Exploiting Bacteriophage Proteomes
The Hidden Biotechnological Potential

Santos, Sílvio B.; Costa, Ana Rita; Carvalho, Carla; Nóbrega, Franklin L.; Azeredo, Joana

DOI
10.1016/j.tibtech.2018.04.006

Publication date
2018

Document Version
Accepted author manuscript

Published in
Trends in Biotechnology

Exploiting bacteriophage proteomes: the hidden biotechnological potential

Sílvio B. Santos, Ana Rita Costa, Carla Carvalho[#], Franklin L. Nóbrega[#], Joana Azeredo[*]

Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal

* Correspondence: jazeredo@deb.uminho.pt (J. Azeredo)
Tel. + 351 253 604 419 Fax. + 351 253 604 429
Present addresses: Carla Carvalho, International Iberian Nanotechnology Laboratory (INL), Braga, Portugal; Franklin L. Nóbrega, Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands
https://www.ceb.uminho.pt/bbig; https://www.ceb.uminho.pt;
https://www.facebook.com/cebuminho

Abstract

Bacteriophages encode many distinct proteins for the successful infection of a bacterial host. Each protein plays a specific role in the phage replication cycle, from host recognition, through takeover of the host machinery, and up to cell lysis for progeny release. As the roles of these proteins are being revealed, more biotechnological applications can be anticipated. Phage-encoded proteins are now being explored for the control, detection and typing of bacteria; as vehicles for drug delivery; and for vaccine development. In this review we discuss how engineering approaches can be used to improve the natural properties of these proteins, and set forth the most innovative applications that demonstrate the unlimited biotechnological potential held by phage-encoded proteins.

Keywords: Phage-encoded proteins; genetic engineering; host specificity; bacteriolytic activity; bacteria control and detection; anti-CRISPR.

Bacteriophage research reloaded

Bacteriophage research is being driven by the global threat of antibiotic resistance, resulting in an increasing wealth of knowledge on phage genes and proteins. Simultaneously, recent progress in sequencing technologies, DNA manipulation and **synthetic biology** approaches has been fostering phage **proteome** exploitation and engineering of specific phage proteins into improved forms. Innovative research on phage-encoded proteins is thus now progressing quickly.

Having been a central part of molecular biology for many years, phage-encoded proteins are now being explored in health, industrial, food, and agricultural settings, for purposes not limited to bacteria control. Many applications have been envisioned and this is only considering the low percentage of phage proteins of known function. Although some of the applications found for phage-encoded proteins may be performed by the phage itself, the use of phage proteins instead may have strong advantages in terms of regulation and public acceptance, and also in manipulating them for improved properties.

Phages have spent billions of years evolving and developing a powerful protein armamentarium to recognize, infect and kill bacteria in a very efficient way. Understanding the phage replication cycle is key to identify the proteins involved (**Box 1**, **Figure 1, Key Figure**), to discern their specific function, and thus to unveil the potential held for biotechnology. The particular applications in which phage proteins can be employed are defined by their intrinsic properties, the technologies at our disposal, and our creativity.

Here we overview the most recent progress reported on the use of phage-encoded proteins and highlight their most innovative uses, showcasing the virtually unlimited biotechnological opportunities hidden in bacteriophage genomes.

Receptor binding proteins

Specificity is a fundamental aspect of phage-host interaction and depends upon the phage receptor binding proteins (RBPs). These highly variable structures are part of the

phage particle (**Figures 1A, 1C and 2A**) and make the first contact with the host, being responsible for recognizing specific receptors on the cell surface. Therefore, RBPs are powerful tools for specific pathogen detection (**Figure 2**), and more recently have shown potential in diagnostics and therapy.

RBPs have been successfully employed for pathogen detection in food in a rapid, sensitive and specific manner using methods as RBP-based magnetic separation combined with PCR [1], Enzyme-Linked Long Tail Fiber Assay (ELLTA) employing RBP-coated paramagnetic beads [2], and RBP-coated long-period gratings [3] and interdigital capacitors [4]. These proteins overcome some of the limitations of antibodies that hamper their use in *in situ* applications, *e.g.* pH, temperature and protease sensitivity, while exhibiting comparable or even superior specificity and affinity. RBP-based detection systems can already be found in the market: bioMérieux commercializes the *Vidas Up* kit for the *in situ* detection of foodborne pathogens.

RBPs are specific enough that they can distinguish **glucosylation** variants of O-antigens when classical methods fail to do so [5]. This is relevant for understanding bacterial **immunogenicity** and spread of disease, and may also be employed for monitoring phase variations during large scale O-antigen generation for vaccine production [5]. In fact, the binding affinity and specificity of RBPs have also found application in bacteria-based processes of therapeutic protein production, for the detection and removal of bacterial endotoxin contaminants [6]. One product with this purpose is commercialized by Hyglos GmbH (http://www.hyglos.de). Alternatively, the specificity of some RBPs for the bacterial **lipopolysaccharide** (LPS) may be explored for modulating and counteracting the effects of LPS-induced inflammatory response *in vivo*[7]. While natural RBPs have proven useful, progress in synthetic biology has created new opportunities for the design of recombinant phage-derived proteins with enhanced properties and novel applications. For example, introduction of specific tags like a Cys-tag [8] to the RBPs originated an oriented immobilization of the proteins onto surfaces, significantly improving their capture efficiency and performance as detection probes. Also, deleting the endorhamnosidase enzymatic activity of a tail spike, a phage RBP (**Figure 1A and 1C**), led to improved performance of the protein as a detection probe [8]. This enzymatic activity is undesirable for diagnostic (and capture) purposes

because it causes hydrolysis of bacterial LPS preventing “irreversible” binding of the RBP to the bacterial surface. Functional analysis of RBPs identified the C-terminal domain as responsible for recognition and binding to the host receptor [9]. These C-terminal domains can thus also be used to develop detection tools (**Figure 2**). This can be advantageous for recombinant expression and application in diagnosis providing a broader host spectrum and higher specificity and sensitivity in the detection [10]. Furthermore, RBPs both in the intact and truncated versions can be engineered to broaden, narrow or acquire new specificities for different targets, by substituting one or a few amino acids allowing a wider range of application [11].

Because RBPs are highly diverse, *in silico* identification may be difficult. Functional analysis is thus usually required which can be time consuming and limit the identification of novel RBPs. Moreover, the structural nature and **multimerization** of RBPs often requires alternative cloning and expression methodologies (*e.g.* inclusion of chaperones) to avoid insoluble and misfolded proteins. These are still a limiting step to the development of novel RBP-based biotechnological applications.

Due to the inherent high specificity and affinity of RBPs, most of their applications are being directed toward pathogen detection, whereas not much has been published for pathogen control. Still, a few studies have explored the use of RBPs with enzymatic activity to reduce bacterial colonization *in vivo* [12]. Given the specificity of RBPs, it is particularly worth hypothesizing their usefulness in the design of targeted and tailor-made antimicrobials by fusing RBPs with unspecific drugs. Also, fusion with a peptide able to induce phagocytosis and/or agglutination could improve the immune response against specific problematic pathogenic bacteria targeted by the RBP.

Depolymerases

Some bacteria have developed a capsular structure, usually composed of **polysaccharides**, intimately associated to the cell surface. The capsule provides multiple advantages to bacteria, like protection against host immunity, antibiotics and desiccation, and increased adherence to host cells and surfaces, thus playing an important role in virulence [13]. Moreover, capsules also protect bacterial cells from

phage infection [14]. Nevertheless some phages have evolved to use the capsule as an adsorption receptor and to degrade its capsular polysaccharides (CPS) [15]. Degradation of the CPS allows phages to penetrate the capsule and gain access to the receptor on the outer membrane of the cell for DNA ejection [16]. These phage-encoded enzymes, known as depolymerases, are typically present as part of the phage structure (*e.g.* as part of RBPs) [17], but may also be in a free form diffused in the medium (*i.e.* depolymerases encoded in the phage genome that are not part of the phage particle, and are released during host cell lysis). Specificity is their main feature, digesting only certain types of polysaccharides [15, 18].

The properties of depolymerases anticipate their use for multiple purposes. Since depolymerases deprive bacteria of their capsule, they reduce bacterial virulence and render the cells sensitive to host defenses such as **phagocytosis**. The enzymes may thus be employed as **adjuvants** of the host immune system, as proven both *in vitro* [15] and *in vivo* with high rates of animal survival [15, 19]. Additionally, released polysaccharides can be used as **immunogens** for **glycoconjugated vaccine** production [20].

There is also growing interest in the use of depolymerases as anti-biofilm agents with applications in health and industrial sectors: most phages infecting exopolysaccharide (EPS)-producing bacteria have depolymerases, and some of them are able to disrupt the **biofilm** by degrading the EPS [21]. By disruption and dispersal of the biofilm matrix, the enzymes can help the activity of antibiotics [22], disinfectants [23] or even other phages against the typically highly resistant biofilms, facilitating the penetration of these agents across the biofilm. Depolymerases were also observed to prevent biofilm formation [24], suggesting the use of depolymerases for surface coating to avoid bacterial colonization.

The main feature of depolymerases, which is their specificity, may constitute a limitation to control bacteria presenting different capsule types. This can be overcome by engineering approaches, or simply by using enzyme cocktails targeting different CPS/EPS. Still, the extraordinary specificity of depolymerases for capsular types can be harnessed for diagnosis and typing applications [15] and as a tool in the rapidly growing field of **glycobiology**, *e.g.* for determining glycan profiles [25].

