ENGINEERING BACTERIOPHAGES TO ENHANCE THEIR POTENTIAL USE IN THERAPY

AURELIJA MARIJA GRIGONYTE

A thesis submitted for the degree of Doctor of Philosophy

School of Life Sciences, University of Warwick

September 2018

Table of Contents

Acknowledgements

I would like to thank to my supervisor Dr. Andrew Millard for being the best supervisor a PhD student could wish for, who without a failure provided the support and tirelessly tried (and succeeded) to shine some light on to my negative nature. His scientific drive and rigour, I will strive to achieve in my future research endeavours.

I would like to thank Dr. Chrystala Constantinidou for being the kindest and the most inspiring female scientist role model with the best sense of humour.

To my PhD committee Prof. Keith Leppard and Dr. Nick Waterfield for pointing me in the right direction at the right time.

I thank my fellow lab mates for the stimulating discussions, numerous cakes, help with experiments and motivation. Slawomir Michniewski for countless discussions on phage biology D-tours and always being in the lab by 8 am. Tamsin Redgwell, a.k.a sequencing wizard, for being a flapjack and coffee provider, always at the right moment. Eduardo Goicoechea Serrano for being a daily Kuala-tea pun provider. Paul MacDonald for being there at the very dark and the very bright moments. My mother for always being my life support: ačiū tau mamiuka, už palaikyma, supratima ir tikejima manimi, be tavęs nieko nebūtu.

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so we may fear less.

Marie Curie

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented in this thesis (including data generation and data analysis) was carried out by the author except in the cases outlined below:

- Data generated in Figures 4-8 and 4-9 by Mark van Raaij's laboratory (at Centro Nacional de biotechnological) members.

- Data generated in Figure 7-3 by Slawomir Michniewski.

Summary

The rise of antibiotic resistance (AMR) is one of the world's major health threats[1]. The infections caused by antimicrobial resistance bacteria are increasing faster than the introduction of new antibiotics[2,3]. The emergence of a multi-drug resistant bacteria has pushed towards bacteriophage therapy as one of the alternatives to antibiotics[4]. Bacteriophages are viruses that infect bacteria. Lytic phages have been examined as a potential therapy against drug resistant bacteria because they propagation involves killing of their bacterial host[5–7]. In addition to vast number of advantages as therapeutics, bacteriophages have several limitations i.e. narrow host range and pharmokinetics – not being able to get high enough phage concentration to the site of infection. A variety of methods have been used for the genetic modification of bacteriophage genomes to undercome bacteriophage limitations, including Yeast Artificial Chromosome (YAC), CRISPR/Cas systems and marker-based methods. However, no direct comparison has been carried to standardise the methods in question. This study aimed to establish the most efficient method by comparing effectiveness of CRISPR/Cas and marker-based systems (Chapter 3) followed by implementation of the method to address narrow host range and tissue binding specificity. I show that the most optimal engineering method is *trxA* marker-based method (Chapter 3). I then modify T7 tail fibers using the method established to retarget an alternative host, *Bordetella pertussis* (Chapter 4) and characterise them on *E. coli* expressing the receptor of interest as well as *Bordetella pertussis* and *Bordetella bronchiseptica* strains (Chapter 4). This uncovered an unexpected T7 infectivity of *Bordetella bronchiseptica* strains. Furthermore, I uncovered a novel way for T7 infection, a co-infection between T7 and genetically modified (chimeric phage) (Chapter 5). I hypothesised that dependence arose from T7 lack of host factor gene (*trxA*) and chimeric phage lack of infective tail fibers. I examined the co-dependency by inducing the bacterial lysis after secondary addition of either of the phages to verify that both phages were required to yield the lysis. I further demonstrate that phage genetic modification can allow phage to carry additional cargo, that can be used to kill bacteria in addition to the phages natural host. To do this I engineered phage T4 to express three different endolysins, and verify the ability of these phage to also lyse *S. aureus*, *C. difficile* and *B. subtilis*. Finally, I examined ways of making T7 phage tissue specific. This was achieved by genetically modifying T7 capsid, major and minor proteins, with small homing peptides that specifically bind to lung epithelial tissue (Chapter 7). This study provides an important leap forward in our understanding of bacteriophage engineering towards therapeutics. In so doing it also uncovers unexpected co-operative phage behaviour upon infection when exposed to limiting resources.

