases (Twort, 1915; d'Herelle, 1917). Throughout his first therapeutic attempts, d'Hérelle even succeeded to cure five children suffering from dysentery. Unfortunately, the viral nature and the lack of consistency in phage therapy results led to an increased doubt regarding its efficacy. Together with the tremendous impact of antibiotic discovery, phage therapeutic potential was cast back into the shades in most western countries but not in some former USSR countries where they continued to be used extensively (Dublanchet & Fruciano, 2008). However, with the increasing concern around microbial resistance, scientists all over the world are regaining interest in these viruses and seen them again as allies in our perpetual battle against bacteria. Additionally, even though phages may have been overshadowed in the medical field, they did not entirely fade from view. Phage convenience and simplicity made them into an emerging model organism in the early 1940s that largely contributed to the development of molecular biology (Keen, 2015). Notably, the discovery that DNA was the hereditary material of life (Hershey and Chase, 1952) and the first genomes (RNA and DNA) to be completely sequenced were all phage-related (Fiers *et al.,* 1976; Sanger *et al*., 1977).

1.1.1. Biology and classification

In addition to being the most abundant entities on earth, another striking feature of these viruses is their remarkable diversity. They have been isolated from every possible environment in which bacteria exist and scientists are virtually sure that at least one phage infects every bacterial strain (Keen, 2015). Thus, it is not surprising that phages play a major role in keeping the ecosystem balanced and take part in several environmental processes such as carbon and oxygen production cycles (Danovaro *et al*., 2011). As a result, phages vary in size, shape and, above all, in the genome they contain, which can be double or single-stranded DNA or RNA. Typically, a phage particle length ranges from 24 to 200 nm and their genome size matches their diversity ranging from 4 to up to 600 kb (Brüssow and Hendrix, 2002).

Given their ubiquitous and abundant nature, deciphering the taxonomic characterization of phages is a challenging task. The International Committee on Taxonomy of Viruses (ICTV), as the official regulatory entity, classifies phages by taking several parameters into consideration such as their morphology (capsid size and shape), type of genetic material, host range, genome size, resistance to organic solvents but mainly genomic comparison (Sharma *et al.*, 2017).

The overwhelming majority of phages are grouped under the *Caudoviridae* order, gathering tailed phages (Fig.I.1). Together they account for 96% of reported bacterial viruses (Ackermann and Prangishvili, 2012). Their head, an icosahedral protein shell, encloses a single linear double strand DNA molecule (Sharma *et al.*, 2017). They use their sophisticated tails to breach their host cell envelope and hence, inject their genetic material into the bacteria upon specific binding to its cell surface receptors. Consequently, they often show great specificity for one bacterial species or even a few strains in it. It is also common to find fibers, spikes and baseplate proteins attached to the tail in order to help the specific binding (Koskella and Meaden, 2013).

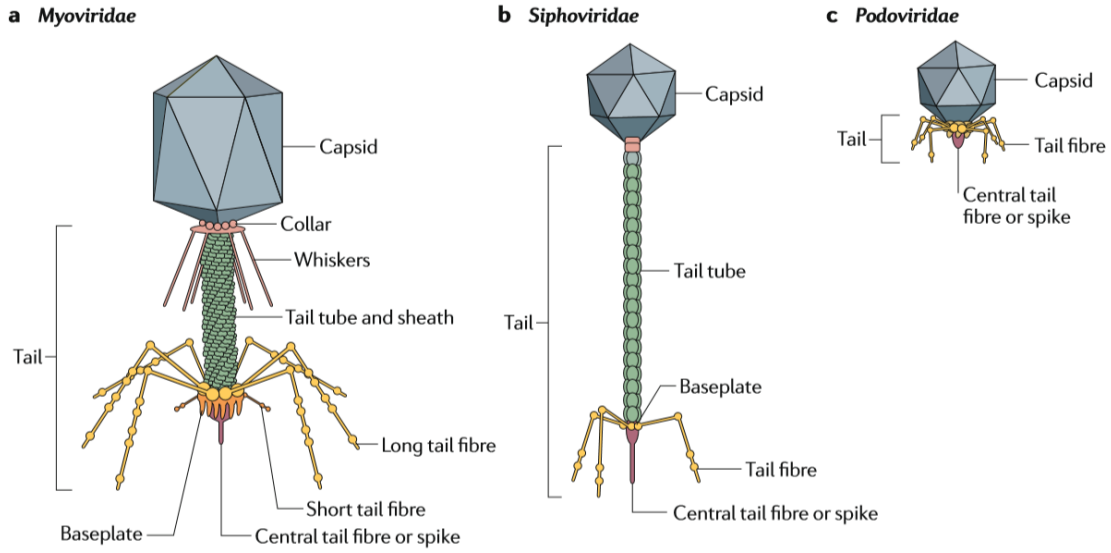

Figure I.1. Schematic representation of tailed phages based on morphological features. Myoviridae possess a long contractile tail contrary to Siphoviridae whose long tail is not contractile. Podoviridae displays a short tail (Nobrega et al., 2018).

For many years, three morphologically based families were recognized under this order, namely the *Myoviridae,* possessing long contractile tails, the *Siphoviridae,* with long non-contractile tails and the *Podoviridae*, featuring short non-contractile tails (Ackermann and Prangishvili, 2012). However, in the last years, genetic analysis of these viral families revealed that inside of the same family some members turned out to be more divergent than expected (Aiewsakun *et al*., 2018). Many reasons reside behind the uncertainty of phage taxonomy, for instance, the unparalleled diversity in genome type and size does not help the task but above all, the lack of universally present genes (*i.e.* analog to 16S RNA in prokaryotes) greatly hinder their accurate classification (Holmes, 2011). Hence, two new families, the *Ackermannviridae* and *Herelleviridae,* were added to the former of the *Caudoviridae* order in 2018.

1.1.2. Two life cycles: Lytic and lysogenic phages

Similar to other viruses, phages are dependent on a host for production of the next phage generation. In order to do so, after complex bacterial recognition processes (detailed in section I.1.2), phages are able to inject their genetic material into their host. At this point, two different life cycles are possible, the lytic or the lysogenic (Fig. I.2.).

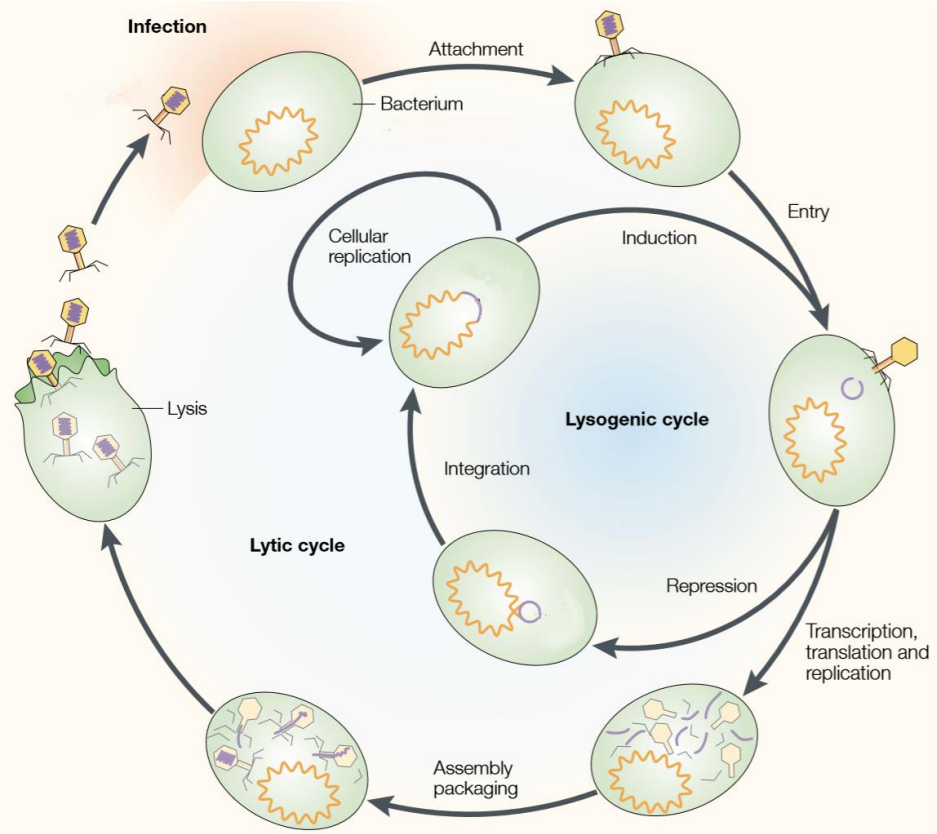

Figure I.2. Lytic and lysogenic cycles. The virus attaches specifically to the bacterial surface and injects its DNA. In the lytic cycle, the host replication machinery is coerced into replicating and transcribing the viral DNA in order to form the new virion particles. In the lysogenic cycle, phage DNA is integrated into the host genome and lysogenic activity is repressed. No viral particles are produced unless lytic cycle is induced (modified from Campbell, 2003).

Lytic cycle

Infections quickly leading to the release of phage progeny are called "productive infections". In the lytic cycle, the infected cell is reprogramed into a phage factory in which components of the cell biosynthetic machinery are hijacked to produce new virions and release them after cell lysis. Generally, DNA replication occurs first and then allows the synthesis of structural proteins needed to assemble the new virions. Before reaching a threshold number of virions, phages produce two main proteins involved in cell lysis. The first one, the holin, is a por- forming protein possessing at least one transmembrane domain whereas the second one, the endolysin, is a peptidoglycan (PG) degrading enzyme that accumulates in the cytoplasm while pathways for virion assembly proceed (Young, 2014). Subsequently, holin forms non-specific channels or micron-scale holes in the cytoplasmic membrane allowing the endolysins to reach the PG leading to the lysis of the host cell (Young, 2014). The lytic cycle can last *ca.* 40 min and produce up to ~ 100 new phages (Campbell, 2003).

Lysogenic cycle

In contrast, the lysogenic cycle performed by the so-called temperate phages, is characterized by a stable and long-lasting relation in which the phage genome is integrated into, and replicated with, the genome of the bacterial host in the form of a prophage. They may also integrate into extrachromosomal plasmids or not integrate at all and remain as linear or circular plasmids (Nobrega *et al.*, 2018). For this persistence infection to be possible, many lytic-life-associated genes are repressed thus allowing the bacterial host to divide and transmit the prophage vertically. However, in response to external stress, they can switch strategies and enter a lytic

cycle in order to lyse the cell and find a more stable host in a process called induction. In this case, the repressive activity stops and transcription is triggered in order to produce the new virions (Gandon, 2016).

1.1.3. Phage applications

The renewed interest for phages in the last decades has shed light on new applications for these viruses. Indeed, the emergence of multidrug-resistant bacteria demands, more than ever, innovative solutions such as exploiting phage-specific antibacterial potential as replacement of antibiotics (Kim *et al.,* 2019). However, the medical field is far from being the only one to harness phage's simplicity and diversity. Phages and their associated enzymes have proven their potential in a wide range of domains such as biotechnologies, food safety and bio-detection, among many others (Nicastro *et al.*, 2016).

Biocontrol potential

The natural specificity of phages for their bacterial targets (*i.e.* one species or a few strains) is the main advantage when compared to antibiotics. They are "precision" antibacterials that have, for instance, no effect on human beneficial microflora (Cieplak *et al.,* 2018). Moreover, they can be considered as a "self-replicating" drug since their effect is amplified at the end of each lytic cycle. Finally, due to their specificity, they do not represent any risk for human health and side effects have been rarely reported (Kim *et al.*, 2019). Despite their advantages, many unresolved difficulties still need to be addressed. For instance, their acute specificity is a double-edged knife since it requires the knowledge of which pathogens are been treated and reduces effectiveness to treat infections colonized with more than one bacterial type (*e.g.* burn wounds). To try to match broad-spectrum antibiotics, phage cocktails have been developed (Lin *et al.*, 2017). Furthermore, phage material is immunogenic, meaning it can trigger host immune response, eventually resulting in inactivation of phage particles. In the light of these issues and thanks to the advancements in genome edition tools, phage engineering is showing as a potential solution. For instance, an amino acid (aa) substitution in a capsid protein of λ phage of *Escherichia coli* enhanced its survival in mouse circulatory system more than a 1000-fold (Vitiello *et al.,* 2005). More recently, the world first controlled clinical trial for CRISPR enhanced bacteriophage therapy was launched against *E. coli* causing urinary tract infections (crPhage™ product by Locus Biosciences, 2020).

Tackling phages potential has been much harder in the medical field than in other sectors with fewer regulatory hurdles such as the food industry where they are already widely used. For instance, the approval of phage-derived products such as SALMONELEX (Micreos Food Safety) for the control of *Salmonella* or EcoShield (Intralytix) targeting *E. coli* O157:H7 as a food processing aid in the USA confirms the feasibility of this application (Rodriguez-Rubio, 2016).

Bio-detection

Standard methods for bacterial detection such as PCR-based, immunological (*e.g.* enzyme-linked immunosorbent assay, ELISA), or simply cultivation techniques tend to be time-consuming, more intricate and often require pre-enrichment steps and expensive machinery (Schmelcher and Loessner, 2014). From the recognition of their target host to the final cell lysis event, every step of the cycle can be harnessed for bacterial detection.

In the so-called phage amplification assays (Fig. I.3), the sensitive detection of specific bacteria can be achieved easily within four hours by following four main steps. Firstly, phage infection of the targeted bacteria is allowed, followed by the addition of a viricide to ensure survival only of phage particles engaged in the infectious process. To amplify the signal of each infection, helper cells (propagating strains for the chosen phage) are subsequently added to the mixture. Finally, counting the number of plaques gives an accurate estimate of the number of targets initially present in the sample. Due to the nature of this assay, only living cells will be detected (Stewart *et al.,* 1998).

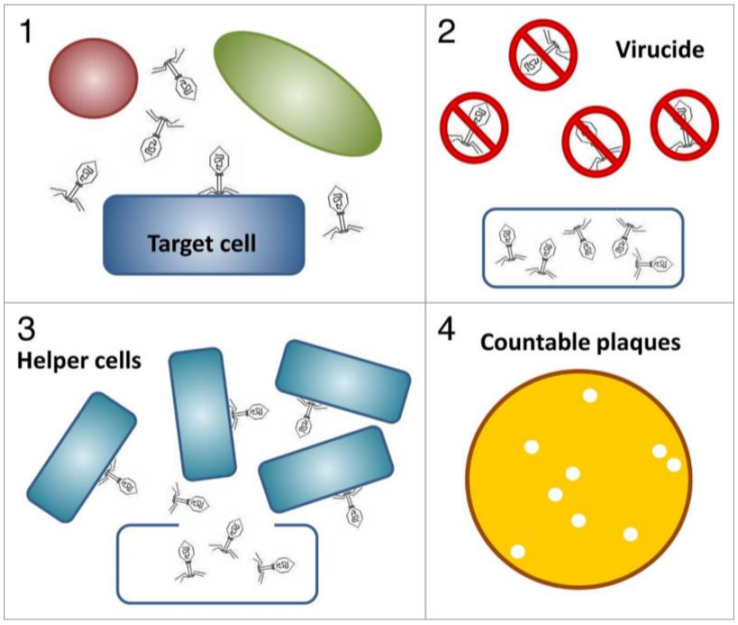

Figure I.3. Phage amplification assay using unmodified phage particles for bacterial detection. 1) Bacterial sample to be tested are mixed with specific phages; 2) extracellular phages are eliminated by adding a viricide; 3) After cell lysis, phage progeny is released and infect added helper cells for amplification of the signal; 4) Each plaque correspond to an initial infection of a bacterial target cell in the sample (Schmelcher and Loessner, 2014).

However, more recent and sophisticated detection strategies can offer real-time results by combining the targeting abilities of phages with a method for detecting the phage itself. This engineered phage is referred to as reporter phage, which is usually produced, by integrating a gene encoding bioluminescent, fluorescent or chromogenic properties into the phage genome. For instance, in a fluorescent-based assay, genes encoding for luciferase or Green Fluorescent Protein (GFP) expressed exclusively upon infection of the bacterial target could be used. (Nicastro *et al*., 2016)

Other methods harnessing phages or phage components as affinity molecules have also emerged. The first ones focus on labeling the entire phage particle in order to detect its attachment to the targeted bacteria without the necessity of infection. *E.coli*- directed phages chemically labeled with horseradish peroxidase (HRP) provide a good example of enzyme-labeled phages, also called Phazymes. Indeed, after Phazymes are added to immobilized cells of a sample, addition of luminescent HRP substrate is sufficient to easily detect targeted *E. coli* cells (Willford *et al.,* 2011). Individual components, rather than entire phage particles, may also be used as affinity molecules in this context, especially phage enzymes such as endolysins and their cell wall binding domains (CBD), which help host recognition. In fact, CBD-coated paramagnetic beads have been used to detect Gram-positive pathogens, such as *Listeria monocytogenes*, in food samples by affinity binding and subsequent recovery of the beads via a magnet. Because

of their high-affinity binding and specificity, CDB-coated beads were able to immobilize and recover 90% of pathogen cells in the sample within 20 to 40 min (Kretzer *et al.,* 2007; Schmelcher and Loessner, 2014).

1.1.4. Coevolution with bacteria

Given their ubiquitous distribution and overwhelming numbers, it is not surprising that phages play important roles in shaping biological interactions and regulating environmental processes. For instance, it is estimated that every day, phages are responsible for the killing of 15% to 40% of the ocean's bacterial population thus influencing many ecological parameters such as oxygen production by phytoplankton and carbon dissolving rate among many others (Danovaro *et al.,* 2011). Since phages constantly apply selective pressure on bacteria and are responsible for a significant part of horizontal gene transfer events (mostly via lysogenic phages), they can be considered top drivers of bacterial evolution (Chibani-Chennoufi *et al.,* 2004). As a result, both parties have developed resistance mechanisms and mechanisms to bypass those of their opponents in an everlasting co-evolutionary arms race.

Main Bacterial resistance mechanisms

Each step of phages replication cycle can be targeted by bacterial resistance mechanisms. The most widespread mechanisms consist in the prevention of phage adsorption (Fig. I.4). For instance, bacteria may modify its cell surface receptors in order to become resistant. Moreover, since phage receptors often have important metabolic roles (*e.g*. *E. coli* FhuA is an iron transporter and also a receptor for coliphages such T5), some bacteria take advantage of it by producing competitive inhibitors making these receptors unavailable for phage adsorption (microbial molecule microcin J25 also uses FhuA as a receptor and can outcompete phage T5). In the same way, *Staphylococcus aureus* can mask the phage receptor by synthesizing protein A (Fig. I.4a). Other bacteria, like *Salmonella* spp., produce extracellular polymers providing a physical barrier between phages and their receptors (Fig. I.4b) (Labrie *et al.,* 2010).

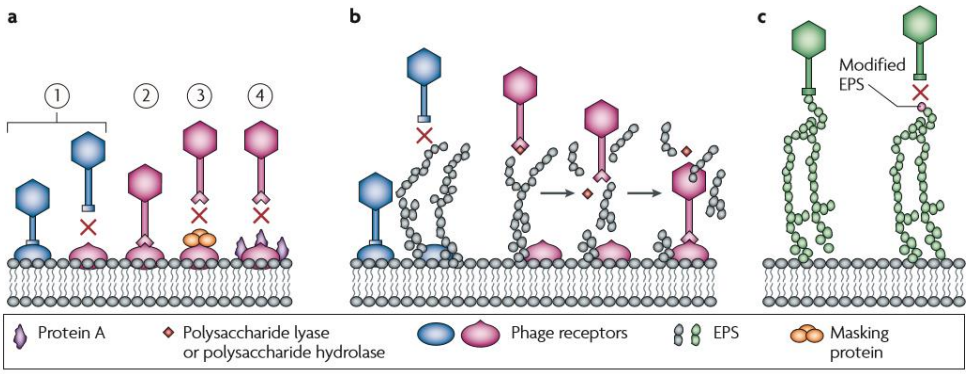

Figure I.4. Bacterial mechanisms for blocking phage adsorption. a) Bacteria can modify their cell surface receptors in order to become resistant (step1); phages can respond by adapting to this newly expressed receptor (step2). Bacteria may also produce receptor-masking proteins (step3) or, in the case of S. aureus, protein A reduces phage adsorption. b) Bacteria can physically separate themselves from phages by producing exopolysaccharides layer. Phages may develop lytic enzymes (i.e. depolymerases) to cleave them and access the receptor. c) Phages may also adapt to specifically recognize polysaccharides directly (Labrie et al., 2010).

Preventing phage DNA entry can be achieved thanks to superinfection exclusion (Sie) systems. These systems are mostly encoded by prophages, which helps them prevent subsequent infection by other similar phages and indicates their importance for phage-phage interactions rather than phage-bacteria interactions. Coliphage T4 has two well-characterized Sie systems, *imm* and *sp*, that block the translocation of related viral DNA and inhibits the activity of PG-degrading T4 lysozyme, respectively (Labrie *et al.,* 2010).

Bacteria also possess systems allowing them to defend themselves even after successful injection of viral DNA. Namely, restriction-modification systems (RM) are present in almost all bacteria and are capable of recognizing and cleaving unmethylated DNA that enters the cell (Labrie *et al.,* 2010). However, phage infection is successful whenever, instead of being enzyme-restricted, phage DNA gets modified (e.g. methylated) and thus, becomes immune to the system. Unsurprisingly, the incidence of this event is much lower than that of the restriction scenario and thus, incoming phage DNA is usually cleaved (Labrie *et al*., 2010). Another mechanism recently described as a primitive bacterial adaptive immunity system is the renowned CRISPR-Cas system (clustered regularly interspaced short palindromic repeats) (Fig, I.5.). Indeed, upon DNA injection, a small sequence of the phage's genome, the spacer, is cut out and integrated into the CRISPR locus, thus providing immunity against that specific phage. This is achieved by using spacers of invading phages to produce small CRISPR RNAs (crRNAs) that act as guides for Cas-encoded nucleases. Upon entrance of a previously encountered phage DNA material, the complex formed by the guide crRNA and the nucleases is able to locate and cut through it, thus inactivating the phage, and preventing infection (Marraffini, 2015).