Endolysins

Double-stranded DNA (dsDNA) phages have evolved a lytic system for the release of newly formed virions trapped inside the bacterial cell after replication, mostly based on two proteins: endolysin and holin. Endolysins are peptidoglycan hydrolases that accumulate in the cytoplasm of the host cell until the holin forms pores in the plasma membrane, giving access to endolysins to degrade the peptidoglycan of the cell wall and cause cell lysis [26].

The inherent bacteriolytic activity of endolysins immediately suggested their antimicrobial potential, which was promptly supported by initial findings on the enzyme´s ability to lyse Gram-positive bacteria when added exogenously [27]. The high potential to control Gram-positive bacteria, coupled with a high specificity, lack of known toxicity, and unlikeliness of development of bacterial resistance [28], has made endolysins the most explored phage-encoded proteins so far. Research has proven the capacity of these enzymes to control localized [29, 30] and systemic [29] infections of Gram-positive pathogenic bacteria both *in vitro* and *in vivo* [31, 32]; to efficiently remove recalcitrant biofilms [33]; to prevent foodborne diseases by controlling for example milk [34] and fruit pathogens [35] contributing thus for food safety; and also as disinfectants on surfaces and equipment [34]. Endolysins can also be used in combination with other treatments, such as **high hydrostatic pressure processing** to control pathogens in low processed, ready-to-eat food products [36].

The modular structure of endolysins (**Figure 3**) targeting Gram-positive bacteria, composed of an enzymatic catalytic domain (ECD) and a cell wall binding domain (CBD) connected by a linker [37], rapidly suggested the opportunity to engineer new endolysins with improved properties by combining different domains. Indeed, swap or combination of CBD domains has proven successful both for expanding the bacteriolytic spectrum [38], improving the anti-biofilm properties [39], and for increasing the activity [40], stability and solubility [41] of endolysins. It is thus foreseeable that engineering approaches will allow for the creation of **chimeric** endolysins targeting any and all desired Gram-positive bacteria.

The success of endolysin engineering is evident, with some formulations currently in clinical trials or reaching the market. Staphefekt SA.100 is an engineered phage endolysin to treat methicillin-sensitive and methicillin-resistant *S. aureus* skin infections [30], commercialized by Micreos and available in Europe in cream or gel formulations (https://www.gladskin.com/en/). Another formulation, SAL200, containing the recombinant endolysin SAL-1 against methicillin-resistant *S. aureus* [42], reported no adverse effects when administered intravenously in a first-in-human phase 1 study [43], and was recently successful in the treatment of chronic *S. aureus*-related dermatoses in three patients, with no signs of resistance [30].

The high efficiency of endolysins against Gram-positive bacteria is not observed for Gram-negative cells due to the existence of an outer membrane protecting the peptidoglycan and obstructing the access of endolysins from the outside. Some strategies have been designed to overcome this limitation (**Figure 3**). For example, the combination of endolysins with outer membrane permeabilizers gave them access to the Gram-negative cell peptidoglycan and rapidly reduced cells by several orders of magnitude [44], proving that endolysins are not limited to Gram-positive bacteria. Recently, genetic engineering gave rise to a new generation of lytic enzymes, known as Artilysins®, with improved antibacterial activity on Gram-negative cells [45, 46]. Artilysins are thus engineered proteins composed of an endolysin and an **outer membrane permeabilizing peptide** or **cell penetrating peptide** able to give the endolysin access to the peptidoglycan [45, 46].

Artilysins have also been created to improve the properties of endolysins targeting Gram-positive bacteria, having shown an improved bactericidal activity and reduced dependence on external conditions [47].

Furthermore, although Artilysins have not been thought for such purpose, fusions of endolysins or other lytic phage proteins with cell penetrating peptides may also have applications in the control of intracellular pathogenic bacteria, which are usually refractory to both antibiotics and the immune system. This strategy explores the capacity of cell penetrating peptides to cross epithelial cell membranes.

Beside their obvious application on bacteria control, endolysins have been used also as tags to improve crystallization [48], solubility and purification of recombinant proteins

[49]. In a distinct approach, the ECD domain of endolysins was used as an alternative to sonication or high pressure homogenization for the release and purification of proteins expressed in *Escherichia coli* [50], a more amenable process for large scale protein isolation. **Bacterial ghosts** can also be produced using endolysins for the development of non-living vaccine candidates [51].

The recognition elements of endolysins, CBDs, are highly specific for certain peptidoglycan types. This specificity has been exploited for the construction of simple, rapid, and cost-effective **biosensors** for bacterial detection, *e.g.* in diagnosis, with results superior to those of antibody-based approaches [52, 53]. The use of a CBD as the recognition element conjugated with colloidal gold nanoparticles (that produce a colorimetric signal) in a nitrocellulose-based lateral flow assays a good example of such biosensors [52].Interestingly, CBDs were reported capable of detecting not only bacteria but also their spores, with important applications for food industry, significantly decreasing the detection time [54]. In a distinct approach, CBDs have been used as the targeting element of antimicrobial nanoparticle conjugates for the specific delivery of antimicrobials to pathogenic bacteria [55]. Sharing similar features, *i.e.* high specificity and affinity, CBDs and RBPs have been explored with similar approaches and are expected to have a comparable role in biotechnology (**Figure 2**).

Holins

As a part of the lytic system of dsDNA phages, holins play two fundamental roles: they create holes in the inner membrane for the release of the endolysins, and determine the timing for the end of the infection cycle. Holins accumulate in the inner membrane of the cell with no effect on its integrity until they reach critical concentration that triggers holin activation [56]. Two types of holins have been described: **canonical holins** that form large pores and **pinholins** that form small pores [57].

Holins can cause cell death independent of endolysins and, unlike these, have a broad-spectrum unspecific antibacterial activity against both Gram-positive and Gram-negative bacteria [58]. These features have attracted interest towards the application of holins in bacterial control, although perhaps limited to disinfection of surfaces or

foodstuff due to the lack of specificity [58]. For applications requiring specificity, it may be possible to fuse holins to a peptide that specifically binds to the target bacteria, since fusion has been shown not to affect holin activity [59]. For this purpose, CBDs of endolysins and phage RBPs may be an option.

The combination of holins with endolysins was shown to be a possible approach to control Gram-positive bacteria with higher efficiency than endolysins alone [60]. A similar strategy, fusing holin and endolysin, may also be an option to control Gram-negative pathogens with high efficiency. Holins would form pores on the outer membrane allowing access of the attached endolysins to the peptidoglycan layer of these bacteria. This **hybrid** protein could be a broader-spectrum alternative to the previously described Artilysins, although not evaluated so far.

The lethality of holins is associated to loss of viability due to the holes formed on the cell membrane, and not to cell burst [63]. While canonical holins form holes that can be crossed by proteins or protein complexes up to 500 kDa [64], pinholins form much smaller pores and may thus be an excellent option to target **endotoxin**-containing bacteria, preventing the release of their toxic content.

The biotechnological prospecting of holins is not limited to bacterial control; they have shown promise also in cancer treatment. Gene therapy using cytotoxic proteins to treat cancer is being intensively studied. The ability of holins to form lesions on the bacterial membrane suggested their cytotoxic activity on eukaryotic tumor cells. The expression of a gene encoding the lambda holin inside eukaryotic cells under a tightly controlled expression system substantially reduced cell viability *in vitro* and inhibited tumor growth *in vivo* demonstrating the potential of holins as a new therapeutic protein for cancer gene therapy [65]. Additionally, the combination of holin and endolysin in plasmids presents a successful alternative for the creation of bacterial ghosts of high immunogenicity for the development of non-living vaccine candidates [61, 62, 66].

Holin/endolysin combinations can be used for induction-controlled delivery of antigens into the cytoplasm of mammalian cells. Attenuated bacteria are engineered to possess the desired antigen and invade mammalian cells due to a natural intracellular parasitism. Bacteria then deliver their antigenic cargo by autolysis caused by

expression of the cloned holin/endolysin genes under the control of specific promoters that respond only to the intracellular environment of mammalian cells [67]. This approach may be used for delivering other cargoes, including bacteriophage-based proteins to control intracellular bacterial pathogens and modulate their causative infection.

The controlled expression of a holin/lysin system was also applied for the purification of **minicells**, themselves used as drug/gene-delivery systems [68]. Often minicell batches are contaminated with parent cells, a critical disadvantage for their practical application due to possible endotoxin release in human cells. Induction of the cloned holin/endolysin system will result in autolysins of the parent cells, thus reducing contamination of the final product. This method not only allows for an efficient separation of highly pure minicells but also avoids the inconvenient conventional multi-step purification approach. The holin/endolysin strategies can also be applied as an economic and simple alternative for the release of products from microbial cells in industrial production of **bio-based chemicals** [69]. The production of intracellular compounds requires cell lysis for product recovery, usually achieved by expensive chemical and mechanical cell disruption methods that add further complexity to the downstream purification processes. The use of phage holin/endolysin systems under the control of a green-light regulated promoter significantly enhances intracellular compound release due to an increased fragility of the cell membrane. This system thus provides recovery of cell compounds with minimal contaminants and energy, avoids the use of conventional chemical inducers using a photosynthetic regulated promoter, and increases safety of **genetically modified organisms** (GMOs) since the modified cells are unable to survive under sunlight if accidentally released in the environment.