List of abbreviations

AMR Antimicrobial Resistance

CBDs Cell Wall Binding Domains

CFU Colony Forming Units

CH Chimeric Phage

CRISPR Clustered Regularly Interspaced Short Palindromic Repeats

dsDNA Double Stranded DNA

EADs Enzymatically Active Domains

EOP Efficiency of Plating

EPS Extracellular Polymeric Substances

FDA Food and Drug Administration

GMP Good Manufacturing Practices

Hoc Highly Antigenic Outer Capsid protein

HP Homing Peptide

HR Homologous Recombination

IM Inner Membrane

LPS Lipopolysaccharide

LTFs Long Shaped Fibers

MGE Mobile Genetic Elements

MOI Multiplicity of Infection

YAC Yeast Artificial Chromosome

Mtd Major Tropism Determinant Protein

OM Outer Membrane

PAM Protospacer Adjacent Motif

PCR Polymerase Chain Reaction

PFU Plaque Forming Units

PGHs Peptidoglycan Hydrolases

PTMPs Phage Therapy Medicinal Products

pre-crRNA CRISPR RNA

Prn Pertactin

RBD Receptor Binding Domain

RBP Receptor Binding Protein

RBS Ribosome Binding Site

Soc Small Outer Capsid Protein

ssDNA Single Stranded DNA

TBD Thioredoxin Binding Domain

TEM Transmission Electron Microscopy

TrxA Thioredoxin

WT Wild Type

List of figures

List of tables

List of Appendices

Chapter 1 Introduction

1.1 Antimicrobial resistance

The rise of antibiotic, or antimicrobial, resistance (AMR) is one of the world's major health threats[1]. The infections caused by antimicrobial resistant bacteria are increasing faster than the introduction of new antibiotics[2,3]. AMR genes have been discovered in bacterial strains that pre-date the first use of antibiotics by humans, suggesting that AMR is a naturally existing phenomenon that was potentially aggravated by societal use of antibiotics[8]. AMR seen in humans is closely related to AMR in other animal populations Figure 1-1[9]. This similarity is due to a close interplay between the environment, animals and humans facilitates the movement of the drugs, the bacteria as well as mobile genetic elements (MGEs)[10]. The complexity of the interaction is scaled up by the broad range of antibiotics that affect multiple bacterial species and the fact that the acquired resistance can be obtained via MGEs. It has been suggested that AMR found in clinical environments is linked to similar mechanisms seen in the environment[11].