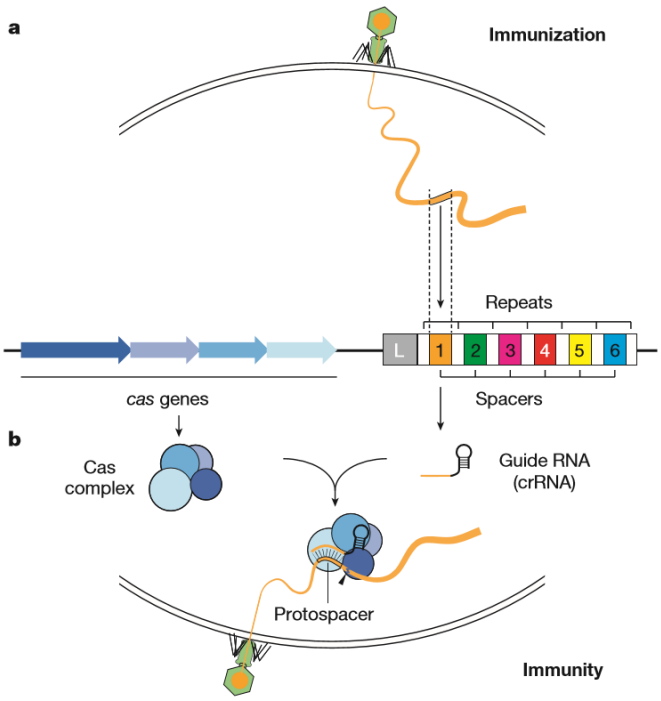

Figure I.5. CRISPR-Cas mediated immunity in prokaryotes. CRISPR loci are composed of alternating short DNA repeats (white) and short spacer sequences of phage origin (colored boxes). CRISPR-associated (Cas) genes (blue arrows) code for enzymes necessary for immunization and immunity processes. During immunization, a spacer sequenced is cut out from the genome of an invading phage and integrated into the CRISPR loci (a). In the immunity context, the previously integrated protospacer is used to build a small CRISPR RNA (crRNA) that acts as a guide for Cas-encoded nucleases and thus allows it to recognize and cleave the target sequence (black arrow) (Marraffini, 2015).

Last but not least, if a phage successfully entered and avoided the previously mentioned defense mechanisms, some bacteria will trigger their own death before contributing to the spread of more infectious particles. These are known as abortive infection (Abi) systems and regroup host mechanisms that interrupt phage development. Although many Abi systems have been identified, their modes of action remain largely unknown. Some, however, have been elucidated like in *E. coli* K-12, where the Lit protease is only activated in the presence of Gol protein, produced by phage T4. Upon infection, Lit becomes active and cleaves Ef-Tu, the translation elongation factor, causing translational arrest and subsequent cell death (Stern and Sorek, 2011).

Main phage counter mechanisms

Given the dynamic relation linking phages and their hosts, it is not surprising to see so many strategies used by phages to subvert these resistance processes. Indeed, they may evolve in order to counter the changes made by bacteria into their original receptor (Fig.I.4a) or by producing depolymerases if the accessibility to the receptor is the one compromised (Fig. I.4b). Lastly, they can also adapt to specifically bind to those polysaccharides (Fig. I.4c) (Labrie *et al.,* 2010).

Concerning RM systems, some phages acquired their own methyltransferase or mimic that of their host in order to protect their genome. Other phages encode proteins targeting the restriction endonucleases responsible for foreign restriction activity. This may be accomplished by producing small DNA-mimicking peptides blocking their active sites or by incorporating unusual bases into their genomes thus hindering their recognition. Regarding the CRISPR-Cas system, phages having integrated mutations, recombined or lost the sequence where their proto-spacer was located, gain CRISPR resistance throughout their second infection (Stern and Sorek, 2011). To combat more directly CRISPR defense mechanism, phages have also developed inhibitor proteins called anti-CRISPRs (Acrs) which inhibit DNA-binding or prevent nuclease-mediated degradation. Interestingly, a study demonstrated that Acrs is an imperfect CRISPR inhibitor as complete inactivation of this system requires an Acrs concentration only achievable by multiple phage genomes (Fig. I.6). This entails that after initial phage infection, CRISPR may be effective in avoiding viral coercion despite Arcs production. However, the accumulation of these proteins enhances the likelihood of subsequent successful phage infection and replication thus showing, in a certain way, the same altruism-like behavior seen in bacteria abortive infection strategy (Borges *et al.,* 2018).

In the same manner, mutations in phage peptides responsible for bacterial abortive mechanism activation result in rescue of the infecting phage. For instance, a mutation in Gol protein of *E.coli* infecting phage T4 suppresses Lit activation and allows the phage lytic cycle to continue (Stern and Sorek, 2011).

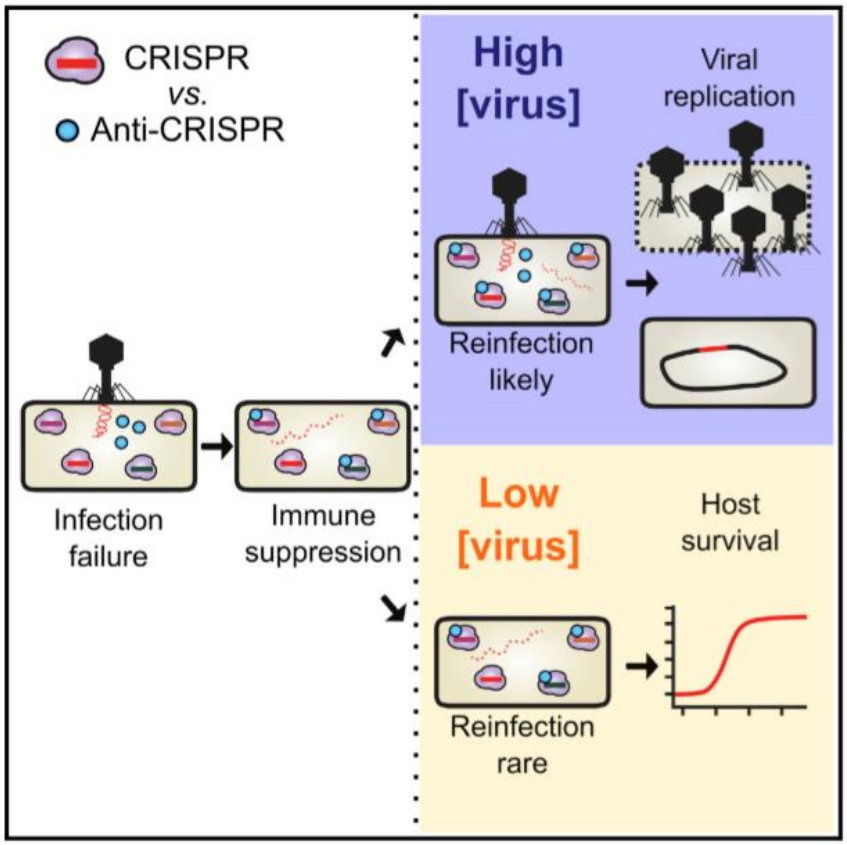

Figure I.6. Anti-CRISPR proteins (Acrs) are imperfect CRISPR inhibitors. A critical level of Acrs proteins is required for effective CRISPR resistance and infection of the immune hosts. This suggesting that Acrs may have emerged in conditions where the likelihood of multiple phage infection is elevated (Borges et al., 2018).

1.2. Phage adsorption process

Phage adsorption to the host cell surface initiates the infection process and is a *sine qua non* condition for its onset. This crucial step is achieved by interaction of the phage receptor binding protein(s) (RBP) and receptors on the bacterial surface, thus defining host range specificity. In this way, viral particles recognize potentially sensitive targets and prepare their subsequent DNA injection. Generally, this process is broken down into three steps: initial contact, reversible binding and ultimately, irreversible attachment (Bertozzi Silva *et al.,* 2016).

Initial contact happens as a result of Brownian motion of phage particles diffusing over the bacterial cell surface and resulting in many collisions. If one of these interactions results in specific “fitting” of the elements, it leads to the formation of a weak and reversible bond. This bond is weak enough to be broken by intense agitation but strong enough to keep the complex together until a more stable and irreversible attachment is managed. A model also pointed out that three-dimensional diffusion is followed by a more effective two-dimensional “random walking” over the bacteria surface in search of the appropriate receptor (Adam and Delbruck, 1968; Storms and Sauvageau, 2015). More recent studies have validated these hypotheses and even showed that some phages, such as *E. coli* phage T4 and T7, have evolved to use their tail fibers to maintain weak and reversible interactions allowing them to “dabble” the bacterial surface until irreversible attachment (Fig. I.7) (Hu *et al.,* 2013).

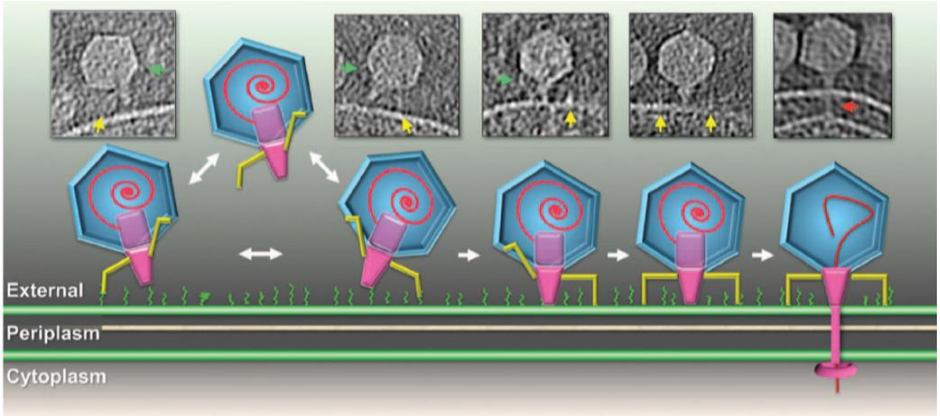

Figure I.7. Schematic model of E. coli phage T7 infection. Cryo-electron tomography technique was used to capture the different orientations of tail fibers (yellow and green arrows show bound and unbound fibers respectively). "Walking" movement allows T7 phage to optimize receptor search (Hu et al., 2013).

Finally, upon irreversible attachment, several conformational rearrangements are triggered in order to allow insertion of the genetic material into the cell host. It is important to note that this specific interaction between receptors and RBPs may be facilitated by enzymatic cleavage (Bertozzi Silva *et al.,* 2016).

1.2.1. Bacterial receptors

A number of studies focusing on host receptors for phage absorption highlighted the variability in the nature of these molecules, which can be proteins and polysaccharides. Moreover, their location may also vary greatly with receptors being identified in cell walls, slime layers, capsules, appendages and even flagella. Sometimes, adsorption to two receptors is needed to trigger infection (Bertozzi Silva *et al.,* 2016). This variability probably emerges from the adaptations developed by phages to overcome the defense mechanisms displayed by their bacterial counterparts. Additionally, the receptors involved in reversible and irreversible binding are occasionally different, with irreversible receptors being more exposed and easier to access in order to increase stability and probability to find the irreversible binding associated receptor (Nobrega *et al.,* 2018).

Bacterial cell type also influences the adsorption of phages due to their important structural differences such as thickness of the PG layer, uniformity and lipid content, among others (Fig. I.8.). In Gram-negative bacteria, which have been historically more studied, the cell surface is constituted by the outer membrane (OM) and the PG layer in between the two membranes is thinner than that of Gram-positive bacteria (Vollmer *et al.,* 2008). Moreover, the outer leaflet of the OM is characterized by the presence of lipopolysaccharides (LPS). The most conserved part of an LPS molecule, the lipid A, is attached to the outermost part, the O-antigen, through the oligosaccharide core. LPS containing all three components are denominated as smooth whereas those lacking the O-antigen are called rough because they often have less regularly shaped colonies than that of their smooth strain counterparts (Letarov and Kulikov, 2017; Nobrega *et al.,* 2018). Given that the saccharides composing the O-antigen are highly variable, unlike the ones that make up the core, phages targeting smooth strains tend to use O-antigens as receptors and consequently have a narrower host range (Rakhuba *et al.,* 2010).

Besides LPS, many phages can also target the variety of exopolysaccharides (*i.e.* capsule and slime) produced by bacteria, whose one of their many roles is precisely to physically separate them from these viruses. For example, *E. coli* strains have more than 80 types of K-antigen (capsule polysaccharides), which is a key factor of pathogenicity and make up a very efficient barrier against phages but can also be used as primary receptors (Whitfield, 2006; Letarov and Kulikov, 2017). Additionally, some phages may use enzymatic activity to degrade capsular polysaccharides and reach cell surface receptors (Porter and Martens, 2017). Adsorption to other receptor types such as porins and other proteins of the outer membrane is quite common. These proteins, allowing the cell to be highly permeable to water and ions, can be harnessed by phages as primary and secondary receptors (*e.g.* phage T4 can use porin OmpC or LPS for adsorption) (Letarov and Kulikov, 2017). Finally, protuberances and appendices on OM surface such as flagella or pili can be used as receptors but they usually work as primary receptors used by phages as conveyance towards the cell surface (Letarov and Kulikov, 2017).

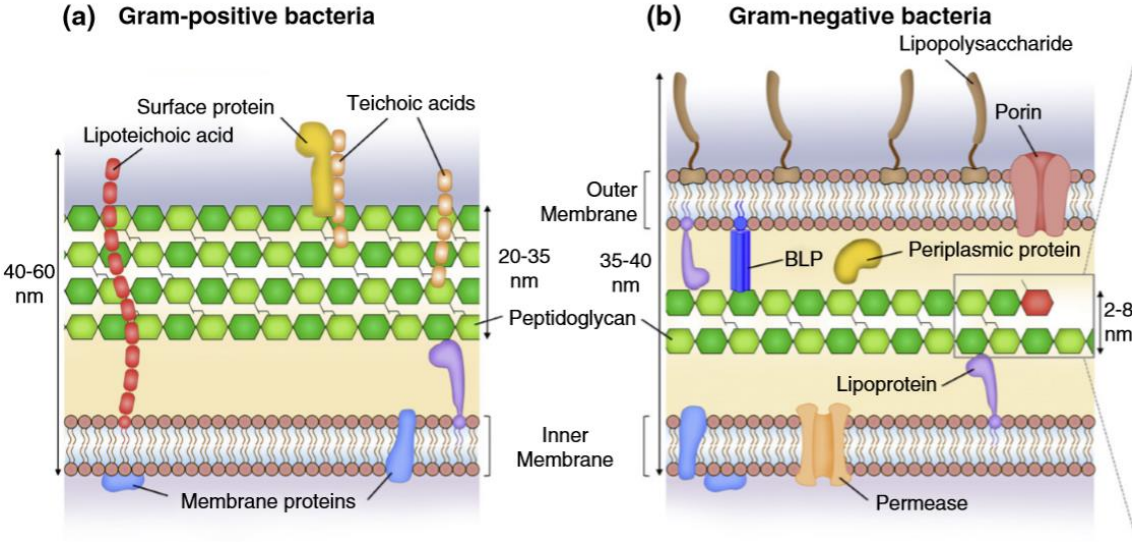

Figure I.8. Schematic organization of Gram-positive (a) and Gram-negative (b) cell wall. The PG layer is composed of linear glycan strands made up by alternating N-acetylglucosamine (NAG) (dark green) and N-acetylmuramic acid (NAM) (light green) residues linked by β(1-4) bonds. These strands are cross-linked by short and variable peptides. BLP stands for Braun's lipoprotein attached to both PG and the outer membrane (Alcorlo et al., 2017).

Gram-positive bacteria cell wall offers other types of structures interesting for phage adsorption. The thick PG layer forms the outermost surface of the cell and envelopes teichoic acid molecules that makeup to 50% of the total wall mass (Letarov and Kulikov, 2017). These teichoic acids (or lipoteichoic acids if bonded to the lipids of the plasma membrane) are often used as receptors for phages. They can mediate irreversible adsorption, but some phages use them only as intermediates (reversible adsorption) to reach their final receptor as in the case of *Bacillus subtilis* SPP1 phage and *Bacillus anthracis γ* phage that need attachment to the extra-membrane domain of YueB and the cell wall protein GamR, respectively, to trigger infection (Davison *et al.,* 2005*;* Baptista *et al.,* 2008*;* Nobrega *et al.,* 2018). As for Gram-negative bacteria, capsular and other surface polysaccharides as well as fimbriae and flagella, are also targeted by phages RBP (Nobrega *et al.,* 2018). Finally, the paracrystalline layer of protein subunits covering the cell surface, or S-layer, produced by both bacterial Gram types is also the target of several phages such as *B. anthracis* phage AP50 (Plaut *et al.,* 2014).

1.2.2. Receptor binding proteins (RBPs)

Their high genetic plasticity enables many tail-associated structures to take the role of RBP in the adsorption process. Even though the taxonomic classification shown in Fig. I.1. is getting obsolete, it will serve to illustrate the different protein structures such as tail fiber, spikes and baseplates associated with this function. RBPs are the primary determinants of the phage host range (Nobrega *et al.,* 2018).

Tail spike. One well-characterized RBP encoded in a tail spike is gene product 9 (gp9) of podophage P22 (Fig. I.9). This *Salmonella* spp. infecting bacteriophage uses its tail appendix to bind to the O-antigen and then, throughout endorhamnosidase enzymatic activity, exposes the outer membrane of its host (Andres *et al.,* 2010). Each RBP is a stable homotrimer of gp9 protein and the attachment to the bacteria is mediated by the variable carboxy (C)-terminal receptor binding domain whereas the conserved amino (N)-terminal domains attach to the tail. A short peptide, called linker, makes the bridge between the two external domains and, through a conformational change, ensures to signal the N-terminal domain when to trigger DNA ejection. In general, more than one spike binding is needed to trigger successful infection, which is subsequently mediated by a gp26 homotrimer forming a "needle" capable of injecting the genetic material by puncturing the cell. This strategy arose from the fact that, as other *Podoviridae* phages, P22 features no contractile tail or baseplate structure (Andres *et al.,* 2010; Nobrega *et al.,* 2018)

Unlike tail fibers, spikes usually have associated enzymatic activity. These tend to be specialized in PG degradation (cf. Ectolysins) but can also target other bacterial structures such as exopolysaccharides discussed below (cf. Depolymerases) (Nobrega *et al.,* 2018).

Tail fiber. Tail fibers tend to be longer and more complex than tail spikes, which often means that chaperones are needed for proper folding. The best-characterized tail fibers are those of *E. coli* phage, T4. Each of the six long tail fibers (LTF) consists of the assembly of four different proteins giving it the leg-like shape: homotrimer of gp34, monomer of gp35, trimer of gp36 and a trimer of gp37. The most distal part is thus composed by a gp37 trimer forming a needle-shaped "foot" containing the receptor-binding domain on its C-terminal end (Fig.I.9). RBPs contained in the LTF have been shown to bind reversibly to either LPS or OmpC. Additionally, the baseplate also features short tail fibers (STF) made of homotrimers of a single protein, gp12. During the recognition phase, LTF act before STF by binding to the bacterial surface and orientating the tail so that STF may be deployed correctly. In fact, at least three LTFs have to bind to the host in order for the signal to be transferred and to enable the conformational change of the baseplate. This process triggers the extension of STF, exposing their C-terminal end and enables the binding of the secondary RBP to the host's LPS (Yap and Rossman, 2014; Hyman and Van Raaij, 2018; Nobrega *et al.,* 2018).

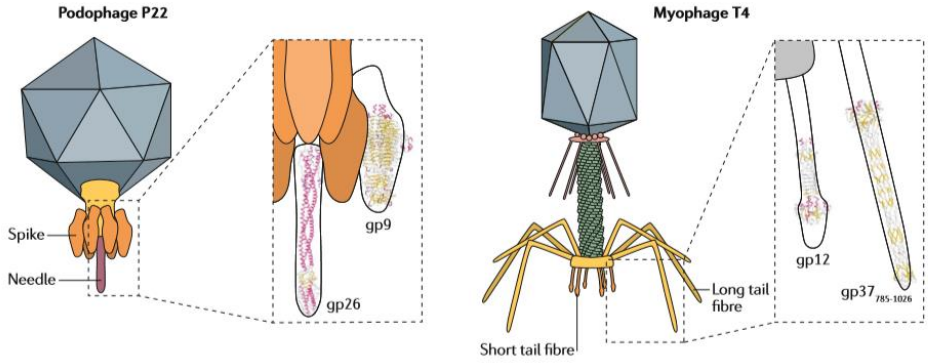

Figure I.9. Receptor binding proteins (RBPs) of tailed phages. Podophage P22 uses its tail spike (gp9) to bind the bacterial receptor and subsequently penetrates the bacterial cell with its needle-like tail tip (left). Myophage T4 has six long tail fibers encoding for RBPs at the most distal part of gp37 protein. Short tail fibers (gp12) also play a role in the adsorption process (Nobrega et al., 2018).

Baseplate. Some phages have their RBPs encoded into the most distal part of the phage tail, termed baseplate. For instance, most lactococcal phages have large baseplate structures that vary significantly in shape and size and form the control center for infectivity by encoding for the RBP. In the case of p2, a lytic phage, and TP901-1, a lysogenic phage infecting *Lactococcus lactis*, baseplates are "ready to adsorb" since no conformational change is needed in its structure in order to trigger infection (Spinelli *et al.,* 2014). Baseplate RBPs have also been identified in *L. monocytogenes* non-contractile phages (*e.g.* A118 and P35). In A118 phage, two proteins, gp19 and gp20, which form the lower and upper baseplate disk, respectively, contribute to host specificity, whereas in phage P35 only one, gp16, was identified. Moreover, the latter also represents the ectolysin of phages P35, thus showing that RBPs are not always located as peripheric structures and might be intimately associated with enzymatic activity (Bielmann *et al.,* 2015).

1.2.3. Depolymerases

As discussed previously, bacteria may exhibit at their surface many types of polysaccharides including capsular polysaccharides (CPS) attached to the cell surface and exopolysaccharides (EPS) unattached and secreted by cells in biofilm formations, in addition to lipopolysaccharides (LPS) in Gram-negative bacteria outer membrane. Many functions like biofilm production, virulence mediation and, of course, bacteriophage interaction are ensured by them. Indeed, capsule components and EPS can efficiently reduce the accessibility to receptors on the cell surface and thus significantly decrease phage adsorption. Therefore, some bacteriophages developed a way to keep their outstanding specificity while binding to these structures and drilling their way toward the cell surface. However, by doing so, most of the time they also became dependent on CPS and EPS for host recognition and infection since these structures became essential primary bacterial receptors. This is also why depolymerase C-termini, featuring the enzymatically active domain, tend to exhibit high specificity for their substrate. Unsurprisingly, the most studied depolymerases are podovirus-derived, given that they rely more on enzyme activity than in their short tails for adsorption (Knecht *et al.*, 2020).

In this context, depolymerases are capable of degrading bacterial polysaccharides, either associated with the cell surface (to facilitate phage adsorption) or involved in biofilm matrix (to facilitate phage diffusion). Normally, phages encode for one or two depolymerase domains in

the same gene, however, some are able to infect more than one bacterial type precisely because they encode several depolymerases in different genes (Pires *et al.,* 2016). As their name suggests, depolymerases cleave repeating units of polysaccharides and, depending on whether or not they release a water molecule while doing so, they are respectively classified as hydrolases or lyases. Hydrolases and lyases can be further divided into subcategories. Their activity hallmark consists of a turbid halo surrounding the clear plaque characteristic of phage infection on bacterial lawns (Knecht *et al.*, 2020).