Further developments on holin applications will depend on progress on the processes of cloning and high-yield expression of these proteins. The lethality of holins to expression cells and their low solubility as membrane proteins are currently a significant limitation.

Structural murein hydrolases

Phages of both Gram-positive and Gram-negative bacteria have been found to employ virion-associated peptidoglycan hydrolases (VAPGHs) at the initial stage of phage infection. These enzymes locally degrade the cell wall peptidoglycan of the host, allowing the phage to eject its genome into the host cell [70]. Whether these enzymes are used in every infection or only under less optimal conditions is still up to debate [71].

Although they are structural enzymes, VAPGHs share some features with endolysins, *i.e.* high substrate specificity, peptidoglycan cleavage mechanisms, and modular structure. This not only suggests an antimicrobial potential similar to endolysins (**Table 1**) [72], but also supports their engineering via domain swapping for improved properties and reduced likelihood of resistant strain development.

One interesting property of VAPGHs is that they tend to exhibit remarkably high thermal stability [73]. This feature suggests potential uses in food technology, where high temperatures are commonly used. This was demonstrated by the high antimicrobial activity achieved in milk pasteurized at 72 ˚C by a CHAP domain (enzymatic motif) of a VAPGH fused to the SH3 domain (binding motif) of lysostaphin [74]. Interestingly, this fusion approach seems to be a common trend in VAPGHs engineering and has consistently broadened the host range of the chimeric enzyme [75] and improved the lytic activity both *in vitro* [75] and *in situ* [76]. The bacteriolytic activity of CHAP-SH3 chimera can be further improved when combined with endolysins [75].

VAPGHs have so far been scarcely explored, but their features similar to endolysins anticipate comparable progress and applications.

Anti-CRISPR proteins

Studies on the mechanisms of defense of bacteria against phages led to the discovery of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated genes (Cas) system (CRISPR-Cas), a prokaryotic version of **adaptive immunity**. In short, when an invading DNA (*e.g.* viral or plasmid) enters a cell containing a CRISPR system, some DNA fragments are captured and incorporated in the CRISPR repeats. When the cell is invaded a second time by the same DNA, the latter is cleaved by the Cas nuclease [77] impairing phage infection.

Six distinct CRISPR-Cas types are currently known, possessing different sets of proteins that enable function [78]. Among these, the Type II CRISPR-Cas system has the advantage of relying on a single protein for function [79]. This protein, Cas9, and its variants have become a potent new tool for targeted mutagenesis and genome editing of all living entities [80]. There are concerns however about gene therapy with Cas9 causing off-target **gene editing** with unwanted side effects, and also about the development of a possible immune reaction against Cas proteins [81].

Recently it was found that some phages are able to counterattack the bacterial CRISPR-Cas system by producing proteins able to block its action: the anti-CRISPR (Acr) proteins [82]. This immediately attracted attention towards the Acr proteins as possible modulators of CRISPR-Cas gene therapy. Those identified so far are able to target the Cas proteins of their phage hosts [83], with a few also targeting the variant SpyCas9, the most used protein for genome editing applications [84]. To control CRISPR-Cas9 gene therapy, the Acr proteins may be delivered a few hours after the Cas9 editing tool, decreasing the off-target gene editing and its unpredicted consequences within cells or tissues [85]. Research on Acr proteins has only just begun, so further understanding and novel fascinating applications are expected to emerge soon.

Other (old) phage-encoded proteins

Phages encode multiple proteins other than those mentioned above, some of which have also found applications in biotechnology. The most notable example is seen in

molecular biology, in which phage-encoded proteins have played a central role for many years. Now, even these old and well-known proteins are finding novel and diverse uses, as summarized in **Box 2**.

Concluding Remarks and Future Perspectives

Here we have discussed the multiple ways in which different phage-encoded proteins have been used for human benefit. From therapy, to bacteria typing and detection, surface disinfection, food decontamination, drug delivery and even vaccine development, the biotechnological potential held by these proteins has been widely demonstrated.

Genetic engineering and biotechnology allowed tailoring of phage proteins for desired properties, leading to further improvements. But the prospects of phage-encoded proteins can be more far-reaching than those achieved so far (see **Outstanding Questions**). Surprising applications are emerging at a fast pace; and this is just considering a small part of the powerful armamentarium phages possess to parasitize bacteria, since only a low percentage of phage genes have a known function.

With the recent available genetic and molecular tools and large datasets of raw sequencing data, research should now center on bioinformatics and functional analysis of phages genes to unveil all possible protein properties, even for those proteins with already known function. For example, a tail tubular protein thought to have only a structural function was recently found also to possess lytic activity with therapeutic potential [86]. Early phage proteins are particularly interesting as they are responsible for hijacking the host machinery to a phage-oriented metabolism. Identification of these proteins and understanding their function is still one of the major challenges of phage research, and knowledge on early phage proteins is key to metabolic manipulation of bacteria with numerous potential biotechnological applications.

In fact, it is plausible to assume that knowledge on the function of most phage genes will arise and that major discoveries are yet to come, some of which undoubtedly undergoing powerful translation into medical, agricultural and industrial biotechnologies.

Box 1. The Bacteriophage Replication Cycle

Bacteriophages are bacterial viruses consisting of a nucleic acid genome enclosed within a proteinaceous coat. Like all viruses, phages are metabolically inert and depend upon infection of a bacterial host for replication.

The beginning: Infection begins with the adsorption of the phage on the host's cell surface, relying on phage proteins that specifically recognize receptor structures. This is the first step defining the range of hosts that can be infected by the phage (host specificity) and may involve the action of phage enzymes able to degrade host membrane structures hiding the cell receptors. After adsorption, the phage ejects its genome into the cell, a process that may also be aided by phage-encoded enzymes able to produce pores in the peptidoglycan layer of the bacteria [87]. Once its genome is ejected into the cell, the phage can adopt distinct replication strategies.

The lytic cycle: After nucleic acid ejection, ***Caudovirales*** phages can assume a lytic replication cycle and expression of phage early genes immediately follows, hijacking the host cellular machinery and redirecting it to phage DNA replication and protein synthesis. Some phage proteins are also involved in the phage genome replication process. After taking control of the cell, the phage genes encoding its structural proteins are expressed as well as all the accessory proteins. The DNA is packaged into the empty heads, and the structural proteins are assembled in a process called maturation [87]**.** The mature phage progeny particles are now able to start a new infection cycle but are trapped inside the host cell. At this moment, late phage proteins, such as holins and endolysins, are produced; these will form pores in the inner membrane and degrade the peptidoglycan leading to cell lysis, and death, for progeny release [88].

The lysogenic cycle: Temperate phages assume a lysogenic cycle characterized by integrating the phage genome into the bacterial chromosome. The so-called prophage is stably replicated in synchrony with the bacterial chromosome, being transmitted to each daughter cell and remaining silent for extended periods. This "dormant" state is maintained while the bacteria (called a lysogen) grows "normally". When exposed to specific stimuli (*e.g.* stressful conditions that cause DNA damage) the prophage is induced, entering a lytic replication cycle and killing the cell as described above [89].

Box 2. Novel Applications for Old Phage-Encoded Proteins

After ejection of a phage genome into a host cell, a fast and complex process occurs towards the takeover of the cell metabolism for the production of new virions; a multitude of proteins and enzymes collaborate for this purpose. Most of these proteins were crucial for the development of molecular biology, and are now finding application in novel techniques. Phage RNA polymerases were recently used to reconstitute *in vitro* genetic circuits for the precise mathematical modelling of biological reactions [90]. The T4 DNA ligase has been employed for the detection of single-nucleotide polymorphisms (SNPs) [91] and to form ligation junctions between DNA segments in circular chromosome conformation capture (4C-seq). 4C-seq is a powerful technique used to study the 3D genome organization in the nuclear space [92]. The exonuclease Red from phage lambda has proven useful for genome modification using techniques of recombineering [93] and multiplex automated genome engineering (MAGE) [94]. More recently it was used to construct single-chain variable fragment antibody libraries [95], and to detect antibiotic residues in foodstuff [96]. Phage scaffolding proteins or procapsids can be used to generate protein-based containers for carrying different cargos. In particular, the procapsid of phage T4 seems to be the most attractive for allowing simultaneous packaging of specific active proteins and DNAs. The *in vitro* packaging of DNA is however limited to linear molecules, a limitation that may be overcome by packing the linear DNA together with a recircularization enzyme, increasing the DNA's biological activity on the target [97]. Capsids of phage P22 have been loaded with contrast agents to increase image contrast in magnetic resonance imaging [98]. Scaffolding proteins and procapsids have also been used for vaccine development [99] and for modulating insulin receptor signaling [100]. The portal protein (DNA packaging motor) of phi29 was recently used for peptide fingerprinting, with suggested application for detection of disease-associated peptide biomarkers [101]. Phage integrases have become valuable tools for precise genome editing using the dual integrase cassette exchange (DICE) system [102] and the recombinase mediated cassette exchange (RCME) system [103], and for the construction of memory genetic logic gates for detecting biological events [104].

Glossary

Adjuvant: substance that enhances the immune response of the body to an antigen; common in vaccines.