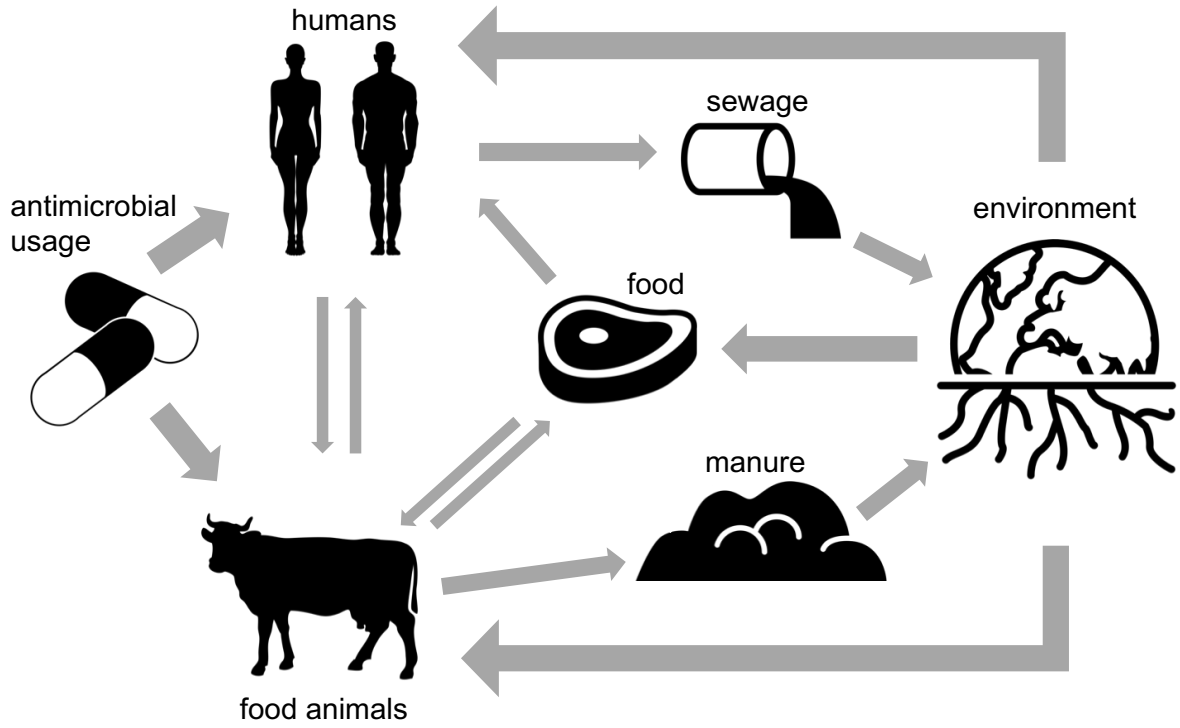


Figure 1-1. Diagram of AMR generation and transmission between humans, farm animals and the environment. Adapted from Woolhouse et al. 2015. High usage of antimicrobials i.e. antibiotics in human disease treatment as well as farming can lead to antibiotic persistence in manure and sewage which is then deposited in the environment. The environmental antibiotic deposits cycle back to humans as well as farming industry.

The introduction of antibiotics in the 1940s, that started with penicillin and streptomycin, revolutionised medicine allowing effective treatment of bacterial infections[12–14]. Antibiotic discovery is a difficult process, mainly due to poor penetration of various compounds into bacterial cells[15]. Most drugs used in today's clinic were discovered by screening cultivable soil microorganisms in the 1960s[12]. However, this limited resource eventually came to an end with no synthetic approaches able to replace the antibiotics previously discovered via natural product mining[15,18]. The deceleration of new drug invention and manufacturing resulted in a lack of financial incentive by pharmaceutical companies[13]. The emergence of multidrug resistant bacteria spurred research into alternative treatments[4]. This lead towards a revival of studies on potential replacements for antibiotics such as probiotics, vaccines, antibodies and immune stimulation, as well as bacteriophage and bacteriophage-based product therapy[4,19–28].

1.2 Bacteriophages

Bacteriophages (phages) are viruses that infect bacteria. They are obligate parasites that are made up of genetic information, encoding all viral characteristics, packaged in a protein coat. They are parasites that exploit their host machinery for their propagation, which can be affected by a variety of factors such as nutrients, temperature and light[29]. Phages are arguably the most abundant biological particles on the planet, with an estimated 10^{31} in the biosphere[30]. Phages have two main lifestyles: lytic and temperate, that will be discussed in sections 1.3 and 1.4. Specifically, lytic phages have been examined as a potential therapy against drug resistant bacteria because their propagation involves the killing of their bacterial host[5–7].

1.3 Lytic or virulent phages

A lytic phage is a phage that induces lysis, or cell death, of its host upon replication. Lytic or virulent phages have been used in a variety of areas including food processing, bacterial biofilm removal and biotechnology[31–34]. Upon bacterial infection a lytic phage replicates faster than its host, with a complete infection to lysis cycle taking between 15 minutes to 36 hours[35–37]. In brief, the main stages for the lytic phage cycle are attachment, replication and lysis. During the attachment step, the phage tail fibers bind to a specific receptor on the surface of the target, or host, bacterium[38]. During the entry step, the phage injects its DNA into the host cytoplasm[39]. The following step is DNA replication, during which phage protein synthesis also takes place[39]. This is followed by assembly of phage capsid proteins and DNA packaging inside the

assembled protein capsids[39]. Late in the lytic cycle the phage expresses genes for holins and lysis that generate holes in the cell membrane of the host cell, resulting in the disbalance of water flowing inside the cell[40]. This causes the host cell to undergo lysis and death, that in parallel liberates progeny phage particles, which are each then ready to start another cycle by infecting new host cells[40]. The lytic phage cycle described above is referred to as 'lysis from within', as bacterial lysis is caused by phage proteins intracellularly[41]. At high concentrations lytic phages can also cause an alternative phage lysis, the phenomenon of 'lysis from without', in which bacteria lyse solely due to adsorption of external agents[42,43], such as phage. This phenomenon has been most frequently studied for T2 and T4 phages[43,44]. It arises when substantial numbers of phages adsorb to the bacterial cell envelope, causing multiple penetrations and cell-wall damage[44]. Even only minimal damage to the cellular wall will result in cellular lysis[45,46].

1.4 Temperate phages

Temperate phages can either act as lytic phages and kill their host upon infection, or form a stable relationship with their host by either integrating into the host chromosome or by forming a replicative plasmid[47–50]. If they form a plasmid they become prophages and replicate together with the bacterial genome upon cell division, a process referred to as the lysogenic cycle[51]. This lysogenic cycle allows phage reproduction without host death. In the lysogenic phage life cycle the attachment and DNA injection steps are the same as for the lytic phages[51]. However, once the DNA has been injected into the host cell, gene transcription and DNA replication does not happen immediately[52]. Instead the phages undergo a recombination event with an insertion site in their host genomes, allowing them to be integrated into the host genome[52]. This allows them to replicate alongside the host for multiple generations, co-existing with instead of lysing the host[29]. This phenomenon is termed 'polylysogeny', only happens in heteroimmune phages, which is a characteristic of two temperate phages where a prophage of one type fails to show immunity against the second time[53]. These phages can undergo polylysogeny because they are different enough and do not display immunity against each other. In contrast, homoimmune phages, show immunity against each other's' prophages and thus cannot undergo polylysogeny ensuring only one phage copy per one host cell, with some exceptions.