Hydrolases. Hydrolases target the glycosidic linkage by catalyzing cleavage of the glycosyl-oxygen bond of mainly capsular polysaccharides or O-antigens on LPS (Drulis-Kawa *et al.,* 2015). In this category, up to six types of enzymes have been identified in phages: endosialidases, levanases, xylosidases, dextranases, rhamnosidases and peptidases (Pires *et al.,* 2016).

Endosialidases recognize and hydrolase the internal α-2,8-linkages in polysialic acid, which is a linear carbohydrate polymer composed of N-acetylneuraminic acid monomers commonly found in some neuropathology causing *E. coli* capsules (K1-antigen) (Drulis-Kawa *et al.,* 2015). Given that the K-1 antigen is also expressed in glycocalyx covering human epithelial cells, these pathogens often elude our immune system vigilance and frequently cause meningitis and septicemia in newborns (Knecht *et al.,* 2020). Several phages such as K1E were discovered to carry a hydrolytic tail spike protein with N-acetylneuraminidase activity (Tomlinson and Taylor, 1985; Knecht *et al.,* 2020). Levanases hydrolyze the β-2,6-linked main chain of levan, a polysaccharide known to contribute to *Bacillus* biofilm formation (Marvasi *et al.,* 2010). Therefore, some of the corresponding phages (*e.g. Bacillus* phage SP10) beneficiate from levan degradation activity (Pires *et al.,* 2016). Xylosidases, dextranases and rhamnosidases are more rarely identified in phages and catalyze hydrolysis of bacterial sugars such as xylan, dextran and rhamnogalacturonan, respectively (Pires *et al.,* 2016). For instance, some phages (*e.g. Shigella* phage Sf6) use rhamnosidases located on their tail spikes for recognition and depolymerization of repeating carbohydrate patterns of LPS O-antigens (Müller *et al.,* 2008; Drulis-Kawa *et al.,* 2015). Finally, some phage-encoded enzymes may also act on peptides (peptidase domain). In fact, some *B. subtilis* strains produce a capsular ɤ-linked polypeptide of glutamate (poly-ɤ-glutamate) that can be degraded by several *Bacillus* phages (*e.g.* phiAGATE, SPP1, phiNIT1) (Kimura and Itoh, 2003; Pires *et al.,* 2016).

Lyases. Lyases, in turn, utilize a β-elimination mechanism to cleave glycosidic β-1,4 bonds in polysaccharides. Three groups of these enzymes, namely hyaluronidases, alginate lyases and pectin lyases, have been identified in phages.

Hyaluronidases are able to digest hyaluronate, a linear unsulfated glycosaminoglucan polymer composing, on its own or dominantly, certain bacterial capsules including those of certain groups of streptococci, staphylococci and clostridia (Hynes and Walton, 2000). For these reasons, some phages like group A streptococcal phage phiNIH have integrated this enzyme in their structure to enable degradation of hyaluronate-containing host capsules (Pires *et al.,* 2016). Alginate is a linear polysaccharide of α-L-guluronate and β-D-mannonate abundant in nature and especially present as EPS in *Pseudomonas aeruginosa* biofilms which is responsible for severe infections and up to 100-fold greater resistance to antibiotics (Wong *et al.,* 2000; Knecht *et al.,* 2020). Podovirus PT-6 is one of the phages known to lyse the alginate capsule of this pathogen with structural alginate lyases (Glonti *et al.,* 2010). Last but not least, pectin lyases catalytically degrade galacturonic acid, the main component of pectin that makes up some bacterial polysaccharide structures (Hugouvieux-Cotte-Pattat *et al.,* 2010). *Pseudomonas putida* phages ϕ15 and AF both encode pectin lyases on their tail spikes, which explains their biofilm-

degrading properties (Cornelissen *et al.,* 2011; Pires *et al.,* 2016). A brief summary of hydrolase and lyase depolymerases can be found in Table I.1.

Table I.1. Mechanism and types of hydrolase and lyase depolymerases.

	Hydrolases	**Lyases**
Mechanisms	Cleave glycosyl-oxygen bond with water molecule	Cleave glycosidic β-1,4 bonds by β-elimination
Types	-Sialidases -Xylosidases -Peptidases	-Hyaluronidases -Alginate lyase -pectin lyase

1.2.4. Ectolysins

Although not indispensable, many phages harbor a variety of PG-degrading enzymes providing an advantage in the initial steps of infection. PG is composed of linear glycans strands made up by alternating N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) residues linked by β(1-4) bonds (Fig. I.10). These strands are cross-linked by short and variable peptides (Alcorlo *et al.,* 2017). More commonly known as ectolysins, these PG-degrading enzymes are also called tail-associated lysins (TALs), virion-associated peptidoglycan hydrolases (VAPGHs) or tail-associated muralytic enzymes (TAMEs). Moreover, they are cataloged depending on the type of bond they cleave in PG and thus be separated into glycosidases, amidases and peptidases (Fernandez and São-José, 2018). Additionally, two types of glycosidic bonds may be cleaved by glycosidases and so, these are subdivided into glucosaminidases which hydrolytically cleave the β(1-4) glycosidic bond between NAG and NAM, and lytic transglycosylases (LTs) together with muramidases which also cleave the β(1-4) glycosidic bond but between sequential NAM and NAG. The difference between the latter two being that muramidases are hydrolytic enzymes while LTs catalyze a non-hydrolytic cleavage (Fig. I.10) (Alcorlo *et al.,* 2017). Amidases target the amide bond linking the first amino acid of the peptide stem (most of the time L-Ala) and NAM. Finally, peptidases cleave bonds between amino acid residues and can be subdivided into endopeptidases (cleave internal bonds of the peptide) and carboxypeptidases (cleave the last C-ter amino acid residue) (Table I.2) (Fernandez and São-José, 2018).

Table I.2. Phage encoded ectolysins examples.

TAL type	Phage / Phage host	Protein name	Location	Reference
Glucosaminidase	T5/ *E. coli*	Pb2 (YP_006968.1)	Tail tip	Boulanger *et al.* (2008)
Muramidase	SP6/ *Salmonella typhimurium*	Gp36 (NP_853596.1)	Tail	Moak and Molineux (2004)
Lytic transglycosylase	T7/ *E. coli*	Gp16 (NP_042004.1)	Tail	Moak and Molineux (2004)
Amidase	Φ11 / *S. aureus*	Gp44 (NP_803297.1)	Baseplate (probably)	Li *et al.* (2016)
Peptidase	K / *Staphylococcus* sp.	ORF56 (AAO47505.1)	Tail	Pau *et al.* (2011)

Ectolysins main role is to facilitate cell wall crossing and tail extension without compromising host cell integrity and thus its viability for the rest of the cycle. Indeed, ectolysins degrade PG only locally, otherwise, the host cell could be lysed prematurely as a result of ectolysin-mediated PG degradation at multiple sites. This phenomenon is called “lysis from without” and is mainly observed following high-multiplicity virion adsorption (*e.g.* gp5 lysozyme in *E. coli* coliphage T4) (Abedon, 2011). Ectolysins have also been shown to confer a significant advantage when the infection takes place under conditions increasing PG cross-linking and thickening such as stationary growth or low temperatures (Piuri and Hatfull, 2006; Stockdale *et al.,* 2013)

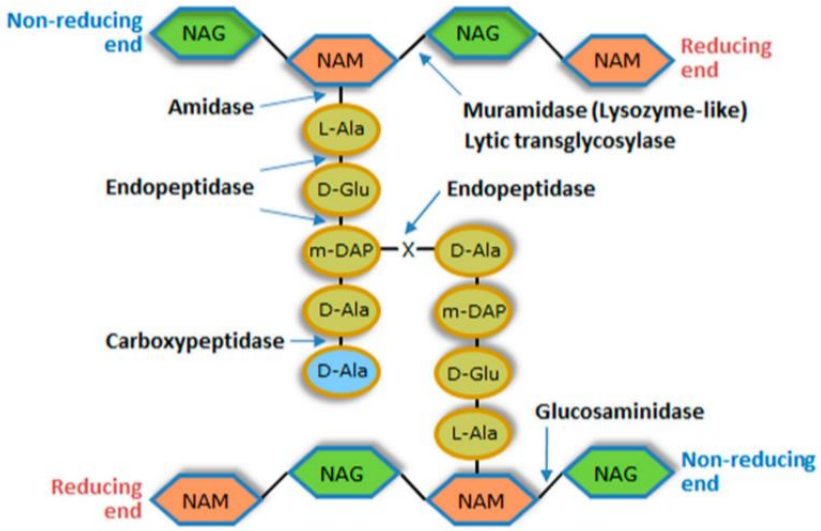

Figure I.10. Schematic representation of cell wall peptidoglycan found in bacteria and the location of enzymatic activities targeting it. In Gram-positive, the meso-diaminopimelic acid (m-DAP) peptide, commonly found in Gram-negative bacteria, is replaced by L-Lys. Crosslinking between one of the above and D-Ala of another chain is usually (symbolized by X) is usually composed of a variable amino acid sequence (Fernandez and São-José, 2018).

Concerning their localization, ectolysins appear to work from various part of the virion. For *Caudovirales*, the most common location is the tail and most precisely in central fibers, tail tips and even as part of the tail measure protein (TMP) which determine the length of the tail in phages (*e.g. E. coli* T5 siphovirus tail fiber acts as TMP and carry muralytic activity) (Boulanger *et al.,* 2008; Latka *et al., 2017)*. Other phage ectolysins, like the previously mentioned Gp5 lysozyme of T4, are located directly into the baseplate as part of a cell puncturing device (Kanamaru *et al.,* 2002). In the case of the tailless PRD1 phage (*Tectiviridae)*, ectolysins are associated with the viral membrane (Rydman and Bamford, 2002) and, in some cases, the ectolysin is located at the phage neck as-is for gp4 in phage P22 infecting *Salmonella* spp. (Tang *et al.,* 2011; Latka *et al., 2017)*.

Just like endolysins, ectolysins can be harnessed as antimicrobial agents due to their high specificity, thermostability, conserved mode of action on unique PG bonds and efficacy even when facing antibiotic resistance mechanism (Drulis-Kawa *et al.,* 2015). All of these make them attractive antimicrobials in the medical sector as well as in the industry. For instance, phage DW2 is not applicable as anti-staphylococcal therapy because of its temperate nature, however, the ectolysin it encodes present strong “enzybiotic” potential (Keary *et al.,* 2014). Ectolysins have also been used in food safety with PG hydrolase HydH5 of phiPLA88 phage used to prevent *S. aureus* growth in milk production (Rodríguez-Rubio *et al.,* 2013).

1.3. Deep Blue, the Herellevirus infecting Bacillus cereus

1.3.1. The *B. cereus* group

The *B. cereus* group, also called *B. cereus sensu lato (s.l.)* is a cluster of low G+C content Gram-positive, rod-shaped, spore-forming, aerobic bacteria that are closely related but exhibit highly diverse pathogenic properties (Liu *et al.,* 2015, Priest *et al.,* 2004). The main members include *B. anthracis, B. cereus sensu stricto (s.s.), Bacillus thuringiensis, Bacillus mycoides, Bacillus cytotoxicus, Bacillus pseudomycoides and Bacillus weihenstephanensis*. However, in the last years, other species such as *Bacillus toyonensis* have been identified and now the group comprises more than eleven suggested species (Liu *et al.,* 2015). Even though intensive work has been performed in deciphering their precise phylogenetic and taxonomic relationship, the group has been strikingly resistant to any satisfactory classification. Indeed, in the past, they were classified on the basis of purely phenotypic criteria but the emergence of molecular techniques (*i.e.* 16S RNA) revealed they were much more closely related than expected and thus blurred the line separating its members (Maughan and Van der Auwera, 2011). For instance, *B. anthracis*, *B. cereus* and *B. thuringiensis* 16S RNA gene sequences share greater than 99% identity (Ash *et al.,* 1991). This is consistent with the fact that distinguishing features between them are mostly encoded by genes located on plasmids (Van der Auwera *et al.,* 2007).

B. anthracis is the causative agent behind the fatal anthrax disease, one of the oldest known to infect man. It represents the only obligate pathogen within the genus *Bacillus*. (Baillie and Read, 2001). Its virulence is based on the presence of two virulence plasmids, namely pXO1 and pXO2. The first plasmid encodes for three toxic factors forming two bipartite exotoxins when associated (the edema and the lethal factors) while the second one encodes a poly-D-glutamic acid capsule conferring resistance against phagocytosis (Jensen *et al.,*2003; Bhatnagar and Batra, 2001).

B. cereus (s.s) is commonly isolated from soil or food and some strains are responsible for two types of gastrointestinal disorders causing emetic intoxications and diarrheal infections (Ehling-Schulz *et al.,* 2004). Both are generally mild and self-limiting even though lethal cases occurred (Mahler *et al.,* 1997; Dierick *et al.,* 2005). Food-borne illness due to this bacterium usually occurs upon spore survival to cooking or pasteurizing and subsequent germination and proliferation. In fact, their highly adhesive endospores are resistant to heat, dehydration and other physical stresses (Arnesen *et al.,* 2008). The emetic toxin, a small dodecadepsipeptide called cereulide, is synthesized by a 25-kb operon located on a large plasmid related to pXO1 (Hoton *et al.,* 2005; Ehling-Schulz *et al.,* 2006a) and is able to resist reheating, proteolysis and gastric acid (Ehling-Schulz *et al.,* 2004). It mostly induces vomiting, however, in serious cases, the toxin causes liver failure due to its action as a potassium ionophore that permeabilizes the mitochondrial membrane and hinders oxidative phosphorylation (Kawamura-Sato *et al.,* 2005). The diarrheal disease is caused by enterotoxins released in the small intestine by vegetative cells following ingestion of viable cells or spores. At least three cytotoxins are considered as potential causative agents of the diarrhoeal disease: haemolysin BL (Hbl), nonhaemolytic enterotoxin (Nhe) and cytotoxin K (CytK) (Arnesen *et al.,* 2008).

B. thuringiensis is an insect pathogen widely used in agriculture as a biopesticide due to its production of diverse crystal toxins, known as Cry and Cyt, synthesized during its sporulation phase of growth. These toxins, also plasmid-encoded, are recognized pore-forming molecules that kill insect larvae epithelium midgut cells by inducing an osmotic shock and thus leading to cell lysis (Bravo *et al.,* 2013). Its success became apparent with the insertion of *cry* genes into species of agricultural interest to create crops that resist insect attack (Bt-crops such as Bt-Maize and Bt-cotton) (Sanahuja *et al.,* 2011).

The other species are discriminated by physiological or morphological traits. *B. weihenstephanensis* strains are psychrotolerant (Lechner *et al.,* 1998), in contrast with *B. cytotoxicus* strains that are thermotolerant and can grow in up to 53°C (Guinebretière *et al.,* 2013). *B. mycoides and B. pseudomycoides* have typical rhizoid growth (Di Franco *et al.,* 2002;) and *B. toyonensis* is known for its probiotic properties and is harnessed in animal feed (Jiménez *et al.,* 2013).

1.3.2. Phages preying on the *B. cereus* group

All identified phages know to prey on *B. cereus* group contain dsDNA genomes and are either classified in the *Caudoviridae* order or *Tectiviridae* family, which includes all icosahedral phages that use an internal lipid vesicle acting as a tail-like structure upon adsorption for genome delivery (Ackermann *et al.,* 2006; Gillis and Mahillon, 2014). It was also noted that, in this context, phages identified as virulent are part of the *Myoviridae* family (several members now reclassified in *Herelleviridae* family) while *Siphoviridae* and *Tectiviridae* family phages are temperate. This supports the hypothesis of a close relationship between life cycle and the particular morphology of phage families infecting this group (Lee *et al.,* 2014). Other categories, proposed by Gillis and Mahillon classify them according to morphotypes, lifestyles and/or lysogenic states such as transducing phages, phages with a chromosomal or plasmidial prophage state, γ-like phages and jumbo-phages (Gillis and Mahillon, 2014).

1.3.3. Phage Deep Blue characterization

Deep-Blue phage is a recently reclassified Herellevirus infecting a few strains of the *B. cereus group* and was isolated from agricultural soil collected in the Belgian countryside (Hock *et al.,* 2019). Indeed, five out of twelve *B. weihenstephanensis* strains (three emetic and two non-emetic strains), two out of five *B. thuringiensis* strains and two out of ten *B. cereus s.s.* strains tested were sensitive to the Deep-Blue phage (Hock *et al.,* 2019). As visible in Fig. I.11., Deep-Blue has an icosahedral head of 48 nm in diameter attached to a 129 nm long contractile tail. Moreover, this virulent phage has a 158,501 base pairs (bp) genome with a G+C content of 39.95% and a coding density of 90%. Its genome contains 226 potential coding sequences CDSs of which 148 have no predicted function (Hock *et al.,* 2016).

B. weihenstephanensis BtB2-4 strain has proven to be its preferred host and, for each lytic cycle, about 23 min are required to reach 90% of adsorption. A latent period of 50 min is then observed between phage adsorption and newly synthesized virion release with an average of 300 virions released per infected cell under the tested conditions (Hock *et al.,* 2016).

Hock *et al.* assessed Deep-Blue biocontrol potential since it fulfilled certain important criteria required for phage application in the food industry. First of all, Deep-Blue is strictly lytic and its genome does not contain bacterial pathogenicity genes and lysogenic determinants risking to enhance bacterial virulence. Moreover, its host spectrum is limited to the *B.* cereus group, which guarantees that it will not interfere with food microflora or gut microbiota. With this in mind, it will have to be included in a phage cocktail in order to obtain complete elimination of all targeted bacteria. Finally, Deep-Blue was proven to be stable during 1h at 40°C and at pH from 5 to 1, thus allowing appropriate biocontrol throughout several food processing and storage steps (Hock *et al.,* 2019). Consequently, Deep-Blue biocontrol potential was evaluated in LB medium. Since psychotropic strains like *B. weihenstephanensis* are regularly responsible for spoilage in the dairy industry and *B. cereus s.s.* for rice and pasta related food intoxications, it was also tested in milk, milk rice and rice (Owusu-Kwarten *et al.*, 2017; Hock *et al.,* 2019). In general, Deep-Blue was able to reduce bacterial growth and its effectiveness was increased with higher phage concentration. Concerning its application to control bacterial biofilm, Deep-Blue showed to be more effective in preventing biofilm development and not so profitable for treating well-established biofilm formations (Hock *et al.,* 2019).

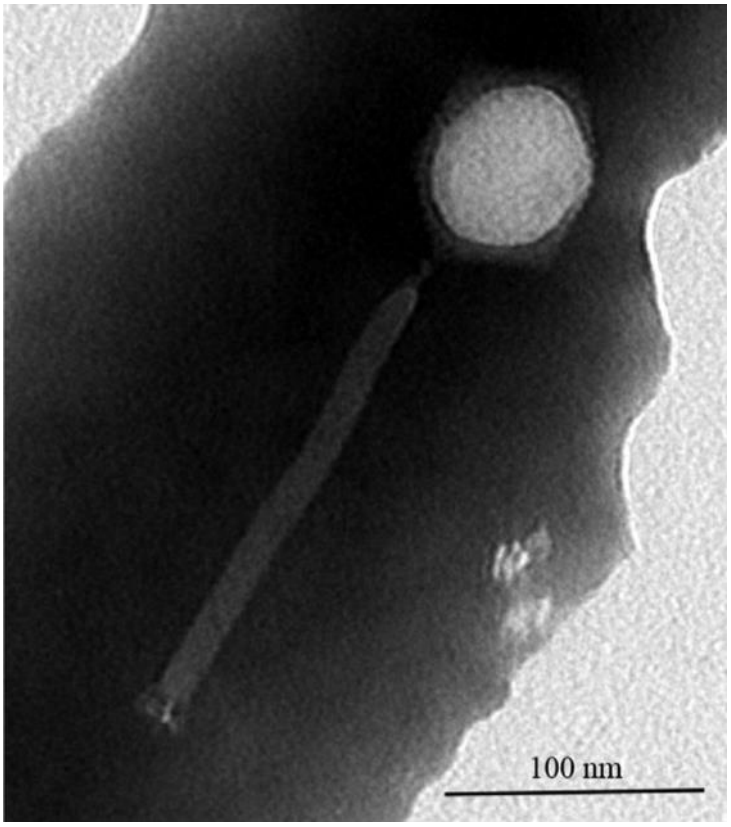

Figure I.11. Phage Deep-Blue picture taken by transmission electron microscopy and negatively stained with 2% potassium phosphotungstate (Hock et al., *2019)*

Objectives

Phages are viruses preying on bacteria, using them to replicate and propagate. They have been known for many years and have had an impact on a variety of fields including molecular biology and medicine, even though antibiotics discovery overshadowed them for a time. Phages are specially harnessed due to their outstanding specificity, which is determined during the adsorption process, the first step of phage infection. This step is completed after the phage Receptor Binding Protein (RBP) specifically and irreversibly binds to the bacterial cell surface receptors. This triggers viral genome injection and starts a lytic or lysogenic cycle. However, the presence of many types of polysaccharides can make reaching the receptors a challenging task. Therefore, some phages carry depolymerases, enzymes capable of degrading bacterial polysaccharides and allowing them to drill their way toward the cell surface. Additionally, many phages harbor peptidoglycan (PG)-degrading enzymes, also called ectolysins or tail-associated lysins (TALs), that facilitate cell wall crossing and tail extension locally.

High specificity, thermostability and efficacy of both enzyme types make them attractive antimicrobial agents in the medical sector as well as in the industry. Since they are mostly used for food safety in the dairy industry, most research has focused on Gram-negative hosts such as lactic acid bacteria thus creating a void in Gram-positive host research.

This work aims to identify the adsorption facilitating enzymes (*i.e.* depolymerases and ectolysins) helping the phage Deep-Blue in the initial steps of the infection process. Deep-Blue phage is a member of the *Herelleviridae* family infecting a few strains of the *B. cereus group* and potentially interesting for biocontrol of emetic *B. cereus* in food products.

The first step towards this goal will be to perform a bioinformatic analysis of Deep-Blue tail proteins and others to select candidate ectolysins and depolymerases. The candidate proteins will then be cloned and expressed in *E. coli*. After induction, soluble proteins will be purified by affinity chromatography in native conditions. Finally, tests will be performed to assess the lytic activity of each candidate on several strains of the *B. cereus group* by turbidity reduction assays (ectolysins) or by spotting them on bacterial layers on agar plates.

Chapter II

Material and methods

2.1. Strains and plasmids

All bacterial strains used in in this work are listed on Table II.1.