Adaptive immunity: component of immunity mediated by lymphocytes, highly specific and adaptable towards a pathogen or toxin, and characterized by immunological memory.

Bacterial ghost: Gram-negative bacterial cell envelope, devoid of all cytoplasmic content but retaining an intact membrane structure and all surface proteins of the original bacteria.

Bacteriophage: virus that specifically infects and replicates within Bacteria and Archaea.

Bio-based chemical: chemical made from substances derived from a biological (living) or renewable source.

Biofilm: community of microorganisms enclosed in a matrix and adhered to biotic or abiotic surfaces, which collaborate closely for survival and persistence.

Biosensor: analytical device that uses a biological component for the detection of a specific analyte, converting a biological response into an electrical signal by a transducer.

Canonical holin: phage protein that forms large pores in the inner membrane of the cell through which endolysins accumulated in the cytoplasm can cross to reach the bacterial peptidoglycan.

Caudovirales: taxonomic order of Virus that consists of three families of bacteriophages with a tail, and which represents 96% of the phages observed at the TEM so far.

Cell penetrating peptide: short peptide able to ubiquitously cross cellular membranes with low toxicity, and transport into the cell a wide variety of biologically active conjugates.

Chimeric: composed of different parts (*e.g.* protein domains) from similar sources.

Endotoxin: toxic heat-stable phospholipid-polysaccharide macromolecule associated with the outer membranes of Gram-negative bacteria, which is released from the cell only upon lysis.

Gene editing: the use of biotechnological techniques to make insertions, deletions or replacements of DNA sequences at specific sites in the genome of an organism or cell.

Genetically modified organism: organism whose genome has been altered using genetic engineering techniques.

Glucosylation: controlled enzymatic modification of a protein by addition of a glucosyl group.

Glycobiology: study of the structure, function, and biology of carbohydrates, molecules relevant in medical, biotechnological and basic research fields.

Glycoconjugated vaccine: vaccines that use carbohydrate antigens chemically coupled to a carrier protein to enhance immunogenicity.

High hydrostatic pressure processing: non-thermal technique for preserving and sterilizing food by subjecting the product to a high level of hydrostatic pressure.

Hybrid: composed of different parts (*e.g.* protein domains) from different sources.

Immunogenicity: ability of a substance to provoke an immune response.

Immunogen: substance that elicits immunogenicity.

Lipopolysaccharide: large molecule consisting of a lipid and a polysaccharide joined by a covalent bond, which can be found in the outer membrane of Gram-negative bacteria.

Lysogenic: bacterium or archaea harboring a temperate bacteriophage as a prophage or plasmid.

Lytic: relating to or causing lysis.

Minicell: small bacterial cell which contains no nuclear material and is unable to grow or divide.

Multimerization: process of assembling multimers of a molecule, in which multimers are aggregates of multiple molecules that are held together with non-covalent bonds.

Outer membrane permeabilizing peptide: a peptide that acts onto the outer membrane of cells making them permeable to other molecules.

Phagocytosis: engulfing and often destruction of microorganisms, other cells or foreign particles by phagocytic cells, *e.g.* macrophages.

Pinholin: holin that forms small pores through which ions move causing depolarization of the cell membrane and consequent activation of a specific type of endolysins anchored to the inner membrane.

Polysaccharide: polymeric carbohydrate molecules composed of long chains of monosaccharide units bound together by glycosidic linkages.

Proteome: entire set of proteins expressed by an organism over its entire life cycle, or at a certain time and under defined conditions.

Synthetic biology: artificial design and engineering of novel biological systems, organisms or devices, for purposes of improving applications for industry or biological/biotechnological research.

Acknowledgements

This study was supported by the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UID/BIO/04469/2013 unit, COMPETE 2020 (POCI-01-0145-FEDER-006684) and the Project PTDC/BBB-BSS/6471/2014 (POCI-01-0145-FEDER-016678). Sílvio B. Santos, Ana Rita Costa and Carla Carvalho were supported by FCT grants SFRH/BPD/75311/2010, SFRH/BPD/94648/2013 and SFRH/BPD/79365/2011, respectively. Franklin L. Nobrega was supported by FCT grant SFRH/BD/86462/2012 and by the Netherlands Organization for Scientific Research (NWO) Veni grant 016.Veni.181.092.

References

1. Poshtiban, S*., et al.* (2013) Phage receptor binding protein-based magnetic enrichment method as an aid for real time PCR detection of foodborne bacteria. *Analyst* 138, 5619-5626
2. Denyes, J.M*., et al.* (2017) Modified bacteriophage S16 long tail fiber proteins for rapid and specific immobilization and detection of Salmonella cells. *Appl Environ Microbiol* 83, e00277-17
3. Brzozowska, E*., et al.* (2016) Label-free Gram-negative bacteria detection using bacteriophage-adhesin-coated long-period gratings. *Biomed Opt Express* 7, 829-840
4. Rydosz, A*., et al.* (2016) A broadband capacitive sensing method for label-free bacterial LPS detection. *Biosens Bioelectron* 75, 328-336
5. Schmidt, A*., et al.* (2016) Bacteriophage tailspike protein based assay to monitor phase variable glucosylations in Salmonella O-antigens. *BMC Microbiol* 16, 207
6. Miller, S*., et al.* Hyglos Invest. Method for detecting and removing endotoxin, US8822641
7. Miernikiewicz, P*., et al.* (2016) T4 Phage tail adhesin gp12 counteracts lps-induced inflammation in vivo. *Front Microbiol* 7, 1112
8. Singh, A*., et al.* (2010) Bacteriophage tailspike proteins as molecular probes for sensitive and selective bacterial detection. *Biosens Bioelectron* 26, 131-138
9. Le, S*., et al.* (2013) Mapping the tail fiber as the receptor binding protein responsible for differential host specificity of Pseudomonas aeruginosa bacteriophages PaP1 and JG004. *PLoS One* 8, e68562
10. Javed, M.A*., et al.* (2013) Bacteriophage receptor binding protein based assays for the simultaneous detection of Campylobacter jejuni and Campylobacter coli. *PLoS One* 8, e69770
11. Uchiyama, J*., et al.* (2011) Improved adsorption of an enterococcus faecalis bacteriophage ΦEF24C with a spontaneous point mutation. *PLoS One* 6, e26648
12. Waseh, S*., et al.* (2010) Orally administered P22 phage tailspike protein reduces salmonella colonization in chickens: prospects of a novel therapy against bacterial infections. *PLoS One* 5, e13904

13. Merino, S. and Tomás, J.M. (2015) Bacterial capsules and evasion of immune responses. *eLS*, 1-10
14. Scholl, D.*, et al.* (2005) Escherichia coli's K1 capsule is a barrier to bacteriophage T7. *Appl Environ Microbiol* 71, 34872-34874
15. Hsieh, P.-F.*, et al.* (2017) Two T7-like bacteriophages, K5-2 and K5-4, each encodes two capsule depolymerases: isolation and functional characterization. *Sci Rep* 7, 4624
16. Rakhuba, D.*, et al.* (2010) Bacteriophage receptors, mechanisms of phage adsorption and penentration into host cell. *Pol J Microbiol* 59, 145-155
17. Pires, D.P.*, et al.* (2016) Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. *Appl Microbiol Biotechnol* 100, 2141-2151
18. Oliveira, H.*, et al.* (2017) Ability of phages to infect Acinetobacter calcoaceticus-Acinetobacter baumannii complex species through acquisition of different pectate lyase depolymerase domains. *Environ Microbiol* 19, 5060-5077
19. Majkowska-Skrobek, G.*, et al.* (2016) Capsule-targeting depolymerase, derived from Klebsiella KP36 phage, as a tool for the development of anti-virulent strategy. *Viruses* 8, 324
20. Lee, I.M.*, et al.* (2017) Structural basis for fragmenting the exopolysaccharide of Acinetobacter baumannii by bacteriophage ΦAB6 tailspike protein. *Sci Rep* 7, 42711
21. Gutiérrez, D.*, et al.* (2015) Role of the pre-neck appendage protein (Dpo7) from phage vB_SepiS-phiIPLA7 as an anti-biofilm agent in Staphylococcal species. *Front Microbiol* 6, 1315
22. Alkawash, M.A.*, et al.* (2006) Alginate lyase enhances antibiotic killing of mucoid Pseudomonas aeruginosa in biofilms. *APMIS* 114, 131-138
23. Tait, K.*, et al.* (2002) The efficacy of bacteriophage as a method of biofilm eradication. *Biofouling* 18, 305-311
24. Guo, Z.*, et al.* (2017) Identification and characterization of Dpo42, a novel depolymerase derived from the Escherichia coli phage vB_EcoM_ECOO78. *Front Microbiol* 8, 1460
25. Dalziel, M.*, et al.* (2014) Emerging principles for the therapeutic exploitation of glycosylation. *Science* 343, 1235681