It has been shown that two Stx phages with identical immunity patterns, φE85539-Stx2[1] and φE85539-Stx2[2], were isolated from the same strain, indicating the homoimmunity rule described above is not always the case[54].When two phages have different immunity profiles,

they are referred to as heteroimmunity phage groups. Two temperate phages that present heteroimmunity to each other are able to have lytic and lysogenic cycles that are independent and do not interfere with each other's repressors[55,56]. Heteroimmunity could potentially lead to superinfection that is where second infection is superimposed on an earlier one, without interference[55].

1.5 Pseudolysogeny

In addition to lytic and lysogenic phage life cycles, pseudolysogeny has been proposed as a third potential path of phage development[57–60]. Although the genetic mechanism of this phage-host interaction has not been characterised, it has been seen and defined experimentally[61,62]. Pseudolysogeny is a carrier state, in which a 'lytic' phage is able to remain inside the bacterial cell until cell conditions are suitable for the phage to enter the lytic infection cycle. It is a state where the phage remains in the host cell without undergoing replication by lysing its host, integrating into the host genome or forming a replicative plasmid[63]. This phage state is caused by unfavourable growth conditions, or stress caused to the host bacterium[64]. Although difficult to study, this was determined after carrying out experiments in *Pseudomonas aeruginosa* strains, where F116 and UT1 phage-postponed lysis was observed under nutrient deficient conditions, followed by addition of nutrients that resulted in immediate cell lysis and phage production[64,65]. Furthermore, pseudolysogeny was identified in other systems including phages infecting *Escherichia coli (E. coli)* and cyanobacteria[58,66]. In the previous work it was demonstrated that *E. coli*-infecting phage T4 was able to form pseudolysogens when exposed to low nutrient levels[58]. Another study suggested that obligately lytic *Synechococcus* phage S-PM2 when used to infect *Synechococcus* sp. WH7803 in phosphate-depleted media resulted in a 80% reduction in burst size and hence the potential formation of pseudolysogens[66].

1.6 Phage therapy as an alternative to antibiotics

Even though phage therapy has only recently received a lot of attention as a possible treatment for multidrug-resistant pathogens, its potential was initially discovered over a century ago[4,67]. Frederick Twort was the first to describe and characterise the signature zones of lysis of phage infection in 1915[67]. However, Felix d'Herelle identified the actual source of this observation and associated it to bacterial viruses, called them "bacteriophage" (literally "bacteria-eater")[67]. He was also the first to notice the potential of phage to be used to treat bacterial infections, and

successfully used phage to treat four paediatric cases of bacterial dysentery[67]. Even though the scientific community condemned Felix d'Herelle`s work and results due to them being poorly carried out, he continued to pioneer phage therapy with treatment of the bubonic plague, dysentery and cholera[68,5]. One of his clinical trials involved treatment of cholera in the Punjab region of India. The study, consisting of 73 treatment subjects and 118 control subjects, showed that treating patients using phage therapy resulted in a 90% reduction in mortality[67].

Recent studies have examined the potential of phage to be used against a range of clinical pathogens. One such study examined ileocecitis caused by *Clostridium difficile* (*C. difficile*) in mice, in which the pathogen and a dose of phage were administrated at the same time, preventing the initiation of the disease[69]. In addition, when a phage was administered post infection with the pathogen only 1 mouse died out of 12, while all 12 mice that were only administered the antibiotic clindamycin died within 96 hours[69]. Bacteremia models have been used to show that intraperitoneal application of phage was enough to generate 100% *Enterococcus faecium (E. faecium)* and *P. aeruginosa* infection free mice[70,71]. In addition to using single phage for treatment in animal models, phage cocktails have been used for antibiotic-resistant *P. aeruginosa* infections present in multiple organs including intestines, skin and lungs[72,73].

Over the last century, phage therapy has been applied to humans in a few institutes, including the Institute of Immunology and Experimental Therapy in Wroclaw, Poland and the Eliava Institute of Bacteriophage in Tbilisi, Georgia[74,75]. The Institute of Immunology and Experimental Therapy in Wroclaw, Poland has a vast range of phages at their disposal[76–78]. The institute continuously isolates new phages to expand its collection[78]. They followed 550 of their patients throughout 1981–1986 and reported cure rates ranging from 75 to 100% depending on infection type[79–83].