Concerning cloning experiments, the plasmid pET30 was chosen as vector. It contains a T7 promotor recognised by *E. coli* T7 polymerase. Regulation is ensured by the Lac operon in which LacI (lac repressor) binds to LacO (lac operator) to inhibit transcription. This inhibition can be removed by addition of IPTG. To facilitate clone selection after transformation a Kanamycin resistant gene (KanR) is also present. Finally, to allow recombinant protein purification, pET30 harbours a 6-His Tag separated from the rest of the protein by an enterokinase cleavage site. The genetic map of this vector is represented in Fig.II.1.

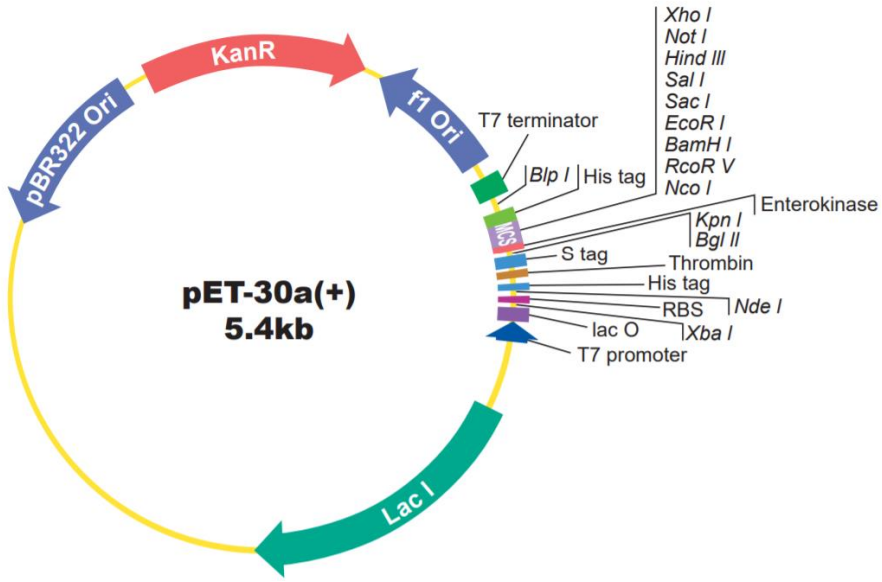

Figure II.1. Genetic map of pET30 expression vector. KanR - Kanamycin resistance ; pBR322 Ori origin of replication (originally presents on pBR322 vector) ; LacI - operon lactose inhibitor ; T7 promoter - promoter for phage T7 RNA polymerase ; LacO - operon lactose operator ; RBS - ribosome binding site ; 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 (Genscript).

Table II.1. Bacterial strains used in this work.

Bacterial strain	Relevant feature	Origin
E. coli 10β	Competent cells	NEB
E. coli BL21(DE3)	Competent cells	Novagen
E. coli BL21(DE3)[pET30::gp189]	Containing the recombinant plasmid pET30::gp189	This work
E. coli BL21(DE3)[pET30::gp189_lytic]	Containing the recombinant plasmid pET30::gp189_lytic	This work
E. coli BL21(DE3)[pET30::gp182]	Containing the recombinant plasmid pET30::gp182	This work
E. coli BL21(DE3)[pET30::gp191_1000]	Containing the recombinant plasmid pET30::gp191_1000	This work
E. coli BL21(DE3)[pET30::gp195]	Containing the recombinant plasmid pET30::gp195	This work
E. coli BL21(DE3)[pET30::gp222]	Containing the recombinant plasmid pET30::gp222	This work
E. coli BL21(DE3)[pET30::gp223]	Containing the recombinant plasmid pET30::gp223	This work
Si0239	*B. weihenstephanensis*	MIAE
ATCC10987	*B. cereus*	MIAE
WS10204	*B. weihenstephanensis*	Lechner S.
KBAB4	*B. weihenstephanensis*	MIAE
MC67-2	*B. weihenstephanensis*	Thorsen
MC118-4	*B. weihenstephanensis*	Thorsen
Aw43	*B. thuringiensis*	MIAE
Sio170	*B. weihenstephanensis*	MIAE
H3081	*B. cereus*	Hoffmaster
GBJ001	*B. thuringiensis*	MIAE
NVH391-98	*B. cytotoxicus*	INRA
BtB2-4	*B. weihenstephanensis*	MIAE

2.2. Culture media, solutions and bacterial growth conditions

2.2.1. Media and antibiotic preparation

Bacterial cultures were grown in Luria-Bertani (LB) medium (Oxoid) except after the heat shock step of the transformation process in which we used the Super Optimal broth with Catabolite repression (SOC) medium. For phage experiments, LB medium supplemented in $MgSO_4$ (VWR Chemicals, 3 mM) and $CaCl_2$ (VWR Chemicals, 3 mM) (LB*) was used. Top agar was used in phage experiments and obtained by supplemented LB medium with 6% agar (Sigma Aldrich) and ions required for phage lytic activity (3 mM MgSO4 and 3 mM CaCl2). Composition of SOC medium, used in the transformation procedures, can be found in Table II.2. *Bacillus* spp. were grown et 30°C whereas *E. coli* at 37°C, with agitation of 120 rpm and 180 rpm, respectively when grown in liquid cultures.

Table II.2. SOC medium composition

Component	Concentration (g/L)
Yeast extract	5
Tryptone	20
NaCl	0.584
$MgSO_4$	2.467
KCl	0.186
$MgCl_2$	0.952
Glucose	3.603
Distilled water	Fill to 1L

Kanamycin (Sigma Aldrich) stock solutions were prepared and sterilized by filtration through 0.22 µm filters to be stored at -20◦C and used in LB medium at a final concentration of 50 µg/mL.

2.2.2. Saline solution

Phosphate Buffered Saline (PBS) solutions were made with Sigma Aldrich PBS tablets (1 tablet for 200 mL of distilled water) and PBS-NaCl solution was prepared by adding NaCl to the concentration of 300 mM.

SM buffer was used in phage experiments and obtained as shown in Table II.3.

Table II.3. SM buffer composition.

Component	Concentration (mM)	Quantity (g or mL)
NaCl	100	5.8
$MgSO_4.7H_2O$	30	7.4
Tris-HCL (1M pH 7.5)	50	50
Distilled water		Fill to 1L

2.2.3. Preparation of competent cells

Competent cells of *E. coli* conserved at -80°C were streak on Luria-Bertani (LB) plate and incubated O/N at 37°C. One single colony was selected to inoculate 10 mL of LB medium and grown O/N at 37°C. The next day, 1L of LB was inoculated with 10 mL of the O/N culture and grown at 37°C with shaking. Continuous measurement of the optical density at 600 nm (OD_{600}) was performed until the culture reached 0.35-0.4 after what it was put on ice. After 20 to 30 min of cooling, the culture was divided into 4 ice-cold 250 mL centrifuge bottles. A first centrifugation was performed at 3 000 g for 15 min at 4◦C to harvest cells. After discarding the supernatant, pellets were resuspended in a total of 400 mL of ice-cold $MgCl_2$ (100 mM). Then, a second spin at 2 000 g for 15 min at 4 °C was performed. The pellet was resuspended in 200 mL of ice-cold $CaCl_2$ (100 mM) and kept on ice at least 20 min before performing a third centrifugation at 2 000 g for 15min at 44 °C. Afterward, cells were resuspended in 50 mL of ice-cold 85 mM $CaCl_2$ supplemented with glycerol (15%). The suspension was then transferred into a cold Falcon tube and centrifuged at 1 000 g for 15 min at 4◦C. The pellet was resuspended in 2 mL of ice-cold 85 mM $CaCl_2$, 15% glycerol, aliquoted in a volume of 50 µl and stored at -80°C.

2.3. Phage multiplication and purification

2.3.1. Multiplication and polyethylene glycol precipitation

Deep-Blue host strain Si0239 was streaked on an LB plate and incubated at 30°C O/N. The following day, 50 mL of LB* were inoculated with this bacteria and incubated at 30°C, 120 rpm for 4 h. Then, the bacterial culture was infected with 500 µL of phage suspension at 10^9 PFU/mL and incubated O/N at 30°C, 120 rpm. Following the O/N incubation, RNase and DNase were added at 1 µg/mL each as well as chloroform (0.2 % v/v) in order to break down nucleic acids and lyse the remaining bacteria, respectively. 0.5 M of NaCl was dissolved in the phage suspension which was incubated for 1 h at 4°C. A centrifugation step at 6000 g, 4°C for 10 min was then needed to precipitate the bacterial debris. The supernatant was transferred into a clean tube to which PEG 6000 or 8000 (10% w/v) was added. The suspension was inverted until the PEG was dissolved and was then incubated at 4°C for at least 1 h. A second centrifugation step (10 000 g, 4°C, 15 min) was necessary to precipitate the phage particles. After discarding the supernatant, the remaining liquid was allowed to flow out by turning over the tubes for 5 min after what each pellet was resuspended in 2 mL of SM buffer. After a last centrifugation at 5000 g, 4°C for 10 min.

2.3.2. Caesium chloride purification

A gradient step centrifugation approach was first attempted. This technique allows to separate phages from bacterial debris based on their differential density. For this matter, a density gradient is created by layering CsCl solutions of increasing density. Once the phage extract is layered on top of this gradient, an ultracentrifugation step separates the two phases into clear bands distant enough to recover the phage-containing one (Protocol 1). However, this technique did not retrieve concentrated enough phage solutions (input of 10^{10} PFU/mL and recovery of 10^4 PFU/mL). The protocol was thus adapted for smaller volumes by using a unique density using an

equilibrium centrifugation approach. For this purpose, 3 g of CsCl were directly added into 4 mL (0,75 g of solid CsCl per ml) of our phage suspension to reach a density of 1,5 g/mL. The solid CsCl was added gradually to prevent inactivation of bacteriophages by osmotic shock. This mix was then transferred into a tube (Ultra-Clear™ centrifuge tubes 13x51 mm -Beckman Coulter®) suited for ultracentrifugation at 35 000 rpm for 24 h at 4°C in an MLS 50 swinging-bucket rotor. The purified phage particles form a band equilibrated at a position corresponding to their proper buoyant density that can be easily collected by inserting a syringe needle (21Gx1'' Terumo® Agani needle) (Protocol 2).

To remove the remaining CsCl and concentrate the recovered phages O/N dialysis tubes in SM buffer and ultrafiltration techniques in Amicon® Ultra-4 Centrifugal filter unit 3000 NMWL (Merck) were tested.

2.4. Phage adsorption test

Bacterial strains to be tested were streaked on LB and incubated at 30°C overnight (O/N). The following day tubes containing 5 mL of LB were inoculated with a single colony from the plated strains and incubated at 30°C, 120 rpm for 3 to 4 h. In a 2 mL Eppendorf tube, 600 µL of LB* were mixed with 600 µL of the bacterial cultures (except for the control tube containing 600 µl of LB* instead of bacteria = phage input). 100 µL of Deep-Blue suspension at 10^7 PFU/mL were added to each tube including the control one and all tubes were then incubated at 30°C for 25 min. Tubes were centrifuged at 10,000 g, 4°C for 10 min. The supernatants, containing unabsorbed phages, were filtered through a 0,45 µm filter and serial dilutions were prepared to assess their concentration in a double layer agars (DLA) assay (Fig. II.2.). In brief, 100 µL of Si0239 (Deep-Blue sensitive strain) and 100 µL of the appropriately diluted supernatant were added to a top agar tube, which was then poured on an LB-agar plate and incubated at 30°C O/N. The following day, phage concentration (Plaque Forming Unit/mL or PFU/mL) was evaluated and the adsorption rates deducted with the following formula:

$$Adsorption\ rate\ [\%] = \frac{[phage\ input] - [phage\ in\ tested\ strain]}{[phage\ input]} \times 100$$

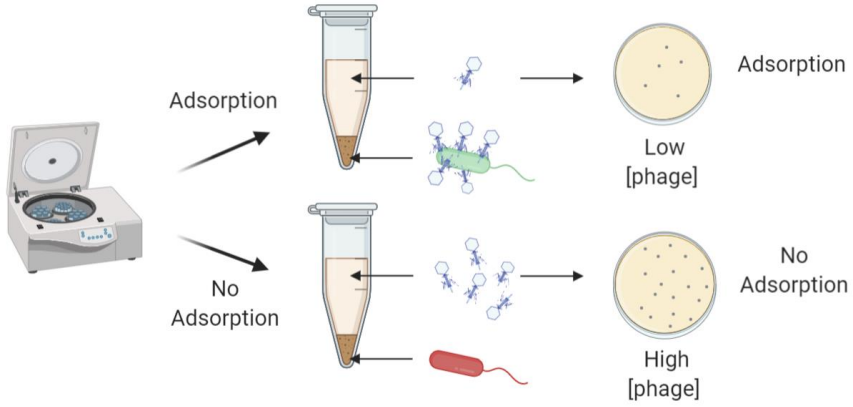

Figure II.2. .Adsorption test principle. If Deep-Blue can adsorb to the strain tested, then the pellet will contain the majority of phages adsorbed to the bacteria and the supernatant will consequently be depleted of viruses. In the opposite scenario, the supernatant will be very concentrated in phages that failed to adsorb to a certain strain.

2.5. Phage concentration estimation

Tubes containing 5 mL of Top agar preparation were inoculated with 200 µL of bacterial culture, spread on an LB plate and left to dry for 15 min. 10-fold serial dilutions of the phage suspension in sterile water were prepared. Once the plate is completely dry, 10 µL droplets of each dilution were spotted on the plates which were incubated O/N at 30°C. The phage concentration (PFU/mL) was estimated by counting the individual phage plaques.

2.6. Molecular methods

2.6.1. Polymerase chain reaction (PCR)

Amplification of candidate genes was done by PCR with the primers obtained from Macrogen and listed on Table II.4.

Routine PCRs were done with the OneTaq® polymerase (NEB) whereas the high-fidelity Q5® polymerase (NEB) was used for DNA amplification of our candidate genes. PCR programs used are reported in Table II.5.

Table II.5. PCR programs used with Q5® and OneTaq® polymerases. Tm was calculated with NEB Tm Calculator.

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

Agarose gel electrophoresis

Agarose gels were prepared by mixing agarose (0,8%, ChemCruz™), 0.5x TAE buffer (Table II.6.) and ethidium bromide (EtBr). Electrophoresis was run for 30 min at 100 V and the DNA bands were revealed under the UV light thanks to the Quantity One R software (Bio-Rad). Unlike the OneTaq® mix, the Q5® polymerase buffer does not contain any tracking dye so 2 µl of Purple loading dye (NEB) was added to 5 µl of PCR product when using this enzyme.

Table II.6. Composition of TAE Buffer stock solution (50x).

	Origin	Concentration	Quantity
Tris-(hydroxymethyl)-aminomethane	Merck	2 M	242 g
$NaOAc.3H_2O$	Merck	1 M	57 mL
EDTA (0.5 M; pH 8)	Sigma-Aldrich	0.05 M	100 mL
Distilled water	-	-	Fill to 1L

Table II.4. Primers used in this study. Direction and restriction enzymes are included in primer names. F = forward primer. R = Reverse primer. Restriction sites are highlighted in bold.

Primer	Primer sequence (5' →3')	DNA target	Construction	Product size (bp)
gp189_BamHI_F	TAT**GGATCC**ACAACGATTGTTAAACGCTATCC	gp189	pET30::gp189	**2251**
gp189_SalI_R	TCT**GTCGAC**TTATCCTGTAAACCTTCTTACGTG	gp189	pET30::gp189	**2251**
gp189_lytic_NcoI_F	TAT**CCATGG**TAGGAGTTGCAATGCAAGCTTT	gp189	pET30::gp189_lyt	**450**
gp191_1000_NcoI_F	TAT**CCATGG**TTAACACGAAACAACGTACTGAAG	gp191_100	pET30::gp191_1000	**863**
gp191_XhoI_R	AAAT**CTCGAG**TTAAGTCATCACCATCTCTTTAGAGTAGAA	gp191_100	pET30::gp191_1000	**863**
gp182_NcoI_F	TAT**CCATGG**TAGCAACTGAAACGATCCAATC	gp182	pET30::gp182	**483**
gp182_XhoI_R	TATA**CTCGAG**TTAAAGGACTCTATGTAGTGAGTATT	gp182	pET30::gp182	**483**
gp195_EagI_F	AT**CGGCCG**AAATTAAAAGCATTAATCACAACAGG	gp195	pET30::gp195	**682**
gp195_NcoI_F	TATC**CATGG**TGAAATTAAAAGCATTAATCACAACAGG	gp195	pET30::gp195	**683**
gp195_XhoI_R	ATAT**CTCGAG**TTACTTACCTAATATCGTAACTTTCAC	gp195	pET30::gp195	**682 or 683**
gp223_NcoI_F	ATA**CCATGG**TTTCGAGAGAGTACTTCGTAGAG	gp223	pET30::gp223	**2391**
gp223_XhoI_R	TTAT**CTCGAG**TTAAACCCATTCCGATCCGTTC	gp223	pET30::gp223	**2391**
gp222_KpnI_F	TT**GGTACC**GCAGAACAAAAGCGTATGA	gp222	pET30::gp222	**483**
gp222_BamHI_R	TCT**GGATCC**TCAAGTTACAACTGTTCCTGC	gp222	pET30::gp222	**483**
Seq_pET30_F	ATACCCACGCCGAAACAA	MCS	-	**Insert + 359**
Seq_pET30_R	CCGGATATAGTTCCTCCTTTCAG	MCS	-	**Insert + 359**

Plasmid extraction and DNA purification

The GenElute™ Plasmid Miniprep kit (Sigma-Aldrich) was used to extract plasmids from *E. coli* cells and the GenElute™ PCR Clean-up kit (Sigma-Aldrich) was used for all DNA purification after PCR and restriction.

Restriction- Ligation

Restrictions were achieved using restriction enzymes (NEB) listed in Table II.8 and using the standard restriction mix showed in Table II.7. Tubes were incubated for 6 h at 37°C after what DNA was purified and the concentration was measured with a Nanodrop ND-1000 spectrophotometer (Isogen Life Sciences).

Table II.7. Standart restriction mix composition

Component	Quantity
DNA	~ 1 µg
CutSmart® Buffer 10x	5 µL
Restriction enzyme 1	1 µL
Restriction enzyme 2	1 µL
Distilled water	Fill to 5 µL

Table II.8. Restriction enzymes (NEB) used in this work.

Enzyme	Restriction site
***Nco*I-HF**	CCATGG
***Eag*I-HF**	CGGCCG
***Eco*RI-HF**	GAATTC
***Kpn*I-HF**	GGTACC
***Xho*I-HF**	CTCGAG
***Bam*HI**	GGATCC
***Sal*I**	GTCGAC

Ligations were performed using the T4 DNA ligase (Promega). NEBioCalculator™ was used to determine the amounts of vector and insert necessary for the ligation mix in order to obtain a vector:insert molar ratio of 1:7. The reaction tubes were then incubated O/N at 4°C. Table II.9. shows the typical composition of a ligation mix.

Table II.9. Standard ligation mix composition.

Component	Quantity
T4 DNA Ligase	1 µl
T4 DNA ligase Buffer 10x	1 µl
DNA (vector + insert)	8µl

Transformation

E. coli transformation was performed by adding 2 µL of ligation product into 50 µL of competent 10β cells. Next, the cells were left 30 min on ice after which a thermal shock at 42°C for 30 sec was conducted. Tubes were then transferred 5 min into ice before adding 950 µL of SOC

medium. Tubes were incubated horizontally at 37°C and 180 rpm for 1 h. In order to select our recombinant clones, cell cultures were spread on Kanamycin-containing agar plates. The following day, a PCR screening was performed to identify the positives clones having integrated the recombinant plasmid.

Sequencing

Plasmids were extracted from "positive" clones and sent to Macrogen Europe B.V for sequencing. For this matter, 7.5 µL of the extracted plasmid was mixed with 2.5 µL of the appropriate primers.

2.7. Recombinant protein expression

2.7.1. Protein induction

Plasmids containing the candidate genes were transformed in the *E. coli* expression strain BL21(DE3). 20 mL of Kanamycin-containing LB was inoculated with a single colony and was incubated O/N at 37°C. The following day, 100 mL of Kanamycin-containing LB was inoculated with 10 mL of the O/N culture and then this new culture was grown at 37◦C with shaking (180 rpm) until the OD_{600} reached 0.5-0.8. Prior to the induction, a 1 mL sample was taken and kept at 4°C for the solubility assessment as non-induced control. To trigger protein production, IPTG (Sigma) was added. Parameters such as IPTG concentration (0.5 or 1 mM), temperature (28, 22, 18°C) and time (3 to O/N) were optimized and will be discussed in the Results section. After the incubation time, two samples of 1 mL were collected for the solubility assessment test as induced control. The remaining culture was poured into 50 mL Falcon tubes. The aliquots were finally collected by centrifugation at 4 000 g, 4°C for 15 min and, after discarding the supernatant, the pellets were stored at -20°C.

2.7.2. Solubility assessment

The 1 mL samples (one non induced and two induced) were thawed on ice for 30 min. Pellets were then resuspended in 80 µL of Tris-10 buffer (Table II.11). Next, 1 mg/mL of lysozyme and 1x protease inhibitor cocktail (Sigma) were added and samples were incubated at 10 °C for 45 min. 3 µL of Benzonase® nuclease (Sigma-Aldrich) were then added to the mixture and incubation was resumed for 15 min. Soluble and insoluble proteins were separated by centrifugation (10 000 g, 4°C, 30 min) in one of the induced control. The supernatant was recovered as soluble fraction and together with the non-induced and induced controls, were used to perform an SDS-PAGE analysis.

2.7.3. Protein gel electrophoresis

The three samples were loaded into a Mini-PROTEAN® Precast protein gel (Bio-Rad) according to the following protocol:

1. Add DTT (Sigma) to Laemmli buffer (Bio-Rad) to obtain a final concentration of 100 mM.

2. Add 20 µl of DTT-Laemmli buffer to 20 µl 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) with distilled water.

4. Carefully remove 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 and make sure it is well positioned.

6. Fill the inner chamber with the buffer up to the top of the wells.

7. Load 20 µl of samples and 2 µl of Ladder (Color Pre-stained Protein Standard, Broad Range (10-250 kDa), NEB) at one end.

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.

13. Mix the Coomassie blue R-250 (800 mg/ml, Merck) with coloration solution (Table II.10) and pour it in the square petri dish and let the gel color for 1 to 3 days.

14. Discolor with MilliQ® water.

Table. II.10. Coloration solution composition. Components should be added in the stated order.