26. Young, R. (2014) Phage lysis: three steps, three choices, one outcome. *J Microbiol* 52, 243-258
27. Nelson, D.*, et al.* (2001) Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proc Natl Acad Sci USA* 98, 4107-4112
28. Rodríguez-Rubio, L.*, et al.* (2013) The phage lytic proteins from the Staphylococcus aureus bacteriophage vB_SauS-phiIPLA88 display multiple active catalytic domains and do not trigger staphylococcal resistance. *PLoS ONE* 8, e64671
29. Chopra, S.*, et al.* (2016) Potential of combination therapy of endolysin MR-10 and minocycline in treating MRSA induced systemic and localized burn wound infections in mice. *Int J Med Microbiol* 306, 707-716
30. Totté, J.E.E.*, et al.* (2017) Successful treatment of chronic Staphylococcus aureus-related dermatoses with the topical endolysin Staphefekt SA.100: A report of 3 cases. *Case Rep Dermatol* 9, 19-25
31. Shen, Y.*, et al.* (2016) A bacteriophage endolysin that eliminates intracellular streptococci. *Elife* 5, e13152
32. Park, S.*, et al.* (2018) Characterisation of the antibacterial properties of the recombinant phage endolysins AP50-31 and LysB4 as potent bactericidal agents against Bacillus anthracis. *Sci Rep* 8, 18
33. Gutiérrez, D.*, et al.* (2014) Effective removal of Staphylococcal biofilms by the endolysin LysH5. *PLoS ONE* 9, e107307
34. Chang, Y.*, et al.* (2017) Characterization of a novel endolysin LysSA11 and its utility as a potent biocontrol agent against Staphylococcus aureus on food and utensils. *Food Microbiol* 68, 112-120
35. Kim, W.-S.*, et al.* (2004) Expression of bacteriophage φEa1h lysozyme in Escherichia coli and its activity in growth inhibition of Erwinia amylovora. *Microbiology* 150, 2707-2714
36. van Nassau, T.J.*, et al.* (2017) Combination of endolysins and high pressure to inactivate Listeria monocytogenes. *Food Microbiol* 68, 81-88
37. Oliveira, H.*, et al.* (2013) Molecular aspects and comparative genomics of bacteriophage endolysins. *J Virol* 87, 4558-4570

625 38. Dong, Q*., et al.* (2015) Construction of a chimeric lysin Ply187N-V12C with
626 extended lytic activity against staphylococci and streptococci. *Microb Biotechnol* 8,
627 210-220
628 39. Fernández, L*., et al.* (2017) Downregulation of autolysin-encoding genes by phage-
629 derived lytic proteins inhibits biofilm formation in Staphylococcus aureus.
630 *Antimicrob Agent Chemother* 61, e02724-16
631 40. Mao, J*., et al.* (2013) Chimeric Ply187 endolysin kills Staphylococcus aureus more
632 effectively than the parental enzyme. *FEMS Microbiol Lett* 342, 30-36
633 41. Haddad Kashani, H*., et al.* (2017) A novel chimeric endolysin with antibacterial
634 activity against methicillin-resistant Staphylococcus aureus. *Front Cell Infect*
635 *Microbiol* 7, 290
636 42. Jun, S.Y*., et al.* (2013) Antibacterial properties of a pre-formulated recombinant
637 phage endolysin, SAL-1. *Int J Antimicrob Agent* 41, 156-161
638 43. Jun, S.Y*., et al.* (2017) Pharmacokinetics and tolerance of the phage endolysin-
639 based candidate drug SAL200 after a single intravenous administration among
640 healthy volunteers. *Antimicrob Agents Chemother* 61, e02629-02616
641 44. Oliveira, H*., et al.* (2016) Structural and enzymatic characterization of ABgp46, a
642 novel phage endolysin with broad anti-gram-negative bacterial activity. *Front*
643 *Microbiol* 7, 208
644 45. Briers, Y*., et al.* (2014) Art-175 is a highly efficient antibacterial against multidrug-
645 resistant strains and persisters of Pseudomonas aeruginosa. *Antimicrob Agents*
646 *Chemother* 58, 3774-3784
647 46. Defraine, V*., et al.* (2016) Efficacy of artilysin Art-175 against resistant and
648 persistent Acinetobacter baumannii. *Antimicrob Agents Chemother* 60, 3480-3488
649 47. Rodríguez-Rubio, L*., et al.* (2016) 'Artilysation' of endolysin λSa2lys strongly
650 improves its enzymatic and antibacterial activity against streptococci. *Sci Rep* 6,
651 35382
652 48. Yang, H*., et al.* (2015) Structural mechanism of ergosterol regulation by fungal
653 sterol transcription factor Upc2. *Nat Commun* 6, 6129
654 49. Boura, E*., et al.* (2017) Metal ions-binding T4 lysozyme as an intramolecular
655 protein purification tag compatible with X-ray crystallography. *Prot Sci* 26, 1116-
656 1123

50. Joshi, H. and Jain, V. (2017) Novel method to rapidly and efficiently lyse Escherichia coli for the isolation of recombinant protein. *Anal Biochem* 528, 1-6
51. Won, G.*, et al.* (2017) A novel method to generate Salmonella Typhi Ty21a ghosts exploiting the λ phage holin-endolysin system. *Oncotarget* 8, 48186-48195
52. Kong, M.*, et al.* (2017) Lateral flow assay-based bacterial detection using engineered cell wall binding domains of a phage endolysin. *Biosens Bioelectron* 96, 173-177
53. Yu, J.*, et al.* (2016) Sensitive and rapid detection of staphylococcus aureus in milk via cell binding domain of lysin. *Biosens Bioelectron* 77, 366-371
54. Gómez-Torres, N.*, et al.* (2017) Development of a specific fluorescent phage endolysin for in situ detection of Clostridium species associated with cheese spoilage. *Microb Biotechnol* (in press)
55. Kim, D.*, et al.* (2017) Targeted killing of pathogenic bacteria with cell wall binding domain (CBD)-antimicrobial nanoparticle conjugates. *AIChE Annual Meeting* 676a
56. Dewey, J.S.*, et al.* (2010) Micron-scale holes terminate the phage infection cycle. *Proc Natl Acad Sci* 107, 2219-2223
57. Catalão, M.J.*, et al.* (2013) Diversity in bacterial lysis systems: bacteriophages show the way. *FEMS Microbiol Rev* 37, 554-571
58. Song, J.*, et al.* (2016) Identification and characterization of HolGH15: the holin of Staphylococcus aureus bacteriophage GH15. *J Gen Virol* 97, 1272-1281
59. White, R.*, et al.* (2011) Holin triggering in real time. *Proc Natl Acad Sci* 108, 798-803
60. Shi, Y.*, et al.* (2012) Combined antibacterial activity of phage lytic proteins holin and lysin from Streptococcus suis bacteriophage SMP. *Curr Microbiol* 65, 28-34
61. Won, G.*, et al.* (2017) Improved lysis efficiency and immunogenicity of Salmonella ghosts mediated by co-expression of λ phage holin-endolysin and ϕX174 gene E. *Sci Rep* 7, 45139
62. Won, G. and Lee, J.H. (2017) Salmonella Typhimurium, the major causative agent of foodborne illness inactivated by a phage lysis system provides effective protection against lethal challenge by induction of robust cell-mediated immune responses and activation of dendritic cells. *Vet Res* 48, 66

688 63. Farkasovská, J.*, et al.* (2004) Identification of a holin encoded by the Streptomyces
689 aureofaciens phage micro1/6; functional analysis in Escherichia coli system. *Folia*
690 *Microbiol* 49, 679-684
691 64. Wang, I.-N.*, et al.* (2003) Sizing the holin lesion with an endolysin-β-galactosidase
692 fusion. *J Bacteriol* 185, 779-787
693 65. Agu, C.A.*, et al.* (2006) The cytotoxic activity of the bacteriophage λ-holin protein
694 reduces tumour growth rates in mammary cancer cell xenograft models. *J Gene*
695 *Med* 8, 229-241
696 66. Horii, T.*, et al.* (2002) Characterization of a holin (HolNU3–1) in methicillin-
697 resistant Staphylococcus aureus host. *FEMS Immunol Med Microbiol* 34, 307-310
698 67. Kuo, C.-Y.*, et al.* (2009) A stably engineered, suicidal strain of Listeria
699 monocytogenes delivers protein and/or DNA to fully differentiated intestinal
700 epithelial monolayers. *Mol Pharm* 6, 1052-1061
701 68. Park, S.-Y.*, et al.* (2011) A coupling process for improving purity of bacterial
702 minicells by holin/lysin. *J Microbiol Meth* 86, 108-110
703 69. Miyake, K.*, et al.* (2014) A green-light inducible lytic system for cyanobacterial
704 cells. *Biotechnol Biofuel* 7, 56-56
705 70. Latka, A.*, et al.* (2017) Bacteriophage-encoded virion-associated enzymes to
706 overcome the carbohydrate barriers during the infection process. *Appl Microbiol*
707 *Biotechnol* 101, 3103-3119
708 71. Rodríguez-Rubio, L.*, et al.* (2013) The peptidoglycan hydrolase of Staphylococcus
709 aureus bacteriophage φ11 plays a structural role in the viral particle. *Appl Environ*
710 *Microbiol* 79, 6187-6190
711 72. Gutiérrez, D.*, et al.* (2017) Real-time assessment of Staphylococcus aureus biofilm
712 disruption by phage-derived proteins. *Front Microbiol* 8, 1632
713 73. Lavigne, R.*, et al.* (2004) Identification and characterization of a highly
714 thermostable bacteriophage lysozyme. *Cell Mol Life Sci* 61, 2753-2759
715 74. Rodríguez-Rubio, L.*, et al.* (2013) Potential of the virion-associated peptidoglycan
716 hydrolase HydH5 and its derivative fusion proteins in milk biopreservation. *PLoS*
717 *ONE* 8, e54828
718 75. Rodríguez-Rubio, L.*, et al.* (2012) Enhanced staphylolytic activity of the
719 Staphylococcus aureus bacteriophage vB_SauS-phiIPLA88 HydH5 virion-associated

peptidoglycan hydrolase: Fusions, deletions, and synergy with LysH5. *Appl Environ Microbiol* 78, 2241-2248