Component	Quantity	Origin
$(NH_4)_2SO_4$	85 g	VWR Chemicals
Distilled water	315 mL	-
Methanol	170 mL	Merck
H_3PO_3	15 mL	Riedel-de Haën

2.7.4. Protein purification

Soluble proteins were purified by affinity chromatography on a nickel column (Ni-NTA Spin Kit from Qiagen), thanks to the six-histidine tag at the N-terminal end of the protein. The clear lysate containing the soluble protein was filtrated through 0.45 µm filters and loaded to a column previously equilibrated with 650 µl Tris-10 (2000 g, 2 min with an open lid). After discarding the flow through, 650 µL of clear lysate were transferred into the column and centrifuged at 2000 rpm for 5 min with closed lid. Clear lysate was loaded twice and the centrifugation step repeated. After discarding the flow-through, the column was washed 2 times, first with 650 µL of Tris-20 and then with 650 µL of Tris-40 interspaced by centrifugations at 2000 rpm for 2 min with closed lid. After discarding the flow through and transferring the column into a clean Eppendorf tube, the proteins were finally eluted twice with 300 µL of Tris-500 buffer by centrifugation at 2000 rpm for 2 min. The composition of all cited buffers can be found in Table II.11. Protein concentration was assessed via a Bradford assay.

Table II.11. Tris buffers composition. Sterilized by filtration on 0.22 µm filters.

Component	Quantity (mg)				Origin
	Tris-10	Tris-20	Tris-40	Tris-5000	
Tris Base	303	303	303	303	Merck
KCl	746	746	746	746	VWR Chemicals
Imidazole	68	136	272	3.4 g	Sigma-Aldrich
NaOH/ HCl	Adjust to pH 8				VWR Chemicals
Distilled water	Fill to 100 mL				

2.7.5. Protein concentration

This step was performed to concentrate proteins before testing them and to change the buffer. Purified proteins were placed in an Amicon® Ultra-4 Centrifugal filter unit 3000 NMWL (Merck) and the tube was filled up to 4 mL with PBS supplemented with 300 mM of NaCl. Tubes were then centrifuged at 7500 g, 4°C for 30min. Flow-through was discarded and volume was increased again to 4 mL. The process was repeated three times until only about 200 µL of highly concentrated protein remained in the tube.

2.8. Cell wall decoration assay

This assay was performed to investigate if Deep Blue RBP (gp185) could bind to the several strains tested. For this purpose, gp185 was coupled to GFP in a previous work in order to highlight the adsorption of the protein to the bacteria. In this assay, 300 µL of a 4-h culture was centrifuged at 13 000 g for 5 min. Next, the pellet was washed with 500 µL SM buffer and the pellet resuspended in 100 µL of SM buffer. Ten µg of gp185-GFP were added to the suspension and incubated 25 min at RT. Thereafter, samples were washed again twice in 200 µL of cold SM buffer to remove the unbound proteins. Finally, samples were resuspended in 50 µL of SM buffer of which 10 µL were taken to be observed under a Leica Microsystems fluorescent microscope, using the LAS AFR® software.

2.9. Enzyme activity testing

2.9.1. Turbidity reduction assay: Ectolysins

Activity tests in liquid conditions were mostly used to test ectolysins candidates. Targeted strains 4-h cultures were first washed in 2 mL of TRIS-HCl 25mM after centrifugation at 10000 g for 5min. After a second centrifugation step cell were resuspended in 1.5 mL of TRIS-HCl 25mM.To start the test with all cultures at the same OD of 0.8-1.0, a first OD_{600} was measured for 200 µL and adjusted cultures to the desired OD using the following formula:

$$Volume\ of\ TRIS - HCl\ to\ add = \frac{Observed\ OD}{Desired\ OD} \times (1.5\text{-}0.2) - (1.5\text{-}0.2)$$

The assay was performed in a standard 96-well titration plate (CellStar® Greiner Bio-one) in which 40 µL of purified protein or TRIS-HCl buffer (control well) were placed into 160 µL of bacterial suspension. The OD_{600} was measured every min during 1 h at 37°C in a MultiskanTM FC Microplate spectrophotometer (Thermofisher scientific).

2.9.2. On Top agar: Depolymerases

To assess the depolymerase activity, 10 µL- droplets of purified protein or TRIS-HCl 25mM (control) were spotted on previously contaminated Top agar plates (200 µL of 4-h bacterial culture in 5 mL Top agar tubes). The presence of a halo around the droplet suggests a depolymerase activity.

2.10. Bioinformatics

Basic local alignment search tool (BLAST) algorithm was used for amino acid sequence analysis and pairwise sequence alignment, more precisely BLASTp was employed to search protein databases using a protein query. Conserved domain searches were conducted using the NCBI Conserved Domain Database and Interpro. HHpred was used for analysis based on structural homology rather than sequence homology. Finally, Phyre[2] , a protein structure modeling and function prediction tool was also used.

Chapter III

Results and discussion

3.1. Bioinformatic analysis of Deep-Blue genome

Since the goal of this work is to identify the adsorption facilitating enzymes of Deep-Blue, a bioinformatic analysis was performed to select potential candidate proteins. Deep-Blue genome contains 226 CDSs of which 148 have no predicted function (Hock *et al.*, 2016). Since previously identified ectolysins and depolymerases were mostly of structural nature, we decided to first focus on the putative structural components of Deep-Blue's tail.

3.1.1. Tail protein region

In genomes of dsDNA phages, the structural proteins are encoded in the late gene region in successive order of the virion assembly process. We were able to delimit the region encoding the tail proteins between the TMP and the helicases, which encodes 16 CDS (Fig. III.1) (Haban *et al.*, 2014). A brief overview of the different tail proteins of Deep-Blue is given here, since a detailed bioinformatic analysis of these proteins has already been done in a previous work (Willocx, 2018). A special focus is made on the candidate ectolysins and depolymerases.

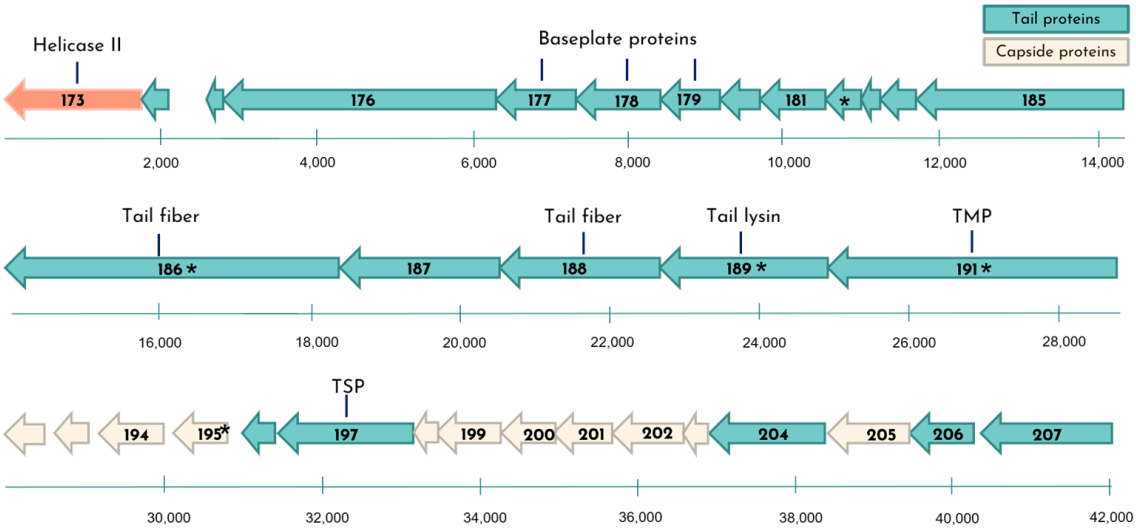

Figure III.1. Region encoding tail proteins in Deep-Blue genome. Blue, white and pink colors indicate tail proteins, capsid proteins and helicase proteins. The stars indicate our candidate proteins. Abbreviations are: TMP- Tape measure proteins; TSP- Tail sheath protein.

Tail fibers and central spikes

Deep-Blue **gp188** BLASTp analysis revealed significant hits with putative glycerophosphoryl diester phosphodiesterase (*i.e.* envelope degradation) and many putative tail fibers. HHpred hits showed homology with several proteins belonging to a class of contractile injection systems. For instance, the membrane-piercing protein gpV of *E. coli* phage P2 and a component of type VI secretion system (T6SS) of *E. coli* T6SS whose role is to efficiently translocate proteins and DNA

across lipid membranes into cells. A common evolutionary history has also been suggested between these systems and cell puncturing phage tails (Leiman *et al.,* 2009). These findings suggest that gp188 may be a central spike participating in host cell puncturing and DNA injection.

The BLASTp analysis of **gp186** was not quite conclusive, and mostly returned matches with minor tail proteins of other *Bacillus* phages. Interestingly though, Interpro analysis revealed the presence of a carbohydrate-binding module (CBM), more precisely a galactose-binding-like domain, whose usual function is to bind to specific ligands, such as cell-surface-attached carbohydrate substrates. Due to its large size (1463 aa) gp186 is likely to be a tail fiber.

Deep-Blue **gp185** was identified experimentally as the RBP in a previous work. This protein possesses an intramolecular chaperone at its C-terminal end. Chaperones are folding-facilitating proteins of which a sub-part are classified as intermolecular chaperones since they are part of the polypeptide chain they help folding and are cleaved off after the folding process is completed. This cleaving process is thought to be achieved only upon folding and trimerization, enabling it to act as a folding sensor (Schulz *et al.,* 2010). As mentioned earlier (*c.f.* Tail fibers), many phage tail fibers are made up of homotrimers and it has been suggested that when they encode the phage RBP, the intramolecular chaperone may cover the binding site and only expose it by auto-proteolysis after correct folding of the homotrimer fiber (Garcia-Doval *et al.,* 2015).

Finally, for **gp176** the sequence homology analysis of BLASTp matched several either putative adsorption-associated protein or hypothetical proteins, including phage A511 gp106, a putative tail fiber.

Baseplate proteins

gp179 and **gp178** tuned out to be well conserved putative baseplate protein. gp179 possesses a conserved 25-like lysozyme domain in the C-terminal region which is part of the GPW_gp25 superfamily and is described as a structural component of the outer wedge of the baseplate that has lysozyme-like activity. For gp178, a Baseplate_J conserved was identified. This domain is described as a protein that lies at the edge of the tail baseplate in *E. coli* phage T4.

BLASTp analysis highlighted the extremely variable nature of Deep-Blue **gp177** C-terminal region. Although no conserved domain could be identified, HHpred revealed homology in the C-terminal region of gp105 of A511 phage, a baseplate protein (Identity 99.91%, E-value 1.9e-23). This all suggests that gp179, gp178 and gp177 are components of Deep-Blue's baseplate structure.

Hypothetical proteins

As regards the other tail proteins, they have not been assigned a function. Importantly, three of them, gp174, gp175 and gp183, are very small in size (less than 100 aa) that limits their potential role as adsorption-facilitating enzymes.

3.1.2. Ectolysin candidates

After considering the previous bioinformatic analysis, four tail proteins were selected as ectolysins candidates. Three of them were located in the tail module whereas one laid outside this region but still in the structural protein-encoding region (*i.e.*, between the tail and the capsid region).

The putative tail lysin gp189

The most promising ectolysin candidate, **gp189** (746 aa), was retained since its BLASTp analysis matched almost exclusively phage tail lysins. The first hits showed homology with putative tail lysins of *Bacillus* phages Bcp1 and BCP8-2 (Coverage 100%, Identity 91,7%, E-value 0.0 and coverage 100%, Identity 91,4%, E-value 0.0, respectively). When looking at conserved domains, both BLAST and INTERPRO matched an endopeptidase domain in the C-terminal end called NlpC/P60, which belongs to a cell-wall hydrolase family of proteins known to cleave peptide cross-bridges between glycan chains. In HHpred, the N-terminal region of gp189 showed homology with several structural associated proteins such as a tail protein of *Shewanella oneidensis* MuSo2 prophage (Identity 99.4%, E-value 8.5e-12) and enterobacteria phage Mu baseplate protein gp44 (Identity 99.34%, E-value 6e-11). Finally, Phyre[2] hits showed the same structural tail protein of *S. oneidensis* prophage MuSo2 in the N-terminal region and an encouraging amount of hydrolase associated activity hits in the C-terminal region. For instance, the best match was a gamma-d-glutamyl-l-diamino acid endopeptidase (npun_r0659) described in the cyanobacteria *Nostoc punctiforme* (100% confidence for homology).

In a nutshell, gp189 was retained due to the high number of similarity hits associated with peptidoglycan degrading enzymes, the presence of an endopeptidase domain on the C-terminal end and a structural one in the N-terminal end. These findings suggest that gp189 could easily fit the function of a structural tail protein involved in local degradation of the peptidoglycan layer prior to genome injection into the host cell (Fig. III.2.).

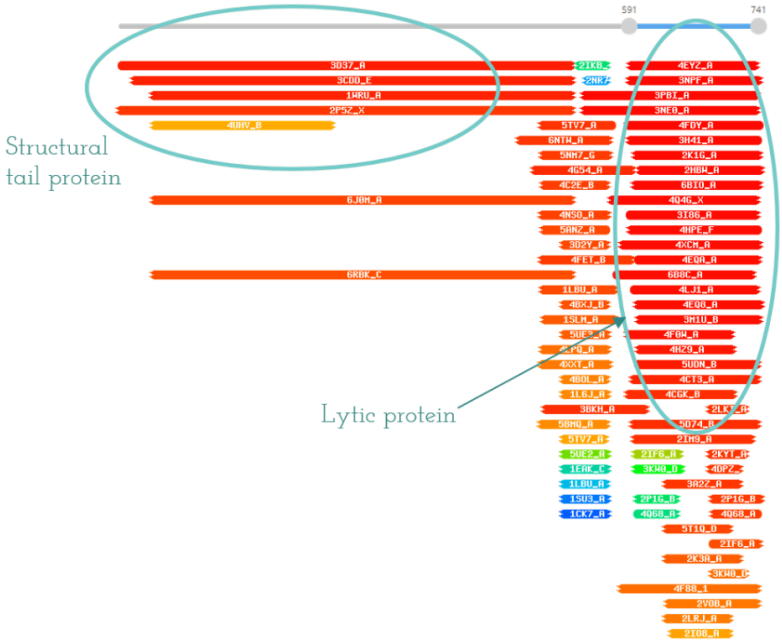

Figure III.2. Graphic summary of HHpred hits for gp189. Blue circles indicate the high amount of hits associated to structural proteins in the N-ter end and to lytic proteins in the C-ter end.

The Tape measure protein gp191

The TMP **gp191** (1281 aa), which is responsible for determining the length of the tail, was also selected. Indeed, it harbors a Mannosyl-glycoprotein endo-beta-N-acetylglucosamidase-like domain that has also been found in flagellar protein J (P75942), known to hydrolyse peptidoglycan (Nambu *et al.,* 1999). Moreover, it is not uncommon to find lytic activity associated with the TMP: *e.g., E. coli* phage T5 TMP protein pb6 is a pore-forming protein and peptidoglycan hydrolase (Zivanovic *et al.* 2014).

The cysteine peptidase gp182

Despite not being able to attribute a function to it, the hypothetical protein **gp182** also drew our attention since several cysteine peptidase associated hits were matched (e.g. cysteine peptidase of *Bacillus* phage Kioshi; Coverage 100%, Identity 91.93%, E-value 0.0). This was confirmed independently when looking at Phyre[2], which also returned cysteine peptidase/ hydrolase headers as first matches (Confidence 99.9%). Moreover, despite its small size of 161 aa, Interpro analysis showed that this protein harbors the same NIpC/P60 endopeptidase domain previously identified in gp189. Nevertheless, this domain occupies the entire length of the proteins, which leaves no place for a structural domain.

Gp195

Finally, gp195 (227 aa), was selected given it was annotated as an L-alanoyl-D-glutamate peptidase in Deep-Blue genome. A 3D domain that consists of three conserved aspartate residues was also identified in the C-terminal end (E-value 1.15e-44). This domain is usually found in active sites of mltA-like lytic transglycosylases (i.e. cleave the beta-1,4 glycosidic linkages between MurNAc and GlcNAc in peptidoglycan). This domain is part of the Rare lipoprotein A (RlpA) domain superfamily of which some proteins are known to contribute to cell separation in *P. aeruginosa* and thus to PG degradation (Jorgenson *et al.* 2014).

A table summarizing these findings and classifying the candidates can be found later in Table III.1.

3.1.3. Depolymerase candidates

In the first part of this research, gp186 was thought to be a candidate for depolymerase activity since it harbors a sugar-binding domain or carbohydrate-binding module (CBM). However, none of the bioinformatic analysis tools was able to detect a lytic domain associated with any depolymerase activity, which is why other depolymerases outside the structural region were also considered. Two candidates were selected and are found in proximity of the gp221 endolysin.

Gp223

gp223 (790 aa) BLASTp analysis returned several right-handed parallel beta-helix repeat-containing protein associated hits in a varied number of strains from the *Bacillus* genus. The C-terminal end was also noted to be poorly conserved (hits below 47% identity), however, a conserved right-handed beta helix region was identified (E-value 2.39e-08, from the 321 to the 467 aa) and consists of parallel beta helix region that shares some similarity with Pectate lyases. This was confirmed independently by Interpro, which identified entries closely related to pectin lyase and composed of the same parallel beta strands stacks. Such a tertiary structure has often been associated with enzymes with polysaccharide substrates (Jenkins *et al.,* 1998). The HHpred search revealed an important number of high probability matches often related to adhesion and anchor proteins. Interestingly, when looking at hits in phages, one tail spike of *Shigella flexneri* phage Sf6 showed high homology (Identity 99.75%, E-value 1.2e-15) for the N-terminal end, which has been associated with endorhamnosidase lytic activity (Müller *et al.,* 2008). Additionally, in the C-terminal end, another match was found with *Salmonella* myovirus Det7 where the Dettilon tail spike protein (gp208) has a sugar-binding ability since it recognizes the *Salmonella anatum* O-antigen (Broeker *et al.* 2019) (Fig.III.3.). These findings suggested that

gp223 could mediate both attachment and degradation of polysaccharides at the surface of Deep-Blue hosts.

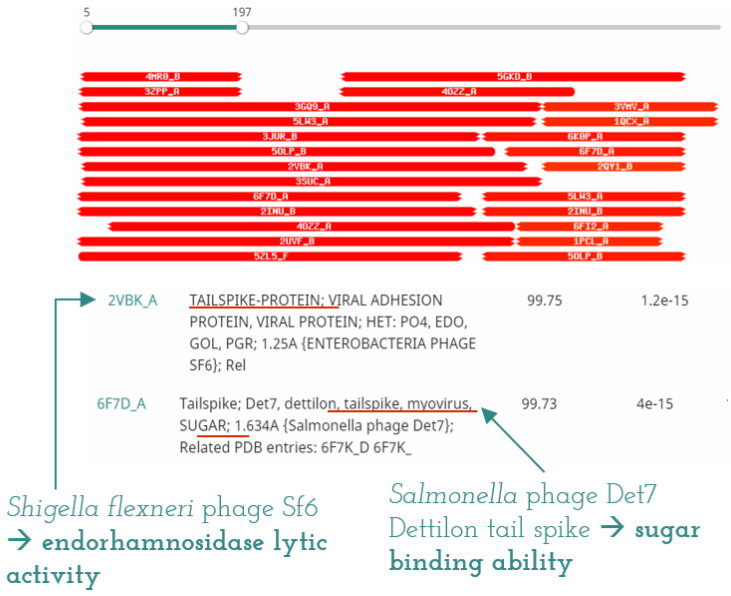

Figure III.3. Graphical summary of HHpred hits for gp223. Two important hits are highlighted by an arrow.

Gp222

For another part, **gp222** (161 aa) BLASTp analysis revealed a very poorly conserved N-terminal end among other *Bacillus* phages (hits below 62.7% identity) but was not able to predict a putative function. Only one significant hit was retrieved using HHpred and consisted of the tail spike protein ORF212 from *E. coli* CBA120 phage that harbors a glycosidase active site in the C-terminal end (Identity 98.45%, E-value 2.5e-7). Phyre[2] indicated structural similarity with the same tail spike protein and confirmed its putative lytic nature (Confidence 99.1%).

Table III.1. Bioinformatic analysis summary of ectolysin and depolymerase candidates of Deep-Blue in order of relevance.

	Size (aa)	**BLASTp**	**InterPro**	**HHpred**	**Phyre[2]**
Ectolysin candidates					
gp189	746	Tail lysins	Endopeptidase NLPC/P60 domain [617-742]	C-ter NLPC/P60 family, peptidase and N-ter structural module	Cell wall hydrolase
gp191	1281	TMP, PG hydrolase, lysin	Mannosyl-glycoprotein endo-beta-N-acetylglucosamidase domain	PG hydrolase, glycosidase in C-ter	-
gp182	161	Cysteine peptidase	Endopeptidase NLPC/P60 domain [5-157]	-	Hydrolase
gp195	227	L-alanoyl-D-glutamate peptidase	C-ter 3D (Asp-Asp-Asp) domain [117-224]	C-ter lytic transglycosylase, PG remodelling	-

Depolymerase candidates					
gp223	790	-	Beta helix repeat containing domain, pectin lyase [321-467]	Alginate lyase, Fibronectin-binding protein A	Exo-poly-α-D-galacturonosidase, Adhesion protein
gp222	161	Hypothetical protein, tail fiber	-	Endo-glycosidase	Tail spike

3.2. Preliminary observations

3.2.1. Degradation Halo: a depolymerase hallmark

Phages equipped with depolymerases form plaques surrounded by a degradation halo. This halo is often seen as an indicator for depolymerase activity and was observed when plating Deep-Blue together with the sensitive strain Si0239 (Fig. III.4.). The size of the halo often increases during prolonged incubation. The origin of this halo is not entirely understood. However, Knecht *et al.* have recently suggested that, since depolymerase activity does not necessarily affect bacterial growth *in vitro* (only removes the capsule), halo formation could be due to the depolymerase activity in diffusing offspring virions. Indeed, bacteria surrounding plaques could exhibit a modified physiology making them insensitive to infection. A second hypothesis states that halo formation could be the result of soluble depolymerases released as free enzymes, in the case of an alternative start codon is used in the translation. Consequently, bacteria stripped from their protective capsule become insensitive towards phage infection (Knecht *et al.,* 2020).

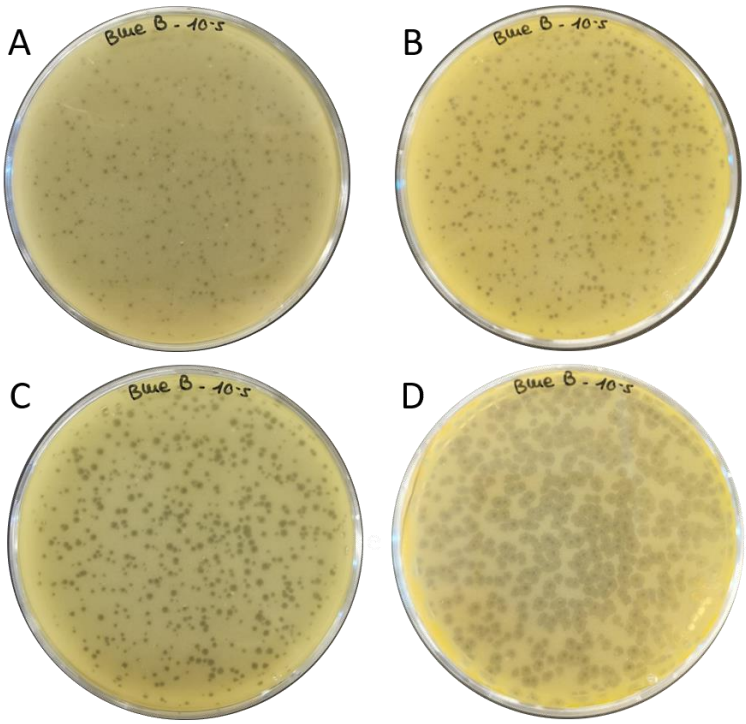

Figure III.4. Double layer agar of Si0239 and Deep-Blue after (A) one, (B) two, (C) three and (D) six days of incubation at 30°C. The size of the halo surrounding lysis plaques increases over time.

3.2.2. Lysis-from-without phenomenon: ectolysin evidence

In phage infection, the virion uses its tail lysin to locally degrade the PG layer and inject its genome. Consequently, after virion production and assembly, the bacterial host cell is lysed thanks to endolysins leading to the release of new viral particles. In the lysis-from-without (LFW) phenomenon, a high virion concentration leads to premature bacterial lysis due to a high number of local degradations although no phages are produced. Thus, it does not require phage infection but instead is directly mediated by extracellular agents called ectolysins (Abedon, 2011).

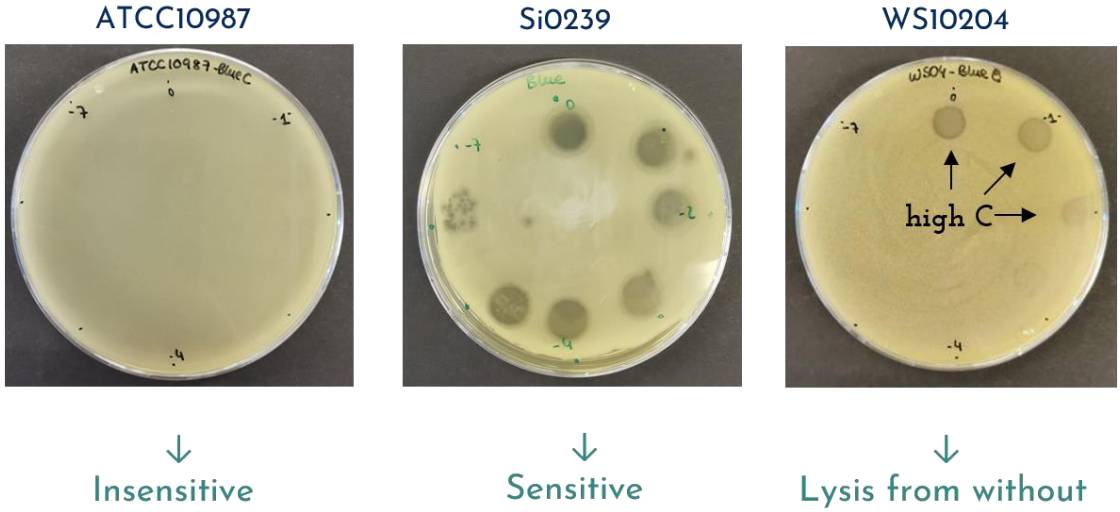

Figure III.5. Insensitive (ATCC10987), sensitive (Si0239) and lysis from without (WS10204) strains plated with different Deep-Blue concentrations. Initial phage concentration is 10^9 PFU/mL.

This phenomenon was observed when spotting droplets of increasing concentration of Deep-Blue on lawns containing *B. cereus* strains. Indeed, LFW strains show a pattern of lysis spots at high concentrations and no individual plaque whereas sensitive strains have both and insensitive strains none (Fig. III.5.) The dependence on this ectolysin-type phenomenon, therefore, is further suggestive of a *bona fide* ectolysin in Deep-Blue.

3.3. Adsorption-independent ectolysin activity

The question of whether the LFW phenomenon, and thus the ectolysin activity, relies on RBP specific attachment to the bacterial surface remained. To address this question, phage adsorption tests were performed on several strains to assess Deep-Blue adsorption capacities. RBP binding was obtained by performing cell wall binding assays relying on the incubation of a given strain with GFP-fused RBP. If the RBP (gp185) is able to adsorb to the strain, the bacteria will appear as green under the fluorescent microscope. By assessing both Deep-Blue adsorption rate (Table III.2) and the RBP binding capacity (Fig. III.6.) to strains affected by lysis from without, we expect to establish if the binding of the RBP is mandatory for the ectolysin to act.

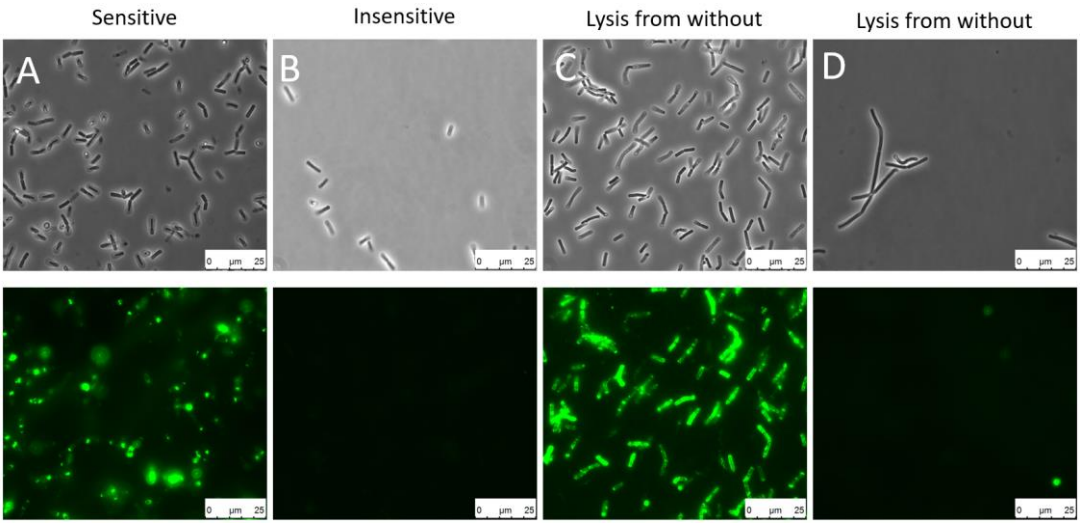

Figure III.6. Cell wall binding assay for Deep-Blue's RBP, gp185 on (A) sensitive Si039, (B) insensitive ATCC10987, (C) lysis from without sensitive WS10204 and (D) lysis from without sensitive MC118-4. Left : Bright field microscopy. Right : fluorescence micrograph.

Globally, the results showed that in our sensitive control, Si0239, the receptor is present, which explains the high adsorption percentage and RBP binding. In contrast, in the insensitive control, ATCC10987, no adsorption was detected and the GFP fused RBP was not able to decorate the cell thus suggesting that no proper receptor was present. However, in the case of LFW sensitive strains two types of behaviors were identified: LFW sensitive strains with no adsorption and no RBP attachment (*e.g.* MC118-4, Fig. III.6D) and LFW sensitive strains with adsorption and RBP binding (*e.g.* WS10204, Fig. III.6C) (Table III.2.). Consistently, all strains showing high adsorption percentages were also tested positive for the RBP binding.

Given that LFW was observed with and without RBP-mediated adsorption, ectolysin activity at the origin of the LFW phenomenon is likely not dependent on the specific attachment of the phage but on its concentration solely. This then raises the question of why the phage cannot multiply in strains where RBP attachment and adsorption are achieved such as WS10204. One possible explanation might be related to a post-infection defense mechanism (*e.g.* CRISPR-Cas or RM system) used by these strains to prevent the phage from entering its lytic cellular cycle.

Table III.2. RBP binding, phage sensitivity and adsorption percentage of Deep-Blue for strains of the B. cereus group. RBP binding experiments were in part performed in a previous study (Willocx, 2018). Green, red and yellow corresponds to sensitive, insensitive and lysis from without character of the strains, respectively. Adsorption percentages were obtained from three independent experiments.

Strain	Species	RBP (gp185) binding	Phage Sensitivity	Adsorption %
Si0239	*B. weihenstephanensis*	+		97.24 ±3.11
ATCC10987	*B. cereus*	-		0 ±0
KBAB4	*B. weihenstephanensis*	-		2.92 ±3.23
MC67-2	*B. weihenstephanensis*	-		9.02 ±8.35
MC118-4	*B. weihenstephanensis*	-		2.30 ±3.99
Aw43	*B. thuringiensis*	+		89.47 ±8.21
WS10204	*B. weihenstephanensis*	+		80.10 ±15.27
Si0170	*B. weihenstephanensis*	+		78.34 ±11.19
H3081	*B. cereus*	+		49.22 ±12.87
GBJ001	*B. thuringiensis*	+		98.89 ±0.62
NVH391-98	*B. cytotoxicus*	+		95.15 ±7.14

3.4. Molecular cloning and induction of candidates

3.4.1. Cloning experiments

To assess their activity, each candidate gene was cloned and expressed in *E. coli*. However, to ease the expression and purification of some candidate proteins, only the lytic domain-containing segments were cloned. This is the case of the TMP gp191 (1281 aa) for which only the C-ter segment (1000 - 1281 aa) was cloned, hereafter called gp191_1000. As for the putative lysin gp189 (746 aa), we parallelly worked with the whole protein and with the N-ter segment (598 - 746 aa) containing the lytic domain, thereafter, called gp189_lyt. Figure III.7. shows a summary of all the constructions.

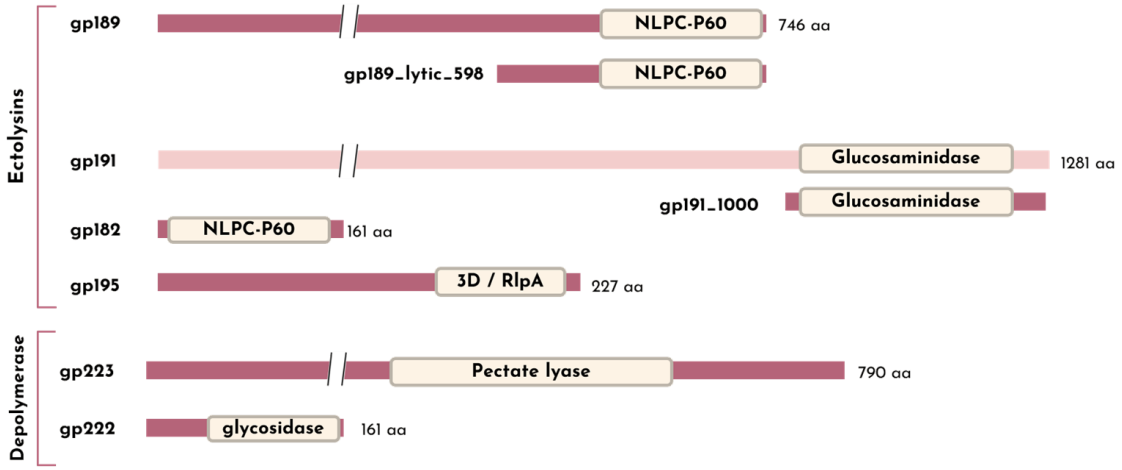

Figure III.7.. Summary of the constructions completed in this study. Light pink stands for non-cloned genes whereas dark pink represents the genes or truncated genes that were cloned and expressed.

3.4.2. Protein induction and assessment of solubility

The recombinant proteins were induced by the addition of IPTG to a culture of *E. coli* BL21(DE3) containing the plasmid constructions. Because every protein is different and might need different induction conditions, parameters of induction (*i.e.*, IPTG concentration, duration and temperature) were optimized to obtain soluble proteins and are summarized in Table III.3.).

Optimal induction conditions could not be found for gp189_lyt, gp182 and gp195. Regarding the latter two, not many conditions were tested since they were not the most promising candidates of this study. Instead, we preferred focusing on gp191 and gp189. In contrast, several conditions were tested for gp189_lyt as shown in Table III.4. However, none of them led to the obtention of a soluble protein. Glucose (at final concentration 0.1 M) was even added to the culture medium to decrease basal protein expression.

Table III.3. Optimal induction parameters for protein candidates.

	[IPTG] (mM)	Time (h)	Temperature (°C)
Ectolysin candidates			
gp189	1	3 to 6	28
gp189_lyt	Not soluble		
gp191_1000	1	6	28
	0.5	O/N	22
gp182	Not soluble		
gp195	Not soluble		
Depolymerase candidates			
gp223	1	O/N	18
	0.5		22
gp222	0.5	O/N	18

Table III.4. Induction conditions tested for gp189_lyt. Glucose was added at 0.1 M.

[IPTG] (mM)	Time (h)	Temperature (°C)
1	6	28
1	3	28
1 + Glucose	6	28
1	O/N	18
0.5	O/N	18

Interestingly, it was noticed that during gp189_lyt induction the culture media became clearer instead of becoming more turbid as the other cultures, thus indicating poor bacterial growth (Fig. III. 8A). The OD was therefore monitored every hour during the induction of gp189_lyt. BL21(DE3) containing the empty vector (pET30) served as control. The following induction conditions were used: 4 h at 28°C with or without 1 mM of (Fig. III.8B). In the absence of IPTG, the culture expressing gp189_lyt shows comparable growth to that of the control with pET30. However, upon IPTG induction, the OD rapidly decreased and was divided in two during the second hour. This observation may suggest that the lytic domain of gp189 is indeed really toxic to *E. coli* and its production may represent a high burden for the cell. Additionally, we were not able to detect the protein on an SDS page gel after induction which may imply that the protein was produced in such a small quantity that it would be necessary to perform a Western blot, much more sensitive, to detect it.

Proteins carrying NlpC/P60 lytic domains can be lethal to bacteria by compromising their cell wall integrity, which is why their activity is tightly controlled. For instance, the atomic structure of the entire protein is often involved in substrate specificity (Xu *et al.,* 2010). Given that the structural N-ter end was suppressed in gp189_lyt it is conceivable that substrate specificity might have been compromised and that lytic activity of the domain became less restricted. To get around this problem, the protein could be expressed in BL21(DE3)pLysS, an expression strain suited to express toxic proteins since it expresses T7 lysozyme to reduce basal expression levels.

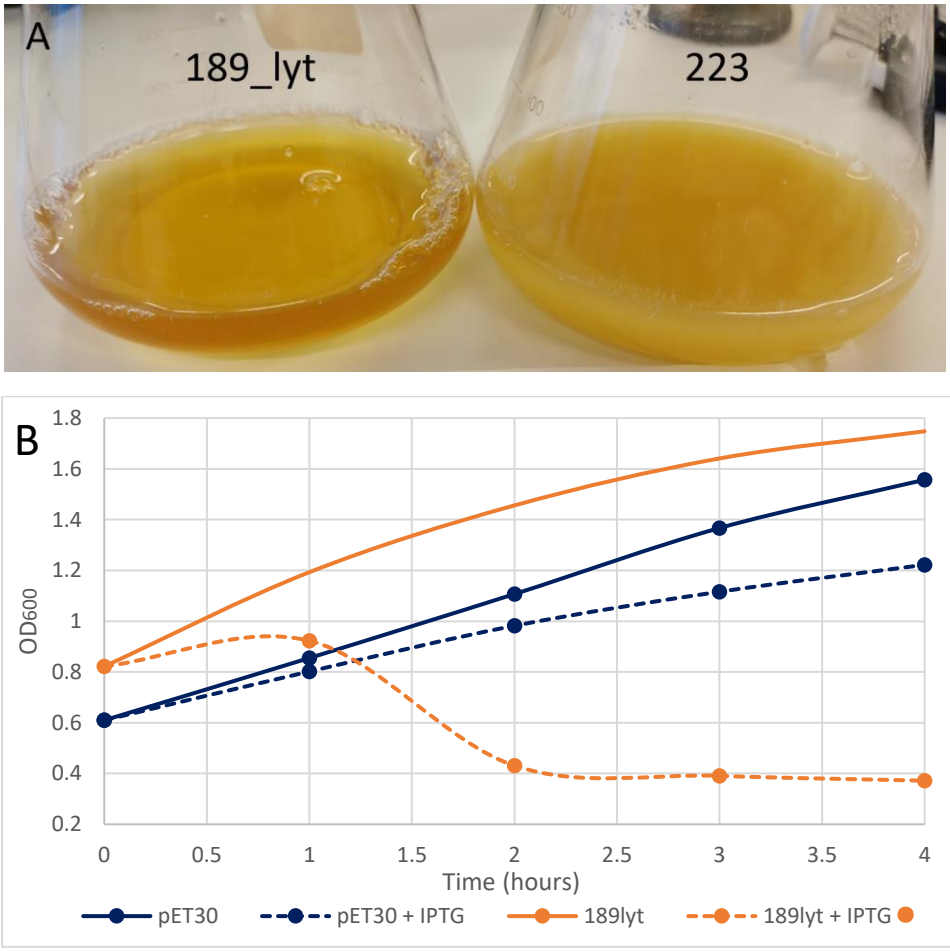

Figure III.8. Toxicity of gp189_lyt upon induction with IPTG. ***(A)*** *Culture media of BL21(DE3)[pET30::189_lyt] (left) and BL21(DE3)[pET30::223] (right) after O/N induction at 18°C and 0.5 mM IPTG.* ***(B)*** *DO_{600} monitoring during induction of BL21(DE3)[pET30::189_lyt] and BL21(DE3)[pET30] for 4 h at 28°C, with or without IPTG (1mM).*

As shown in Fig. III.9., the protein was suited for purification when the corresponding band was found in the soluble fraction (SF). This was the case for gp189 (87,6 kDa), gp191_1000 (34,7 kDa), gp222 (20 kDa) and gp223 (92,9 kDa) whereas for gp189_lyt (21,4 kDa), gp182 (23,7 kDa) and gp195 (25.5 kDa) (latter two not shown) the proteins were not soluble and no band was observed in the SF column.

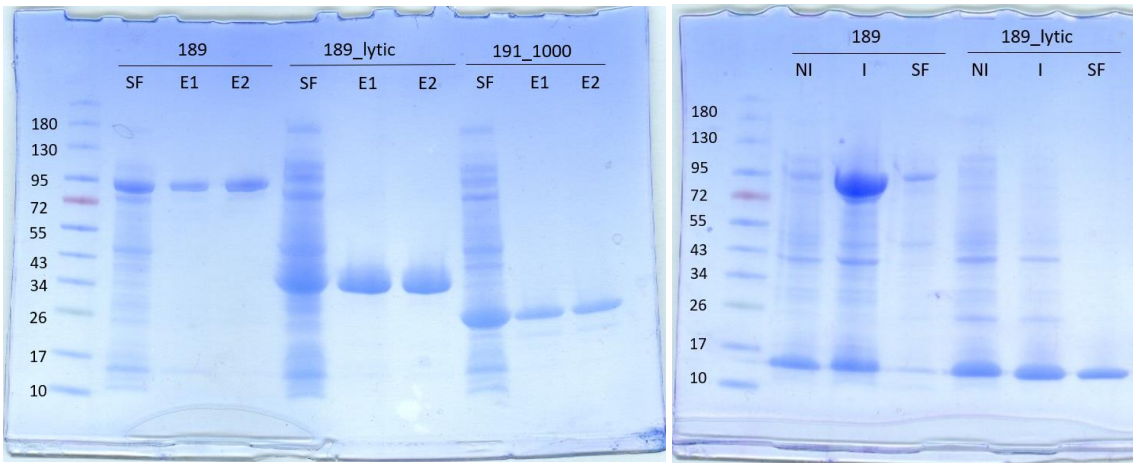

Figure III.9. SDS-PAGE gels of induced recombinant proteins. Induction parameters are 6 h at 28 °C with 1 mM of IPTG. First column of the gel contains the protein ladder reference, NI: Non-Induced, I: Induced, SF: Soluble Fraction, E1 or E2: Eluate 1 or 2.

3.4.3. Protein purification and concentration

Purification was performed in native conditions to preserve protein biological activity and to avoid a renaturation step at the end of purification. The Soluble proteins gp189, gp191_1000, gp222 and gp223 were purified on Ni-NTA columns thanks to their N-terminal 6-histidine tag, and purity was verified by SDS-PAGE (Fig.III.10.). Purified proteins gave concentrations around 0,2 to 0,5 mg/mL Thus, proteins were concentrated by ultrafiltration to obtain concentrations around 0,5 to 1 mg/mL.

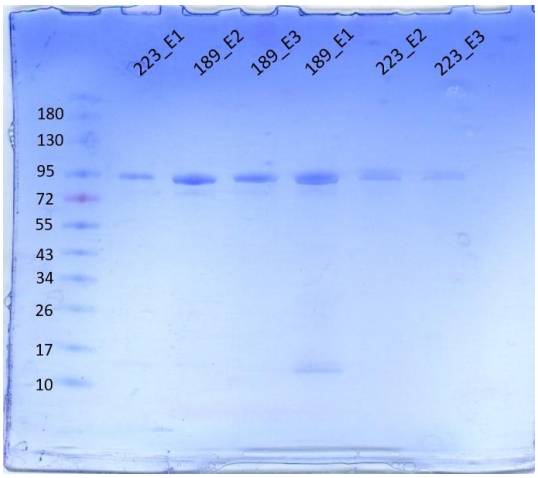

Figure III.10. SDS-PAGE of gp189 and gp223 purification fractions. Abbreviations are: E1: elution 1; E2: elution 2.

3.5. Assessment of the activity of ectolysin and depolymerase candidates

3.5.1. The putative ectolysins gp189 and gp191_1000

The lytic activity of the ectolysin candidates was assessed in a turbidity reduction assays where the OD was measured every minute for 30 min. The lytic activity was tested in different strains of the *B. cereus* group (Table III.5.). Turbidity reduction assay in Si0239 for both candidates are shown in Fig. III.11. Only two repetitions were performed for gp191_1000 since no lytic activity could be observed. However, in WS10204, some lytic activity (inhibition of 28.3 %) was registered (Fig. S1) but when repetition was attempted this result was no longer observed. Thus, the results provided by this type of test are not as stable as desired, which explains the standard deviation values obtained. Therefore several congruent outputs need to be obtained for a solid hypothesis to be legitimated. This is the case for gp189, whose lytic activity showed consistency across repetitions and acceptable standard deviation values.