76. Vipra, A.A*., et al.* (2012) Antistaphylococcal activity of bacteriophage derived chimeric protein P128. *BMC Microbiol* 12, 41
77. Mohanraju, P*., et al.* (2016) Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353, aad5147.
78. Marraffini, L.A. (2015) CRISPR-Cas immunity in prokaryotes. *Nature* 526, 55-61
79. Barrangou, R. and Doudna, J.A. (2016) Applications of CRISPR technologies in research and beyond. *Nat Biotech* 34, 933-941
80. Ma, H*., et al.* (2017) Correction of a pathogenic gene mutation in human embryos. *Nature* 548, 413-419
81. Zischewski, J*., et al.* (2017) Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnol Adv* 35, 95-104
82. Hynes, A.P*., et al.* (2017) An anti-CRISPR from a virulent streptococcal phage inhibits Streptococcus pyogenes Cas9. *Nature Microbiol* 2, 1364-1380
83. Pawluk, A*., et al.* (2017) Disabling a type I-E CRISPR-Cas nuclease with a bacteriophage-encoded anti-CRISPR protein. *mBio* 8, e01751-17
84. Rauch, B.J*., et al.* (2017) Inhibition of CRISPR-Cas9 with bacteriophage proteins. *Cell* 168, 150-158.e110
85. Shin, J*., et al.* (2017) Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci Adv* 3, e1701620
86. Pyra, A*., et al.* (2017) Tail tubular protein A: a dual-function tail protein of Klebsiella pneumoniae bacteriophage KP32. *Sci Rep* 7, 2223
87. Kutter, E. and Sulakvelidze, A. (2004) *Bacteriophages: Biology and Applications* CRC Press

88 Drulis-Kawa, Z*., et al.* (2012) Learning from Bacteriophages - Advantages and Limitations of Phage and Phage-Encoded Protein Applications. *Current Protein and Peptide Science* 13, 699-722

89. Drulis-Kawa, Z*., et al.* (2012) Learning from bacteriophages - Advantages and limitations of phage and phage-encoded protein applications. *Curr Prot Pept Sci* 13, 699-722

90. Citorik, R.J., *et al.* (2014) Bacteriophage-based synthetic biology for the study of infectious diseases. *Curr Opin Microbiol* 19, 59-69
91. Zeng, L. and Xiao, Z. (2017) A lateral flow biosensor for the detection of single nucleotide polymorphisms. In *Biosensors and Biodetection: Methods and Protocols, Volume 2: Electrochemical, Bioelectronic, Piezoelectric, Cellular and Molecular Biosensors* (Prickril, B. and Rasooly, A., eds), pp. 421-430, Springer New York
92. Schwartz, M., *et al.* (2015) Comparative analysis of T4 DNA ligases and DNA polymerases used in chromosome conformation capture assays. *Biotechniques* 58, 195-199
93. Czarniak, F. and Hensel, M. (2015) Red-mediated recombineering of Salmonella enterica genomes. In *Salmonella: Methods and Protocols* (Schatten, H. and Eisenstark, A., eds), pp. 63-79, Springer New York
94. Gallagher, R.R., *et al.* (2014) Rapid editing and evolution of bacterial genomes using libraries of synthetic DNA. *Nat. Protocols* 9, 2301-2316
95. Kato, M. and Hanyu, Y. (2017) Enzymatic assembly for scFv library construction. In *Synthetic Antibodies: Methods and Protocols* (Tiller, T., ed), pp. 31-44, Springer New York
96. Wang, X., *et al.* (2016) Highly sensitive homogeneous electrochemical aptasensor for antibiotic residues detection based on dual recycling amplification strategy. *Biosens Bioelectron* 82, 49-54
97. Liu, J.L., *et al.* (2014) Viral nanoparticle-encapsidated enzyme and restructured DNA for cell delivery and gene expression. *Proc Natl Acad Sci* 111, 13319-13324
98. Usselman, R.J., *et al.* (2015) Gadolinium-loaded viral capsids as magnetic resonance imaging contrast agents. *Appl Magn Reson* 46, 349-355
99. Mattiacio, J.L., *et al.* (2017) Display of HIV-1 envelope protein on lambda phage scaffold as a vaccine platform. In *Recombinant Virus Vaccines: Methods and Protocols* (Ferran, M.C. and Skuse, G.R., eds), pp. 245-253, Springer New York
100. Chan, J.Y., *et al.* (2017) Targeting insulin receptor in breast cancer using small engineered protein scaffolds. *Mol Cancer Ther* 16, 1324-1334
101. Ji, Z., *et al.* (2016) Fingerprinting of peptides with a large channel of bacteriophage Phi29 dna packaging motor. *Small* 12, 4572-4578

784 102. Zhu, F., *et al.* (2014) DICE, an efficient system for iterative genomic editing in
785 human pluripotent stem cells. *Nucleic Acids Res* 42, e34-e34
786 103. Inniss, M.C., *et al.* (2017) A novel Bxb1 integrase RMCE system for high fidelity
787 site-specific integration of mAb expression cassette in CHO Cells. *Biotechnol*
788 *Bioeng* 114, 1837-1846
789 104. Hsiao, V., *et al.* (2016) A population-based temporal logic gate for timing and
790 recording chemical events. *Mol Sys Biol* 12, 869
791 105. Olszak, T., *et al.* (2017) The O-specific polysaccharide lyase from the phage LKA1
792 tailspike reduces Pseudomonas virulence. *Sci Rep* 7, 16302

793

Figure legends

Figure 1, Key Figure. Proteins with biotechnological applications expressed during the course of a bacteriophage replication cycle. (A) Structural composition of bacterium and phage, and symbols representing phage-encoded proteins. (B) Phage infection begins with adsorption, the interaction of the phage with a specific receptor on the bacterial surface, typically involving (C) the activity of receptor binding proteins (e.g. tail fibers and tail spikes) and depolymerases. (D) This interaction leads to phage genome ejection into the bacterial cytoplasm, which may (E) be aided by VAPGHs that break the bacterial peptidoglycan layer. (F) Once inside, the phage genome may encounter bacterial defenses that aim to degrade it, and counteract it with proteins as anti-CRISPRs (Acr). At this stage (G) the phage may opt for two distinct life cycles: the lysogenic life cycle (H) where the phage genome integrates the bacterial genome with the help of integrases originating bacterial lysogens. The resulting prophage (I) replicates together with the bacterial genome for several generations; or the lytic life cycle (J) where the phage highjacks the bacterial molecular machinery for genome replication and protein expression, using also (K) its own proteins. (L) The structural proteins are then expressed and (M) new phage virions are assembled. To release the newly formed virions (N) late phage proteins as endolysins and holins/pinholins act on the cell membrane to pierce it and (O) cause cell lysis, allowing for a new round of phage infection. (P) Symbol code for the biotechnological applications of phage-encoded proteins indicated in the life cycle.

Figure 2. Biotechnological applications of phage-derived cell binding proteins. (A) Phage receptor binding proteins as tail spikes and tail fibers (or domains thereof), and the cell wall binding domain (CBD) of endolysins have been used as specific cell binding peptides (CBP) in combination with different components for bacterial detection, capture and targeting. The CBP can be combined with (B) fluorescent proteins (e.g. GFP) for the detection of desired bacteria using fluorescent microscopy and/or flow cytometry; (C) horseradish peroxidase (HRP) for the detection of specific bacteria by

enzyme-linked immunosorbent assay (ELISA) after adding an HPR substrate; (D) magnetic nanoparticles for the capture of specific bacteria using a magnetic field, and subsequent bacterial enrichment; (E) gold nanoparticles to create a biosensor-detectable signal after binding to the target bacteria; (F) unspecific anti-microbial peptides (AMP) for targeting and elimination of specific bacteria; and (G) a biosensor that senses modifications on the CBP caused by its interaction with specific bacteria and consequently produces a detectable signal.

Figure 3. Strategies for endolysin engineering. Endolysins of bacteriophages infecting Gram-positive bacteria have a modular structure composed of an enzymatic catalytic domain (ECD) and a cell wall binding domain (CBD) connected by a linker (L). The modular structure of these enzymes led to strategies of engineering based on domain swapping or combination of CBDs with other hydrolytic enzymes. By doing so, endolysins with high catalytic activity and low specificity (ECD1-CBD1), or a strong hydrolytic enzyme (hydrolase), may be combined with endolysins with low catalytic activity and high specificity (ECD2-CBD2) to obtain a chimeric protein either with high catalytic activity and specificity (ECD1-CBD2) to strongly target specific bacteria or with unspecific and high catalytic activity (ECD1) for disinfection purposes. Instead, endolysins of phages infecting Gram-negative bacteria typically have a globular structure composed of a single ECD, although more rarely they may also have a modular structure similar to endolysins targeting Gram-positive bacteria (in an opposite arrangement). Engineering of these endolysins has mostly consisted on their combination with extra peptides (EP-ECD and ECD-EP), such as outer membrane permeabilizing peptides or cell penetrating peptides (mostly polycationic) to give the endolysin access to the peptidoglycan layer of the Gram-negative bacteria.