Table III.5. Inhibition percentages of gp191_1000 and gp189 calculated based on turbidity reduction assay. Each percentage is obtained by averaging two (gp191_1000) or three (gp189) independent measurements. Phage sensitivity and inhibition standard deviation values are also showed. Abbreviations are: N.D for Non Determined. Green, red and yellow correspond to sensitive, insensitive and lysis from without character of the strains, respectively.

	Si0239	BtB2-4	WS10204	GBJ001	ATCC10987
Phage sensitivity					
gp191_1000	4.08 ± 5.8	N.D	14.5 ± 20.0	0	0
gp189	44.4 ± 12.4	39.0 ± 2.5	40.5 ± 8.4	42.9 ± 8.9	10.04 ± 7.9

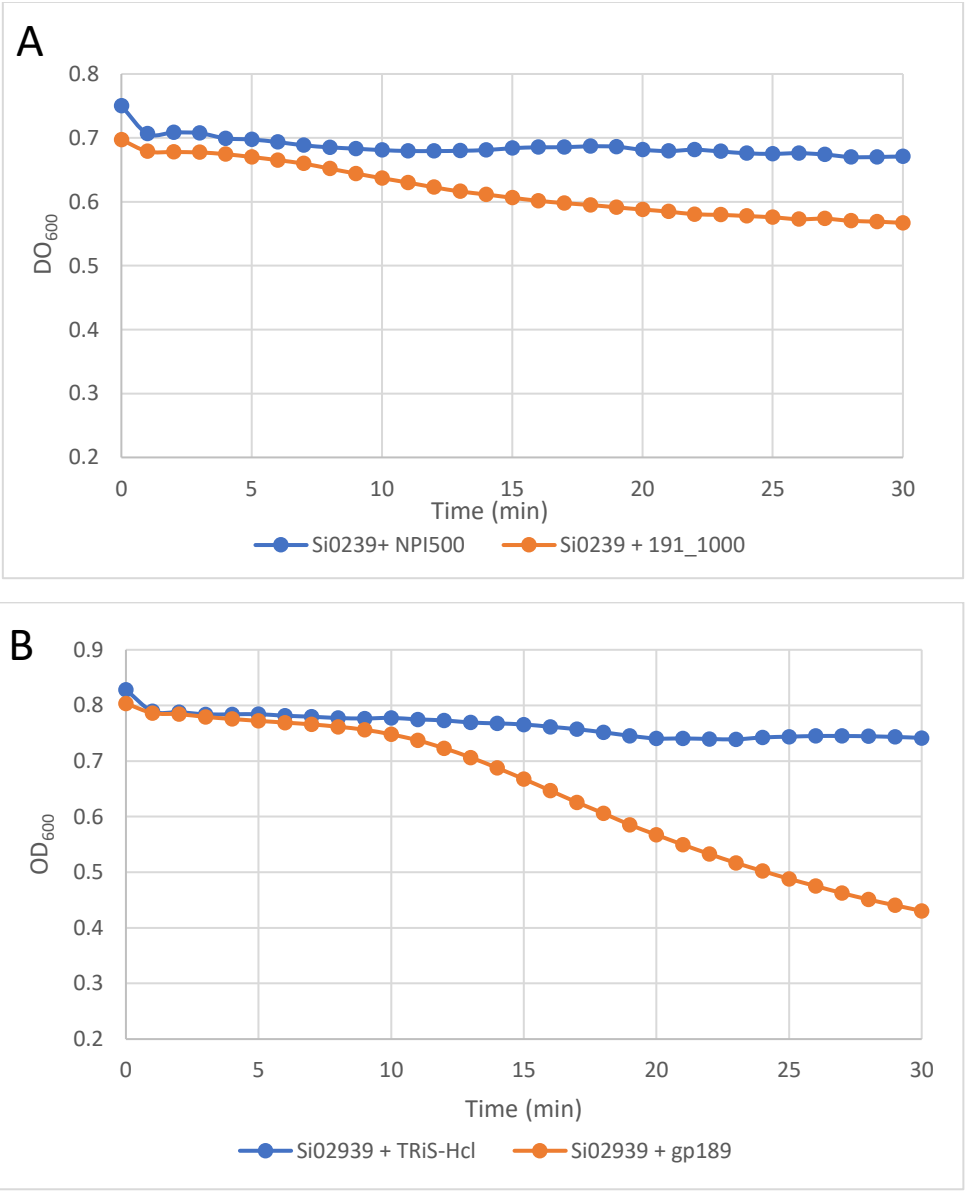

*Figure III.11. Turbidity reduction assay of **(A)** gp191_1000 and **(B)** gp189 on Si0239 conducted for 30 min. Protein concentrations tested are approximately 0.5 mg/mL.*

The overall results indicate a strong inhibition of cell growth by gp189 and none by gp191_1000, regardless of the strain. Regarding gp189, all strains, except for the insensitive strain ATCC10987, showed important inhibition percentages (from 39 to 44.4 % and 10.4 %,

respectively). Being the most promising candidate from the start of this study, these results strongly suggest that gp189 is most likely Deep-Blue's ectolysin.

To evaluate the effect of gp189 concentration on the tested strains, increasing concentrations of this protein were added to a Si0239 cell culture undergoing the same turbidity reduction assay. As shown in Fig.III.12, higher protein concentrations led to higher inhibition percentages, in what can be described as a dose-response behavior. After 20 min, an inhibition of 38.4 % was reached when treated with 100 µg/mL of gp189 whereas 12.7 % inhibition was measured when treated with 12 µg/mL.

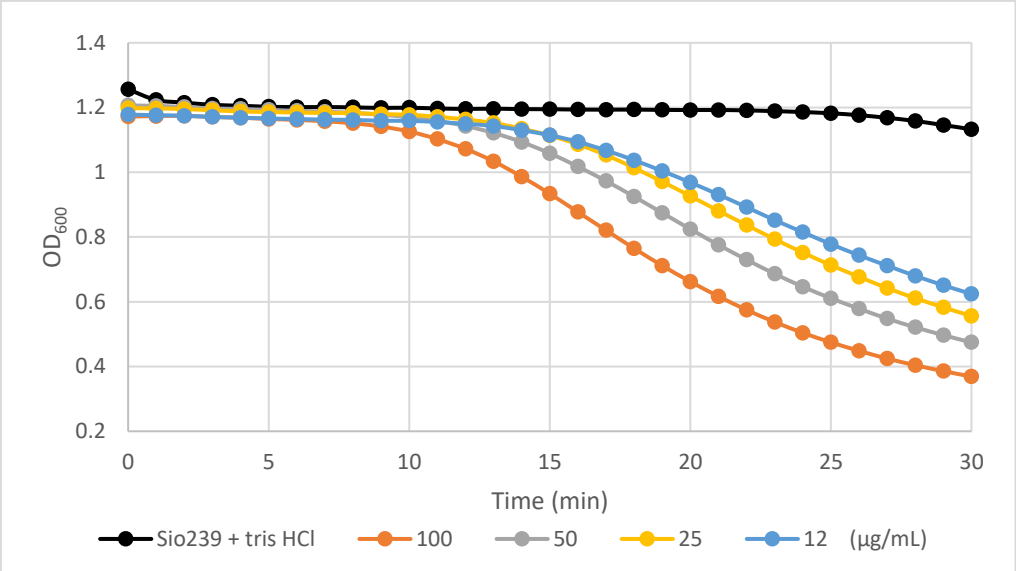

Figure III.12. Turbidity reduction assay of decreasing concentrations of gp189 (ranging from 100 to 12 µg/mL) on Si0239 culture.

3.5.2. The putative depolymerases gp222 and gp223

Given that phages using depolymerases show a degradation halo around their lysis plaques, the purified depolymerase is expected to have the same effect when spotted on a sensitive bacterial lawn. Therefore, activity tests for gp222 and gp223 were conducted by spotting the concentrated purified candidates (between 0.25 and 1 mg/mL) on several different bacterial lawn top agars (Si0239, VD21, WS10204, BtB2-4,GBJ001 and ATCC10987) next to a drop of the elution buffer used in purification and concentration (PBS-NaCl). Unfortunately, no difference in clearing of the bacterial lawns was ever observed compared to the control. This test was also performed on bacterial lawns incubated for 6 or 24 h prior to protein spotting. Despite these attempts, no difference was observed between the control and both proteins.

If the purified proteins are active this would probably mean that these candidates do not encode for Deep-Blue depolymerase. Moreover, it is unlikely that an eventual effect was missed due to a short observation time since halos have been observed just 30 min post-spotting for other phages depolymerases (*e.g. P. aeruginosa* phage IME180) (Mi *et al.*, 2019).
However, it is always plausible that the proteins tested were not active, explaining why the depolymerase effect could not be observed. To remediate that, several supplementary experiments could be performed. For instance, the His-Tag used for purification can sometimes hinder protein activity and its cleavage by an enterokinase could be considered. Similarly, some proteins are less stable in some buffers and a change in the purification and concentration buffer could also be envisioned, for instance with the TRIS-HCl buffer used in ectolysin activity tests.

3.6. Phage CsCl purification

One of the aims of this thesis was also to demonstrate the structural nature of the enzyme candidates thanks to a mass spectrometry (MS) analysis of whole Deep-Blue viral particles. If the bands corresponding to these proteins were identified, their structural nature would be confirmed. For this purpose, phage particles were multiplied and further precipitated thanks to polyethylene glycol (PEG) addition. Given that MS analysis requires extremely pure and concentrated samples, phages were subjected to a CsCl purification process. Two protocols were tested in order to obtain a phage concentration of at least 10^{10} PFU/ml.

When using the gradient step centrifugation technique (Protocol 1), phages are sedimented through a discontinuous step gradient formed by layers of CsCl solutions of increasing density stacked on top of each other (Boulanger, 2009). Their migration stops at around 1.45 and 1.52 g/mL since known tailed phages infecting a variety of Gram-positive and Gram-negative bacteria have that buoyant density in CsCl (Carlson, 2005). This protocol was performed multiple times by varying the optimization parameters such as the CsCl layer densities (1.6, 1.5, 1.3 or 1.6, 1.5, 1.4, 1.3 or 1.7, 1.5, 1.45) and ultracentrifugation duration (2 or 3 h). Reducing the total volume of the centrifugation tube (from 15 mL to 4,5 mL) while keeping the CsCl layer was also attempted with no successful results.

In the second protocol tested, using the equilibrium centrifugation technique, instead of depositing the collected phages on top of increasing-density CsCl layers, CsCl was directly added to the phages in a unique density-layer at 0.75 g/mL concentration. Centrifugation of a homogeneous density CsCl solution creates a continuous density gradient allowing phages to migrate (Boulanger, 2009). The ultracentrifugation parameters were adapted to last 24 h. In spite of these modifications, phage purification continued to retrieve low concentrations of phage (C_{final} = 10^3 to 10^6 PFU/mL) despite high concentrations inputs prior to the ultracentrifugation step ($C_{initial}$ = 10^8 to 10^{13} PFU/mL).

Purification by CsCl ultracentrifugation not being an especially difficult technique to implement, it was considered that Deep blue may not be compatible with this technique. In fact, this approach has been reported to be inefficient for a number of phages that either get damaged by the centrifugal forces, suffer from osmotic shock or interact with CsCl and lose their infectivity (Carlson, 2005). As a result, a drop in the viable phages retrieved is observed in many cases as it was observed in this work and as in *B. thurengiensis* phage 0305Ø8-36 CsCl purification (Pathria *et al.,*2012). For this reason, phage behavior in CsCl solutions should be investigated prior to using this purification technique in the future.

Some alternatives can still be implemented to purify Deep-Blue. In the case of phage destruction by the centrifugal forces required by ultracentrifugation (above 100.000 *g*), protocols using general centrifugation (maximum 40.000 to 50.000 *g*) demonstrated to have almost the same concentration yield and purity as conventional density gradient ultracentrifugation (Nasukawa *et al.,* 2017). For phages where a low CsCl tolerance has been evidenced, the use of another type of gradient is conceivable. For example, CsCl purification of *Bacillus megaterium* phage G triggers DNA expulsion from the capsid and its purification is therefore performed in a sucrose gradient (Serwer *et al.,* 1978). Another alternative to both CsCl and ultracentrifugation is purification by monolithic ion-exchange chromatography, which has reported high yields of infectious particles with concentrations up to 5 mg/mL, sufficient for all types of biochemical analysis (Adriaenssens *et al.,* 2012).

Chapter IV

Conclusions and perspectives

Ever since antibiotic resistant became a worldwide concern, interest for bacteriophages exponentially reemerged due to their interesting antimicrobial properties. Phages and their lytic proteins can be harnessed in a variety of fields including food safety, medecine and overall scientific research. Thus, understanding how these viruses prey on their specific hosts is essential. Combined with their high specificity, phages can be compared with “precision” weapons against pathogenic bacteria and the threat posed by their constant evolution.

The purpose of this master thesis was to identify such antibacterial proteins and more specifically, depolymerases and ectolysins, involved in the adsorption of the *B. cereus* phage Deep-Blue. These enzymes, not necessarily present in all phages, allow surface polysaccharide and peptidoglycan degradation thus facilitating diffusion towards the RBP and injection of viral genome through the cell wall, respectively.

The first step towards this goal was an in-depth bioinformatic analysis of the region encoding structural proteins of the tail given that most adsorption-facilitating enzymes previously identified were of structural nature. Four ectolysin candidates, gp189, gp191, gp182 and gp195 were retained since they contained PG-degradation associated domains (*e.g*. NlpC/P60 domain for gp189 and gp182). Only two depolymerase candidates, gp222 and gp223, located outside the structural region, were selected for their consistent matches with sugar-binding and polysaccharide degrading associated hits (*e.g.* Pectate lyase domain in gp223).

Candidate proteins were then cloned into the expression strain BL21(DE3). Truncated versions of gp189 (gp189_lyt) and gp191 (gp191_1000) were also constructed to facilitate purification. Optimal induction conditions could not be found for gp182 and gp195, however further parameter optimization should be sufficient to obtain soluble proteins. Induction of gp189_lyt turned out to be lethal for the expression strain. It is, therefore, worth exploring if the expression of gp189_lyt can be achieved by using strains better suited for toxic protein expression such as BL21(DE3)pLysS, expressing T7 lysozyme reducing basal expression levels.

For ectolysins, the activity test consisted in a turbidity reduction assay. Preliminary results showed that gp189 was the only ectolysin candidate displaying consistent lytic activity on all sensitive strains tested, with up to 44.4 % of inhibition on Si0239. Thus, it is likely that gp189 is Deep-Blue’s ectolysin. To go one step further and confirm gp189 target, the same tests could be carried out directly on purified peptidoglycan. A biochemical analysis of the protein would also provide information on gp189 stability in different pH and temperature conditions as well as the effect of salinity on its activity. Finally, the catalytic mechanism and the amino acids involved in PG degradation could also be investigated. The NlpC/P60 lytic domain carried by gp189 (617-742 aa) is indeed a cysteine peptidase, capable of hydrolyzing peptide bonds using the thiol group of a cysteine residue as a nucleophile (Barrett and Rawlings, 2001). Within this interval, only two cysteines are encoded, Cys638 and Cys645. Therefore, it could be interesting to mutate or delete these residues to assess their impact on gp189 lytic activity.

Additionally, the lysis from without phenomenon, in which a high virion concentration causes bacterial lysis but no phage progeny production, was also studied since it constitutes an indicator of ectolysin activity. It was demonstrated that specific adsorption onto the bacterial surface was not essential for ectolysin activity given that the lysis from without phenomenon was observed for strains lacking the RBP.

Concerning depolymerases, purified candidates were expected to produce a degradation halo when spotted on a sensitive bacterial lawn. Despite many different bacterial lawns with different incubation time were used, no difference in clearing of the bacterial lawns was ever observed compared to the control. In the case of the proteins being in their active form, this would probably indicate that these candidates are not depolymerases of Deep-Blue. However, it always remains possible that the proteins tested were not active, explaining why no depolymerase activity could not be observed. To remediate that, several supplementary experiments could be performed such as cleavage of the His-Tag or finding a buffer increasing the protein stability (*e.g*. Tris-HCl buffer used in ectolysin activity tests).

To prove the structural nature of the selected candidates, it was considered to perform an MS analysis of Deep-Blue viral particles. A concentrated phage sample was obtained through PEG precipitation of the multiplied phages. To obtain a sample pure enough to be used in MS, the classic CsCl gradient purification by ultracentrifugation was attempted. Even though two different protocols were tested and adjusted several times, phage purification continued to retrieve phage concentrations insufficient for MS ($C_{initial}$=10^8-10^{13} and C_{final}=10^3-10^6 PFU/mL). It was thus hypothesized that Deep-Blue may not be compatible with this technique as it has been the case for many other phages. However, other alternatives to ultracentrifugation and CsCl, which may be the cause of the observed drop in viral particles, could be implemented (*e.g.* general centrifugation and sucrose gradient, respectively).

This master thesis focused on studying two enzymes used by Deep-Blue, a phage praying on *B. cereus* group, to improve their adsorption and infection chances. In the end, although more research is needed to confirm these results, the work carried out throughout this study strongly suggests gp189 to be Deep-Blue's *bona fide* ectolysin. Further work is required to better understand the functioning of this enzyme; however, it was established that its action was independent of RPB-receptor interaction even though some strains remain completely insensitive. Unfortunately, the depolymerase function could not be assigned to either of the two candidates investigated. Nevertheless, this research has contributed, on its own level, to a better understanding of the studied phage and leaves a solid base for further investigation, including the identification of its depolymerase, whose identity remains an unresolved mystery.

Supplementary Figure

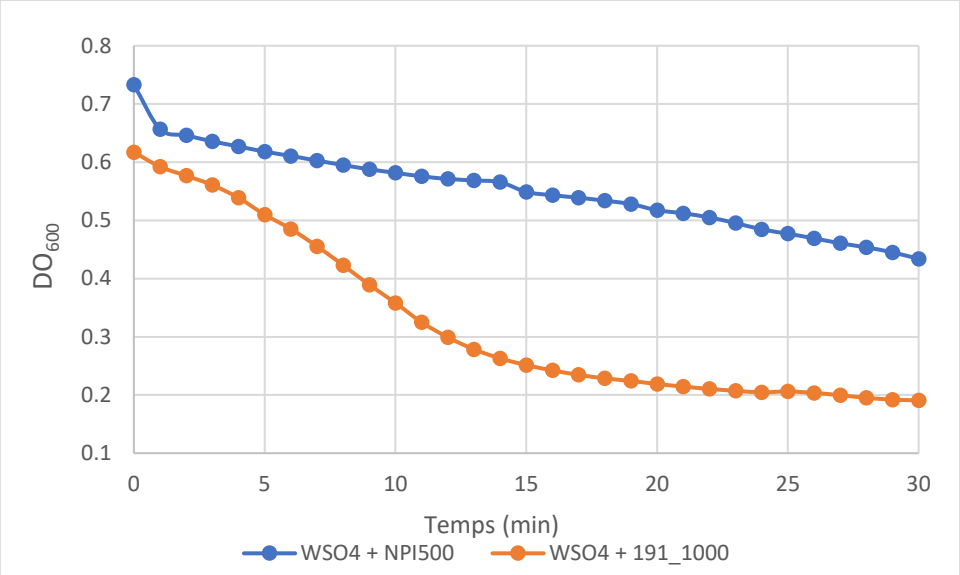

Figure S1. Turbidity reduction assay of gp191_1000 on WS10204 conducted for 30 min. Protein concentration tested are ca. 0.5 mg/mL.

References

Abedon, S.T. (2011). Lysis from without. Bacteriophage 1, 46–49.

Ackermann, H.-W., and Prangishvili, D. (2012). Prokaryote viruses studied by electron microscopy. Arch Virol 157, 1843–1849.

Adriaenssens, E.M., Lehman, S.M., Vandersteegen, K., Vandenheuvel, D., Philippe, D.L., Cornelissen, A., Clokie, M.R.J., García, A.J., De Proft, M., Maes, M., *et al*. (2012). CIM® monolithic anion-exchange chromatography as a useful alternative to CsCl gradient purification of bacteriophage particles. Virology 434, 265–270.

Aiewsakun, P., Adriaenssens, E., Lavigne, R., Kropinski, A.M., Simmonds, P. (2018). Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: steps towards a unified taxonomy. J Gen Virol 99, 1331–1343.

Alcorlo, M., Martínez-Caballero, S., Molina, R., and Hermoso, J.A. (2017). Carbohydrate recognition and lysis by bacterial peptidoglycan hydrolases. Curr Op Struct Biol 44, 87–100.

Andres, D., Hanke, C., Baxa, U., Seul, A., Barbirz, S., and Seckler, R. (2010). Tailspike Interactions with Lipopolysaccharide Effect DNA Ejection from Phage P22 Particles in Vitro. J Biol Chem 285, 36768–36775.

Arnesen, L.P.S., Fagerlund, A., and Granum, P.E. (2008). From soil to gut: *Bacillus cereus* and its food poisoning toxins. FEMS Microbiol Rev 32, 579–606.

Ash, C., Farrow, J.A.E., Dorsch, M., Stackebrandt, E., and Collins, M.D. (1991). Comparative Analysis of *Bacillus anthracis, Bacillus cereus*, and Related Species on the Basis of Reverse Transcriptase Sequencing of 16S rRNA. Int J Syst Bacteriol 41, 343–346.

Baptista, C., Santos, M.A., and São-José, C. (2008). Phage SPP1 Reversible Adsorption to *Bacillus subtilis* Cell Wall Teichoic Acids Accelerates Virus Recognition of Membrane Receptor YueB. J Bacteriol 190, 4989–4996.

Barrett, A.J., and Rawlings, N.D. (2001). Evolutionary Lines of Cysteine Peptidases. Biol Chem 382, 727–734.

Bertozzi Silva, J., Storms, Z., and Sauvageau, D. (2016). Host receptors for bacteriophage adsorption. FEMS Microbiol Lett 363.

Bielmann, R., Habann, M., Eugster, M.R., Lurz, R., Calendar, R., Klumpp, J., and Loessner, M.J. (2015). Receptor binding proteins of *Listeria monocytogenes* bacteriophages A118 and P35 recognize serovar-specific teichoic acids. Virology 477, 110–118.

Boulanger, P. (2009). Purification of bacteriophages and SDS-PAGE analysis of phage structural proteins from ghost particles. Meth Mol Biol 502, 227–238.

Boulanger, P., Jacquot, P., Plançon, L., Chami, M., Engel, A., Parquet, C., Herbeuval, C., and Letellier, L. (2008). Phage T5 Straight Tail Fiber Is a Multifunctional Protein Acting as a Tape Measure and Carrying Fusogenic and Muralytic Activities. J Biol Chem 283, 13556–13564.