Tables

Table 1. Summary of the most recent biotechnological applications of phage-encoded proteins

Protein (source)	Application	Description	Year	Ref
RBP				
Long tail fiber (gp37-gp38) (Salmonella phage S16)	Biosensors for whole cell detection	Recombinant long tail fiber (LTF) proteins were used to coat paramagnetic beads for the efficient capture of *Salmonella* Typhimurium cells from food samples. By integrating this LTF-based enrichment method with horseradish peroxidase-conjugated LTF, a new method named Enzyme-linked LTF assay (EELTA) was created which detects as few as 10^2 CFU/mL of S. Typhimurium in 2 h.	2017	[2]
Tail spike LKA1gp49 (Pseudomonas phage LKA1)	Anti-virulence strategies	The tail spike protein (TSP) specifically binds and cleaves B-band lipopolysaccharide (LPS) of *Pseudomonas aeruginosa* PAO1. When employed in an *in vivo Galleria mellonella* model, the enzyme reduce *P. aeruginosa* virulence and sensitized the bacteria to serum complement activity.	2017	[105]
Tail fiber protein tip (gp37) (Escherichia phage T4)	Biosensors for LPS detection	The adhesin was used as the recognition element of a new highly sensitive label-free microwave sensor. The adhesin specifically recognizes the LPS of *Escherichia coli*, causing a change in the capacitance and conductance of the sensor, used as an indicator of LPS detection.	2016	[4]
Tail spikes (Salmonella phages 9NA and P22)	Serotyping	TSPs specifically distinguished glucosylation phenotypes of *Salmonella* O-antigens, when classical methods failed to do so. Variations in glucosylated O-antigens are related to immunogenicity, so the TSP can be used to monitor *Salmonella* epidemiology. It can also be useful for monitoring phase variations during large scale preparation of O-antigens for vaccine production.	2016	[5]
Short tail fiber protein (gp12) (Escherichia phage T4)	Modulator of LPS-induced inflammatory effects.	The tail fiber administered together with LPS in a murine model decrease the inflammatory response to LPS. This suggests the use of the tail fiber as a potential tool for modulating and counteracting LPS-related immune responses.	2016	[7]
Tail fiber protein tip (gp37) (Escherichia phage T4)	Biosensors for whole cell detection	The adhesin was used to create a highly sensitive sensor for the detection of Gram-negative bacteria containing OmpC. The sensor, based on long-period gratings coated with the adhesin, demonstrated sensitivity higher than reference tests.	2016	[3]
Depolymerase				
Depolymerases ORF37 and ORF38 (Klebsiella phages K5-2 and K5-4)	Anti-virulence strategies	The depolymerases were active against the capsules of *Klebsiella*, suggesting their use as adjuvants of the host immune system by decreasing capsule-associated virulence.	2017	[15]
Depolymerase (orf40) (Acinetobacter phage phiAB6)	Development of glycoconjugate vaccines	The depolymerase specifically hydrolysed the exopolysaccharides of *Acinetobacter baumannii*. The released polysaccharides may be used as immunogens for glycoconjugated vaccine production.	2017	[20]
Depolymerase Dpo42 (Escherichia phage vB_EcoM_ECOO78)	Biofilm prevention	The depolymerase degraded the capsular polysaccharides surrounding *E. coli* cells, and exhibited a dose-dependent capacity to prevent biofilm formation.	2017	[24]
Depolymerase depoKP36 (Klebsiella phage KP36)	Anti-virulence strategies	The depolymerase was active both *in vitro* and *in vivo*, significantly inhibiting the mortality of *Galleria mellonella* larvae induced by *Klebsiella pneumoniae*. The depolymerase stability over a broad range of conditions makes it suitable for the development of new treatments for *K. pneumoniae* infections.	2016	[19]
Depolymerase Dpo7 (Staphylococcal phage vB_SepiS-phiIPLA7)	Biofilm prevention and removal	The depolymerase efficiently prevented and removed biofilm-attached staphylococcal cells, although restricted to polysaccharide-producer strains.	2015	[21]

Endolysin				
Endolysin SA.100 (Staphylococcal phage) *Commercialized by Micreos*	Localized antibacterial therapy	The endolysin product Staphefekt SA.100 was successfully used to treat patients with chronic and recurrent *S. aureus*-related dermatoses. There were no signs of induction of resistance by the protein.	2017	[30]
Endolysins PlyP40, Ply511, and PlyP825 (Listeria phages P40, A511, and ProCC P825)	Combinational food preservation techniques	The endolysins were individually combined with high hydrostatic pressure processing, resulting in the synergistic killing of *Listeria monocytogenes*. The results suggest the use of combined processes for the inactivation of *L. monocytogenes* in low processed, ready-to-eat food products.	2017	[36]
Endolysin LysSA11 (Staphylococcal phage SA11)	Food biocontrol and utensil sanitization	Endolysin LysSA11 was active against methicillin-resistant *Staphylococcus aureus* contaminating both food and utensils. The endolysin was similarly active at refrigeration and room temperatures.	2017	[34]
Endolysins AP50-31 and LysB4 (Bacillus phages AP50 and B4)	Antibacterial therapy	The endolysins demonstrated rapid and broad bacteriolytic activity in vitro against strains within the Bacillus genus. Intranasal administration of LysB4 protected mice from death after infection with *Bacillus anthracis* Sterne spores.	2018	[35]
ECD of endolysin A (Mycobacterium phage D29)	Protein purification	The ECD was used as an alternative to sonication or high pressure homogenization for the efficient lysis of *E. coli* cells during protein purification. The ECD is expressed intracellularly and remains non-toxic until chloroform is added to the culture medium. This permeabilizes the bacterial cell membrane allowing diffusion of ECD to the peptidoglycan layer where it acts causing cell lysis. The method is applicable in high-throughput and large-scale protein purification.	2017	[50]
Chimeric endolysin LysK (Staphylococcal phage K)	Antibacterial therapy	The properties of primary, secondary and tertiary structure of endolysin LysK were improved using *in silico* design, and resulted in enhanced stability, solubility and antibacterial activity of the enzyme against *S. aureus*, *S. epidermidis* and *Enterococcus*.	2017	[41]
Recombinant endolysin Sal-1 (Staphylococcal phage SAP-1)	Antibacterial therapy	A recombinant form of endolysin SAL-1 was used for the development of drug SAL200, for the treatment of antibiotic-resistant staphylococcal infections. Phase 1 studies of SAL200 administered intravenously reported no serious adverse effects, supporting the progress of the drug for later phase studies.	2017	[43]
Modified lysozyme mbT4L (Escherichia phage T4)	Purification tag	A metal ions-binding mutant of phage T4 lysozyme (mbT4L) was used as a purification tag in immobilized-metal affinity chromatography (IMAC), proving advantageous over the conventional IMAC technique. The mbT4L protein is suggested to be compatible also with X-ray crystallography.	2017	[49]
Chimeric protein CHAPSH3b (Staphylococcal phage vB_SauS-phiIPLA88)	Biofilm control	The chimeric protein is a fusion of the ECD of VAPGH HydH5 with the CBD of lysostaphin. The protein was able to control biofilm-embedded *S. aureus* and decrease biofilm formation by some strains.	2017	[39]
Endolysin MR-10 (Staphylococcus phage MR-10)	Localized and systemic combinational antibacterial therapy with antibiotics	The endolysin was combined with the antibiotic minocycline, in a single dose, resulting in complete survival of mice with systemic methicillin-resistant *S. aureus* infection. Encouraging results were also obtained when applying the combined therapy to localized burn wound infections.	2016	[29]
Endolysin PlyC (Streptococcal phage C1)	Intracellular antibacterial therapy	The endolysin PlyC was able to control intracellular *Streptococcus pyogenes* by crossing epithelial cell membranes. By doing so, the endolysin creates new opportunities to avoid refractory infections caused by the internalized pathogen.	2016	[31]
Artilysin Art-175 (Pseudomonas phage varphiKZ)	Antibacterial therapy	Artilysin Art-175 instantaneously killed stationary-phase cells of multidrug-resistant *A. baumannii*, with no sign of development of resistance.	2016	[46]
Artilysin Art-240 (Streptococcal phage λSa2)	Antibacterial therapy	Fusion of endolysin λSa2lys with the polycationic peptide PCNP generated Artilysin Art-240, a hybrid protein with specificity similar to the parental enzyme, but increased stability and	2016	[47]