Broeker, N.K., Roske, Y., Valleriani, A., Stephan, M.S., Andres, D., Koetz, J., Heinemann, U., and Barbirz, S. (2019). Time-resolved DNA release from an O-antigen–specific *Salmonella* bacteriophage with a contractile tail. J Biol Chem 294, 11751–11761.

Browning, C., Shneider, M.M., Bowman, V.D., Schwarzer, D., and Leiman, P.G. (2012). Phage Pierces the Host Cell Membrane with the Iron-Loaded Spike. Structure 20, 326–339.

Brüssow, H., and Hendrix, R.W. (2002). Phage Genomics: Small Is Beautiful. Cell 108, 13–16.

Campbell, A. (2003). The future of bacteriophage biology. Nature Rev Genet 4, 471–477.

Carlson K. (2005) Appendix: working with bacteriophages: common techniques and methodological approaches in E.M. Kutter, A. Sulakvelidze, Bacteriophages: Biology and Application, 2005, CRC Press, Boca Raton, FL, pp. 437-494

Chibani-Chennoufi, S., Bruttin, A., Dillmann, M.-L., and Brüssow, H. (2004). Phage-Host Interaction: an Ecological Perspective. J Bacteriol 186, 3677–3686.

Cieplak, T., Soffer, N., Sulakvelidze, A., and Nielsen, D.S. (2018). A bacteriophage cocktail targeting *Escherichia coli* reduces *E. coli* in simulated gut conditions, while preserving a non-targeted representative commensal normal microbiota. Gut Microbes 9, 391–399.

Cornelissen, A., Ceyssens, P.-J., T'Syen, J., Praet, H.V., Noben, J.-P., Shaburova, O.V., Krylov, V.N., Volckaert, G., and Lavigne, R. (2011). The T7-Related *Pseudomonas* putida Phage ϕ15 Displays Virion-Associated Biofilm Degradation Properties. PLOS ONE 6, e18597.

Danovaro, R., Corinaldesi, C., Dell'Anno, A., Fuhrman, J.A., Middelburg, J.J., Noble, R.T., and Suttle, C.A. (2011). Marine viruses and global climate change. FEMS Microbiol Rev 35, 993–1034.

Davison, S., Couture-Tosi, E., Candela, T., Mock, M., and Fouet, A. (2005). Identification of the *Bacillus anthracis* γ Phage Receptor. J Bacteriol 187, 6742–6749.

Di Franco, C., Beccari, E., Santini, T., Pisaneschi, G., and Tecce, G. (2002). Colony shape as a genetic trait in the pattern-forming *Bacillus mycoides*. BMC Microbiology 2, 33.

Dierick, K., Van Coillie, E., Swiecicka, I., Meyfroidt, G., Devlieger, H., Meulemans, A., Hoedemaekers, G., Fourie, L., Heyndrickx, M., and Mahillon, J. (2005). Fatal Family Outbreak of *Bacillus cereus*-Associated Food Poisoning. J Clin Microbiol 43, 4277–4279.

Drulis-Kawa, Z., and Maciejewska, G.M.-S. and B. (2015). Bacteriophages and Phage-Derived Proteins – Application Approaches.

Dublanchet, A., and Fruciano, E. (2008). Brève histoire de la phagothérapie. Médecine et Maladies Infectieuses 38, 415–420.

Ehling-Schulz, M., Fricker, M., and Scherer, S. (2004). *Bacillus cereus*, the causative agent of an emetic type of food-borne illness. Mol Nutr Food Res 48, 479–487.

Ehling-Schulz, M., Fricker, M., Grallert, H., Rieck, P., Wagner, M., and Scherer, S. (2006). Cereulide synthetase gene cluster from emetic *Bacillus cereus*: Structure and location on a mega virulence plasmid related to *Bacillus anthracis* toxin plasmid pXO1. BMC Microbiol 6, 20.

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., *et al*. (1976). Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. Nature 260, 500–507.

Gandon, S. (2016). Why Be Temperate: Lessons from Bacteriophage λ. Trends Microbiol 24, 356–365.

Garcia-Doval, C., Castón, J.R., Luque, D., Granell, M., Otero, J.M., Llamas-Saiz, A.L., Renouard, M., Boulanger, P., and Van Raaij, M.J. (2015). Structure of the Receptor-Binding Carboxy-Terminal Domain of the Bacteriophage T5 L-Shaped Tail Fibre with and without Its Intra-Molecular Chaperone. Viruses 7, 6424–6440.

Ge, P., Scholl, D., Prokhorov, N.S., Avaylon, J., Shneider, M.M., Browning, C., Buth, S.A., Plattner, M., Chakraborty, U., Ding, K., *et al*. (2020). Action of a minimal contractile bactericidal nanomachine. Nature 580, 658–662.

Gillis, A., and Mahillon, J. (2014). Phages Preying on *Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis*: Past, Present and Future. Viruses 6, 2623–2672.

Glonti, T., Chanishvili, N., and Taylor, P.W. (2010). Bacteriophage-derived enzyme that depolymerizes the alginic acid capsule associated with cystic fibrosis isolates of *Pseudomonas aeruginosa*. J Appl Microbiol 108, 695–702.

Guinebretière, M.-H., Auger, S., Galleron, N., Contzen, M., De Sarrau, B., De Buyser, M.-L., Lamberet, G., Fagerlund, A., Granum, P.E., Lereclus, D., *et al*. (2013). *Bacillus cytotoxicus sp*. nov. is a novel thermotolerant species of the *Bacillus cereus* Group occasionally associated with food poisoning. Int J Syst Evolut Microbiol, 63, 31–40.

Habann, M., Leiman, P.G., Vandersteegen, K., Bossche, A.V. den, Lavigne, R., Shneider, M.M., Bielmann, R., Eugster, M.R., Loessner, M.J., and Klumpp, J. (2014). *Listeria* phage A511, a model for the contractile tail machineries of SPO1-related bacteriophages. Mol Microbiol 92, 84–99.

Hatfull, G.F. (2015). Dark Matter of the Biosphere: the Amazing World of Bacteriophage Diversity. J Virol 89, 8107–8110.

Hershey, A.D., and Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage. J Gen Physiol 36, 39–56.

Hock, L., Gillis, A., and Mahillon, J. (2016). Complete Genome Sequence of Bacteriophage Deep-Blue Infecting Emetic *Bacillus cereus*. Genome Announc. 4.

Hoffmaster, A.R., Novak, R.T., Marston, C.K., Gee, J.E., Helsel, L., Pruckler, J.M., and Wilkins, P.P. (2008). Genetic diversity of clinical isolates of *Bacillus cereus using* multilocus sequence typing. BMC Microbiol 8, 191.

Holmes, E.C. (2011). What Does Virus Evolution Tell Us about Virus Origins? ▿ . J Virol 85, 5247–5251.

Hoton, F.M., Andrup, L., Swiecicka, I., and Mahillon, J. (2005). The cereulide genetic determinants of emetic *Bacillus cereus* are plasmid-borne. Microbiology, 151, 2121–2124.

Hu, B., Margolin, W., Molineux, I.J., and Liu, J. (2013). The Bacteriophage T7 Virion Undergoes Extensive Structural Remodeling During Infection. Science 339.

Hugouvieux-Cotte-Pattat, N., Condemine, G., and Shevchik, V.E. (2014). Bacterial pectate lyases, structural and functional diversity. Env Microbiol Rep 6, 427–440.

Hyman, P., and van Raaij, M. (2018). Bacteriophage T4 long tail fiber domains. Biophys Rev 10, 463–471.

Jenkins, J., Mayans, O., and Pickersgill, R. (1998). Structure and evolution of parallel beta-helix proteins. J Struct Biol 122, 236–246.

Jensen, G.B., Hansen, B.M., Eilenberg, J., and Mahillon, J. (2003). The hidden lifestyles of *Bacillus cereus* and relatives. Env Microbiol 5, 631–640.

Jorgenson, M.A., Chen, Y., Yahashiri, A., Popham, D.L., and Weiss, D.S. (2014). The bacterial septal ring protein RlpA is a lytic transglycosylase that contributes to rod shape and daughter cell separation in *Pseudomonas aeruginosa*. Mol Microbiol 93, 113–128.

Kawamura-Sato, K., Hirama, Y., Agata, N., Ito, H., Torii, K., Takeno, A., Hasegawa, T., Shimomura, Y., and Ohta, M. (2005). Quantitative Analysis of Cereulide, an Emetic Toxin of *Bacillus cereus*, by Using Rat Liver Mitochondria. Microbiol Imm 49, 25–30.

Keary, R., McAuliffe, O., Ross, R.P., Hill, C., O'Mahony, J., and Coffey, A. (2014). Genome analysis of the staphylococcal temperate phage DW2 and functional studies on the endolysin and tail hydrolase. Bacteriophage 4, e28451.

Keen, E.C. (2015). A century of phage research: Bacteriophages and the shaping of modern biology. Bioessays 37, 6–9.

Kim, B., Kim, E.S., Yoo, Y.-J., Bae, H.-W., Chung, I.-Y., and Cho, Y.-H. (2019). Phage-Derived Antibacterials: Harnessing the Simplicity, Plasticity, and Diversity of Phages. Viruses 11, 268.

Kimura, K., and Itoh, Y. (2003). Characterization of Poly-γ-Glutamate Hydrolase Encoded by a Bacteriophage Genome: Possible Role in Phage Infection of *Bacillus subtilis* Encapsulated with Poly-γ-Glutamate. Appl Environ Microbiol 69, 2491–2497.

Knecht, L.E., Veljkovic, M., and Fieseler, L. (2020). Diversity and Function of Phage Encoded Depolymerases. Front Microbiol 10:2949.

Koskella, B., and Meaden, S. (2013). Understanding Bacteriophage Specificity in Natural Microbial Communities. Viruses 5, 806–823.

Kretzer, J.W., Lehmann, R., Schmelcher, M., Banz, M., Kim, K.-P., Korn, C., and Loessner, M.J. (2007). Use of High-Affinity Cell Wall-Binding Domains of Bacteriophage Endolysins for Immobilization and Separation of Bacterial Cells. Appl Environ Microbiol 73, 1992–2000.

Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. Nature Rev Microbiol 8, 317–327.

Latka, A., Maciejewska, B., Majkowska-Skrobek, G., Briers, Y., and Drulis-Kawa, Z. (2017). Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. Appl Microbiol Biotechnol 101, 3103–3119.

Lechner, S., Mayr, R., Francis, K.P., Prüss, B.M., Kaplan, T., Wiessner-Gunkel, E., Stewart, G.S., and Scherer, S. (1998). *Bacillus weihenstephanensis sp*. nov. is a new psychrotolerant species of the *Bacillus cereus* group. Int J Syst Bacteriol 48, 1373–1382.

Lee, J.-H., Shin, H., and Ryu, S. (2014). Characterization and comparative genomic analysis of bacteriophages infecting members of the *Bacillus cereus* group. Arch Virol 159, 871–884.

Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., Sauder, J.M., Pukatzki, S., Burley, S.K., Almo, S.C., and Mekalanos, J.J. (2009). Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. Proc Natl Acad Sc 106, 4154–4159.

Letarov, A.V., and Kulikov, E.E. (2017). Adsorption of bacteriophages on bacterial cells. Biochemistry Moscow 82, 1632–1658.

Li, X., Koç, C., Kühner, P., Stierhof, Y.-D., Krismer, B., Enright, M.C., Penadés, J.R., Wolz, C., Stehle, T., Cambillau, C., *et al*. (2016). An essential role for the baseplate protein Gp45 in phage adsorption to *Staphylococcus aureus*. Sci Rep 6, 26455.

Lin, D.M., Koskella, B., and Lin, H.C. (2017). Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. World J Gastrointest Pharmacol Ther 8, 162–173.

Liu, Y., Lai, Q., Göker, M., Meier-Kolthoff, J.P., Wang, M., Sun, Y., Wang, L., and Shao, Z. (2015). Genomic insights into the taxonomic status of the *Bacillus cereus* group. Sci Rep 5, 14082.

Locus Biosciences initiates world's first controlled clinical trial for a CRISPR enhanced bacteriophage therapy (2020) Locus Biosciences.

Mahler, H., Pasi, A., Kramer, J.M., Schulte, P., Scoging, A.C., Bär, W., and Krähenbühl, S. (1997). Fulminant Liver Failure in Association with the Emetic Toxin of *Bacillus cereus*. New Engl J Med 336, 1142–1148.

Marraffini, L.A. (2015). CRISPR-Cas immunity in prokaryotes. Nature 526, 55–61.

Marvasi, M., Visscher, P.T., and Martinez, L.C. (2010). Exopolymeric substances (EPS) from *Bacillus subtilis* : polymers and genes encoding their synthesis. FEMS Microbiol Lett 313, 1–9.

Maughan, H., and Van der Auwera, G. (2011). *Bacillus* taxonomy in the genomic era finds phenotypes to be essential though often misleading. Inf Genet Evol 11, 789–797.

Mi, L., Liu, Y., Wang, C., He, T., Gao, S., Xing, S., Huang, Y., Fan, H., Zhang, X., Yu, W., *et al*. (2019). Identification of a lytic *Pseudomonas aeruginosa* phage depolymerase and its anti-biofilm effect and bactericidal contribution to serum. Virus Genes 55, 394–405.

Moak, M., and Molineux, I.J. (2004). Peptidoglycan hydrolytic activities associated with bacteriophage virions. Mol Microbiol 51, 1169–1183.

Müller, J.J., Barbirz, S., Heinle, K., Freiberg, A., Seckler, R., and Heinemann, U. (2008). An Intersubunit Active Site between Supercoiled Parallel β Helices in the Trimeric Tailspike Endorhamnosidase of *Shigella flexneri* Phage Sf6. Structure 16, 766–775.

Nambu, T., Minamino, T., Macnab, R.M., and Kutsukake, K. (1999). Peptidoglycan-Hydrolyzing Activity of the FlgJ Protein, Essential for Flagellar Rod Formation in *Salmonella typhimurium*. J Bacteriol 181, 1555–1561.

Nasukawa, T., Uchiyama, J., Taharaguchi, S., Ota, S., Ujihara, T., Matsuzaki, S., Murakami, H., Mizukami, K., and Sakaguchi, M. (2017). Virus purification by CsCl density gradient using general centrifugation. Arch Virol 162, 3523–3528.

Nobrega, F.L., Vlot, M., de Jonge, P.A., Dreesens, L.L., Beaumont, H.J.E., Lavigne, R., Dutilh, B.E., and Brouns, S.J.J. (2018). Targeting mechanisms of tailed bacteriophages. Nat Rev Microbiol 16, 760–773.

Owusu-Kwarteng, J., Wuni, A., Akabanda, F., Tano-Debrah, K., and Jespersen, L. (2017). Prevalence, virulence factor genes and antibiotic resistance of *Bacillus cereus sensu lato* isolated from dairy farms and traditional dairy products. BMC Microbiol 17, 65.

Pathria, S., Rolando, M., Lieman, K., Hayes, S., Hardies, S., and Serwer, P. (2012). Islands of non-essential genes, including a DNA translocation operon, in the genome of bacteriophage 0305φ8-36. Bacteriophage 2, 25–35.

Paul, V.D., Rajagopalan, S.S., Sundarrajan, S., George, S.E., Asrani, J.Y., Pillai, R., Chikkamadaiah, R., Durgaiah, M., Sriram, B., and Padmanabhan, S. (2011). A novel bacteriophage Tail-Associated Muralytic Enzyme (TAME) from Phage K and its development into a potent antistaphylococcal protein. BMC Microbiol 11, 226.

Pires, D.P., Oliveira, H., Melo, L.D.R., Sillankorva, S., and Azeredo, J. (2016). Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. Appl Microbiol Biotechnol 100, 2141–2151.

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

Plaut, R.D., Beaber, J.W., Zemansky, J., Kaur, A.P., George, M., Biswas, B., Henry, M., Bishop-Lilly, K.A., Mokashi, V., Hannah, R.M., *et al*. (2014). Genetic Evidence for the Involvement of the S-Layer Protein Gene sap and the Sporulation Genes spo0A, spo0B, and spo0F in Phage AP50c Infection of *Bacillus anthracis*. J Bacteriol 196, 1143–1154.

Porter, N.T., and Martens, E.C. (2017). The Critical Roles of Polysaccharides in Gut Microbial Ecology and Physiology. Annu Rev Microbiol 71, 349–369.

Priest, F.G., Barker, M., Baillie, L.W.J., Holmes, E.C., and Maiden, M.C.J. (2004). Population structure and evolution of the *Bacillus cereus* group. J Bacteriol 186, 7959–7970.

Rakhuba, D.V., Kolomiets, E.I., Dey, E.S., and Novik, G.I. (2010). Bacteriophage Receptors, Mechanisms of Phage Adsorption and Penetration into Host Cell. Pol J Microbiol 59, 145–155.

Rodríguez-Rubio, L., Martínez, B., Donovan, D.M., García, P., and Rodríguez, A. (2013). Potential of the Virion-Associated Peptidoglycan Hydrolase HydH5 and Its Derivative Fusion Proteins in Milk Biopreservation. PLOS ONE 8, e54828.

Rydman, P.S., and Bamford, D.H. (2002). The Lytic Enzyme of Bacteriophage PRD1 Is Associated with the Viral Membrane. J Bacteriol 184, 104–110.

Sanahuja, G., Banakar, R., Twyman, R.M., Capell, T., and Christou, P. (2011). *Bacillus thuringiensis:* a century of research, development and commercial applications: A century of Bacillus thuringiensis. Plant Biotechnol J 9, 283–300.

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchison, C.A., Slocombe, P.M., and Smith, M. (1977). Nucleotide sequence of bacteriophage ϕX174 DNA. Nature 265, 687–695.

Schmelcher, M., and Loessner, M.J. (2014). Application of bacteriophages for detection of foodborne pathogens. Bacteriophage 4, e28137.

Schulz, E.C., Dickmanns, A., Urlaub, H., Schmitt, A., Mühlenhoff, M., Stummeyer, K., Schwarzer, D., Gerardy-Schahn, R., and Ficner, R. (2010). Crystal structure of an intramolecular chaperone mediating triple–β-helix folding. Nature Str Mol Biol 17, 210–215.

Serwer, P., Graef, P.R., and Garrison, P.N. (1978). Use of ethidium bromide fluorescence enhancement to detect duplex DNA and DNA bacteriophages during zone sedimentation in

sucrose gradients: molecular weight of DNA as a function of sedimentation rate. Biochem 17, 1166–1170.

Sharma, S., Chatterjee, S., Datta, S., Prasad, R., Dubey, D., Prasad, R.K., and Vairale, M.G. (2017). Bacteriophages and its applications: an overview. Folia Microbiol 62, 17–55.

Spinelli, S., Veesler, D., Bebeacua, C., and Cambillau, C. (2014). Structures and host-adhesion mechanisms of lactococcal siphophages. Front Microbiol 5, 2014.03.

Stern, A., and Sorek, R. (2011). The phage-host arms race: Shaping the evolution of microbes. BioEssays 33, 43–51.

Stewart, Jassim, Denyer, Newby, Linley, and Dhir (1998). The specific and sensitive detection of bacterial pathogens within 4 h using bacteriophage amplification. J Appl Microbiol 84, 777–783.

Stockdale, S.R., Mahony, J., Courtin, P., Chapot-Chartier, M.-P., Pijkeren, J.-P. van, Britton, R.A., Neve, H., Heller, K.J., Aideh, B., Vogensen, F.K., *et al*. (2013). The Lactococcal Phages Tuc2009 and TP901-1 Incorporate Two Alternate Forms of Their Tail Fiber into Their Virions for Infection Specialization. J Biol Chem 288, 5581–5590.

Storms, Z.J., and Sauvageau, D. (2015). Modeling tailed bacteriophage adsorption: Insight into mechanisms. Virology 485, 355–362.

Tang, J., Lander, G.C., Olia, A., Li, R., Casjens, S., Prevelige, P., Cingolani, G., Baker, T.S., and Johnson, J.E. (2011). Peering Down the Barrel of a Bacteriophage Portal: The Genome Packaging and Release Valve in P22. Structure 19, 496–502.

Tomlinson, S., and Taylor, P.W. (1985). Neuraminidase associated with coliphage E that specifically depolymerizes the *Escherichia coli* K1 capsular polysaccharide. J Virol 55, 374–378.

Twort, F.W. (1915). An investigation on the nature of ultra-microscopic viruses. The Lancet 186, 1241–1243.

Van der Auwera, G.A., Timmery, S., Hoton, F., and Mahillon, J. (2007). Plasmid exchanges among members of the *Bacillus cereus* group in foodstuffs. Int J Food Microbiol 113, 164–172.

Vitiello, C.L., Merril, C.R., and Adhya, S. (2005). An amino acid substitution in a capsid protein enhances phage survival in mouse circulatory system more than a 1000-fold. Virus Res 114, 101–103.

Vollmer, W., Blanot, D., and De Pedro, M.A. (2008). Peptidoglycan structure and architecture. FEMS Microbiol Rev 32, 149–167.

Willford, J.D., Bisha, B., Bolenbaugh, K.E., and Goodridge, L.D. (2011). Luminescence based enzyme-labeled phage (Phazyme) assays for rapid detection of Shiga toxin producing *Escherichia coli* serogroups. Bacteriophage 1, 101–110.

Wong, T.Y., Preston, L.A., and Schiller, N.L. (2000). Alginate Lyase: Review of Major Sources and Enzyme Characteristics, Structure-Function Analysis, Biological Roles, and Applications. Annu Rev Microbiol 54, 289–340.

Xu, Q., Abdubek, P., Astakhova, T., Axelrod, H.L., Bakolitsa, C., Cai, X., Carlton, D., Chen, C., Chiu, H.-J., Chiu, M., *et al*. (2010). Structure of the γ-d-glutamyl-l-diamino acid endopeptidase YkfC from *Bacillus cereus* in complex with l-Ala-γ-d-Glu: insights into substrate recognition by NlpC/P60 cysteine peptidases. Acta Cryst F 66, 1354–1364.

Yap, M.L., and Rossmann, M.G. (2014). Structure and function of bacteriophage T4. Future Microbiol 9, 1319–1327.

Young, R. (2014). Phage lysis: Three steps, three choices, one outcome. J Microbiol 52, 243–258.

Zivanovic, Y., Confalonieri, F., Ponchon, L., Lurz, R., Chami, M., Flayhan, A., Renouard, M., Huet, A., Decottignies, P., Davidson, A.R., *et al*. (2014). Insights into Bacteriophage T5 Structure from Analysis of Its Morphogenesis Genes and Protein Components. J Virol 88, 1162–1174.