		bactericidal activity.		
Endolysin ABgp46 (Acinetobacter phage vB_AbaP_CEB1)	Antibacterial therapy	The endolysin ABgp46 demonstrated antibacterial activity against several multidrug resistant *A. baumannii* strains. The activity of the endolysin was broadened to other Gram-negative pathogens, including *P. aeruginosa* and *S.* Typhimurium, when combined with the outer membrane permeabilizing agents citric and malic acid.	2016	[44]
Endolysin E (Escherichia phage T4)	Crystallization	The endolysin was fused to the ligand-binding domain (LBD) of the fungal sterol transcription factor Upc2, improving expression and crystallization of Upc2 LBD.	2015	[48]
Chimeric endolysin Ply187N-V12C (Staphylococcal phage 187, enterococcal phage phi1)	Antibacterial therapy	Fusion of the ECD of staphylococcal endolysin Ply187 with the CBD of enterococcal endolysin PlyV12 extended the lytic activity of Ply187 to streptococci and enterococci.	2015	[38]
CBD				
CBD of endolysin LysB4 (Bacillus phage B4)	Biosensors for whole cell detection	The CBD was used as the recognition element of a nitrocellulose-based lateral flow assay, which employed colloidal gold nanoparticles as a colorimetric signal for bacterial detection. The method provided results superior to those of antibody-based approaches, detecting 10^4 CFU/mL of *Bacillus cereus* in 20 min.	2017	[52]
CBD of endolysin CTP1L (Clostridium phage phiCTP1)	Biosensors for whole cell detection	The CBD was fused to green fluorescent protein (GFP) for the specific detection of dairy-related *Clostridium* species by fluorescence microscopy. The GFP-CBD was also capable of binding to clostridial spores, and permitted the visualization of vegetative cells of *Clostridium tyrobutyricum* directly in the matrix of late blowing defect cheese.	2017	[54]
CBD (NA)	Targeted delivery of antimicrobials	The CBD was fused to antimicrobial silver nanoparticles (AgNPs) for their targeted delivery. The construct improved specificity and killing efficiency of the target bacteria, when compared to bare AgNPs.	2017	[55]
CBD of endolysin PlyV12 (Enterococcal phage Φ1)	Biosensors for whole cell detection	CBD of endolysin PlyV12 was coupled to immunomagnetic separation for the rapid and sensitive detection of *S. aureus* cells in spiked milk. The sensitivity of detection (detection limit of $4x10^3$ CFU/mL) is improved by the large number of binding sites available at the cell surface for CBD attachment.	2016	[53]
Holin				
Holin HolGH15 (Staphylococcal phage GH15)	Antibacterial therapy	The holin has shown efficient antibacterial activity against a broad range of species, including *S. aureus*, *L. monocytogenes*, *Bacillus subtilis*, *P. aeruginosa*, *K. pneumoniae* and E. coli.	2016	[58]
Endolysin and holin				
Endolysin R and holin S (Escherichia phage lambda)	Production of bacterial ghosts for vaccine production	The holin and endolysin of phage lambda were used to construct novel plasmids for the production of *Salmonella* ghosts of high immunogenicity. This strategy is expected to contribute to the development of novel and safe non-living vaccine candidates.	2017	[51, 61, 62]
Endolysin E and holin T (Escherichia phage T4)	Biofuel recovery	The endolysin and holin genes of phage T4 were introduced in cyanobacteria under the control of a promoter regulated by a physical signal. When the cells are exposed to the signal, the lysis proteins are expressed causing cell death. This strategy is suggested for the recovery of biofuels and related compounds without resorting to chemical inducers and mechanical disruption, with the advantage of controlling the accidental release of cyanobacteria.	2014	[69]
VAPGH				
VAPGH HydH5 (Staphylococcal phage vB_SauS-phiIPLA88)	Biocontrol in dairy	The VAPGH and its variants fused to lysostaphin (or domains thereof) were shown to have efficient lytic activity against *S. aureus* in both raw milk and milk pasteurized at 72 ˚C.	2013	[74]
Anti-CRISPR				

anti-CRISPR proteins AcrIIA2 and AcrIIA4 (Listeria prophages)	Regulation of CRISPR-Cas9 genome editing	The anti-CRISPR proteins were shown to inhibit the widely used *Streptococcus pyogenes* Cas9 when assayed in bacteria and human cells.	2017	[84]
anti-CRISPR protein AcrIIA5 (Streptococcal phage D4276)	Regulation of CRISPR-Cas9 genome editing	The first anti-CRISPR protein isolated from a virulent phage was able to completely inhibit *S. pyogenes* Cas9 (SpCas9) activity *in vitro*.	2017	[82]
anti-CRISPR protein AcrIIA4 (Listeria monocytogenes prophages)	Regulation of CRISPR-Cas9 genome editing	The anti-CRISPR protein bound only to assembled Cas9-single-guide RNA complexes and not to the Cas9 protein alone. AcrIIA4 delivered a few hours after the Cas9 editing tool in human cells allowed on-target Cas9-mediated gene editing and reduced off-target editing and its unpredicted side effects.	2017	[85]
anti-CRISPR AcrE1 (gp34) (Pseudomonas phage JBD5)	Regulation of type I-E CRISPR-Cas	The anti-CRISPR bound Cas3 and inactivated the type I-E CRISPR-Cas system in *P. aeruginosa*. AcrE1 can convert the endogenous type I-E CRISPR system into a programmable transcriptional repressor, providing a new biotechnological tool for genetic studies of bacteria encoding this CRISPR system.	2017	[83]
Other				
P22 capsid (Enterobacteria phage P22)	Improvement of contrast agents	The viral capsids were loaded with the paramagnetic gadolinium ion (positive contrast agent) to use as contrast agents in magnetic resonance imaging. The capsids efficiently enhanced the relaxivity of the contrast agent, which is expected to increase image contrast.	2017	[98]
Lambda capsid (Escherichia phage lambda)	Vaccine development	The capsid was used as a scaffold for the display of the human immunodeficiency virus envelope spike protein. This vector can be used for vaccine development, with advantages over mammalian virus vectors of genetic tractability, inexpensive production, aptness for scale-up, and stability.	2017	[99]
T4 procapsid (Escherichia phage T4)	Targeted gene/cancer therapy	The procapsid was packaged *in vitro* with a mCherry expression plasmid and *in vivo* with the active Cre recombinase. The capsid-based nanoparticles were delivered into cancer cells, in which the Cre recombinase circularized the linear expression plasmid, resulting in enhanced expression of mCherry. This strategy overcomes the main limitation of *in vitro* DNA packaging, its restriction to linear molecules, which hinders the biological activity of the DNA on the target.	2014	[97]
BxB1 integrase (Mycobacterium phage BxB1)	Cell line engineering	The integrase was used to build a novel recombinase mediated cassette exchange (RMCE) system, with fidelity of RMCE events higher than those obtained with the common Flp/FRT RMCE system. This system provides a novel tool for the engineering cell lines for biotherapeutic production.	2017	[103]
Serine integrases TP901-1 (intA) and Bxb1 (intB) (Lactococcus phage TP901-1 and Mycobacterium phage Bxb1)	Construction of memory genetic logic gates for detection of biological events	The integrases were used to create a two-input temporal logic gate capable of sensing and recording the order of inputs, the timing between inputs, and the duration of input pulses. The integrases were specifically used for unidirectional DNA recombination to detect and encode sequences of input events.	2016	[104]
PhiC31 and BxB1 integrases (Streptomyces phage phiC31, Mycobacterium phage BxB1)	Genetic modification of human stem cells	The integrases were used for the development of DICE, Dual integrase cassette exchange. The system offers rapid, efficient and precise gene insertion in stem cells, and is particularly well suited for repeated modifications of the same locus.	2014	[102]
Red recombinase (Escherichia phage lambda)	Engineering of bacterial genome	The Red recombinase of phage lambda was used for enterobacteria genome mutagenesis, such as rapid generation of genome deletions, site-directed mutagenesis, generation of reporter fusions or chimeric genes, and transplantation of regulatory elements into the cell chromosome.	2015	[93]
T7 modified RNA polymerase inhibitor	Modulation of insulin receptor	The *E. coli* RNA polymerase inhibitor was subjected to direct evolution for the development of variants able to inhibit insulin	2017	[100]

(modified gp2) (Escherichia phage T7)	signalling	receptors. The variants inhibited insulin-mediated proliferation of breast cancer cells, without downregulating the expression of the insulin receptor.		
T4 DNA ligase (Escherichia phage T4)	Single nucleotide polymorphism (SNP) analysis	The DNA ligase was used for the development of a simple and robust lateral flow biosensor method for the detection of SNPs. The method is suggested to have great potential for the detection of genetic diseases, cancer-related mutations, and drug-resistant mutations of infectious agents, and for the development of personalized medicine.	2017	[91]
T7 exonuclease (Escherichia phage T7)	Detection of antibiotic residues	The exonuclease was used for developing a highly sensitive homogeneous electrochemical strategy for the detection of ampicillin residues, which can be applied both for clinical and food safety purposes. The exonuclease provides amplification of electrochemical signals with a limit of detection superior to those reported in the literature.	2016	[96]
phi29 motor channel (gp10 portal protein) (Bacillus phage phi29)	Fingerprinting of peptides	The motor channel was used for peptide fingerprinting in single molecule electrophysiological assays. The protein is used to generate peaks of current blockage that serve as typical fingerprint for peptides with high confidence. The results demonstrate the potential of the motor channel for detection of disease-associated peptide biomarkers.	2016	[101]
Exonuclease Exo, and DNA-binding proteins Beta and Gam (Escherichia phage lambda)	Engineering of bacterial genome	The Exo, Beta, and Gam proteins were exploited for the development of MAGE, Multiplex Automated Genome Engineering. MAGE uses synthetic single-stranded DNA to introduce targeted modifications into the chromosome of *E. coli*, to generate combinatorial genetic diversity in a cell population, or for genome-wide editing.	2014	[94]

853