There have been multiple attempts to classify phage-based treatment within EU framework[84–87]. The current agreement among the European Medicines Agency is to monitor phage-based therapeutics as biological medicinal products[88]. One of the key issues for phages is the indecision as to whether phage therapy medicinal products (PTMPs) require patenting and marketing authorisation[89]. Since phage cocktails would potentially be customised, the PTMPs are placed somewhere between industrially made medicinal products and magistral formulas. This classification uncertainty is due to a contradiction within the specifics of the treatment[89].

A phage cocktail and its constituent parts would be tailored towards each individual patient, while at the same time each of the active ingredients would have to fulfil the characteristics of an industrial process in compliance with the Good Manufacturing Practices (GMP) requirements[89]. Even though phage therapy has not been accepted as a licensed drug, EU legislation as well as the USA Food and Drug Administration (FDA), allows their use as a compassionate, last resort medicine[89,90].

The Eliava Institute has an extensive collection of phages that work against most common bacterial pathogens including *P. aeruginosa, S. aureus, E. coli* as well as species of *Proteus*, *Salmonella*, *Enterococcus*, *Streptococcus*[91]. The phages collected are mainly used for prophylactic and therapeutic applications, including treatment of gastroenterological issues and post-surgical infections[91]. In six cases antibiotic-unresponsive diabetic foot ulcers were successfully treated using pathogen specific, topically applied phage[92]. Recently phage therapy has been used as a last resort medicine in Belgium as well as the USA[93–95]. In one instance, a 2-year old experiencing multidrug-resistant *Pseudomonas aeruginosa* infection with bacteremia was treated using a two phage cocktail that resulted in clearing of bacteremia[94]. Currently no licensed products are approved for human treatment either in the United States or Europe. However, this is not the case when it comes to the food industry. There are multiple commercial phage products against meticillin-resistant *Staphylococcus aureus* (MRSA), *Listeria monocytogenes*, *E. coli* O157:H7, *Mycobacterium tuberculosis*, *Salmonella* spp, that have been approved by the FDA[96–99]. The use of naturally occurring phages for use in the food industry is also allowed in Europe[100].

1.7 Bacteriophage Host Range

A phage's 'host range' consists of the pathogens which a phage is able to infect[101]. Phage host range determination can be difficult due to the fact that measured host ranges rely on the technique used[101]. It has been stated that "host range is often, but not always, determined by success or failure of adsorption" but for phage therapy, host cell lysis rather than phage adsorption to the host plays a key role in host range determination[101,102]. Phage host range can be limited to a specific species of bacteria[101]. However, there are multiple examples where phages evolved to expand their host range within a species and to a different species by binding to a new host receptor or overcoming restriction immunity (see section 1.23)[103–105]. Moreover, there are phages that can even adapt to infect new genera[106,107]. Phage binding specificity to its host comes from the phage receptor binding domain (RBD) of its tail fiber protein or receptor binding protein (RBP)[108]. The interaction between RBP and its bacterial cell surface receptor is one of the key parameters that determine phage infection kinetics[109]. While host specificity might be useful in certain scenarios, a broader host range may be desired in a clinical setting. This is one of the reasons why the use of broad-spectrum antibiotics was preferred to that of specific host targeting phage therapeutics in the early 1940s[110]. To overcome the narrow host range of some phages, cocktails containing several phages have been used to provide pharmacologically diverse formulations[111].

Both phage and their hosts undergo selective pressure, where cells evolve resistance to phage infection and phage adapt to overcome this, in a continuous arms race known as the red queen principle[112]. There are multiple ways which can be employed by a phage to expand its host range[113]. Point mutations in the phage tail fiber gene is one example of host range expansion[114]. This has been seen experimentally in tail fiber mutant T7 variants which were shown to infect LPS mutant *E. coli* strains that were immune to the wild type T7[114]. Another possibility is one phage obtaining a fragment of a tail fiber gene from another phage, allowing phage host expansion between unrelated phages. This has been observed in such as phages P2, Mu and P1[115]. Furthermore, phage host range can be expanded by picking up a piece of genetic debris that remained in the host chromosome from another bacteriophage, as demonstrated by T2 and K3 phages[116]. Phages may also be able to recognize multiple receptors via the same tail fiber. For instance, the T4 phage can recognize both the outer membrane protein OmpC and the lipopolysaccharide as a receptor[117]. In some instances phages may have two adsorption systems, allowing different receptor recognition[118]. One such example is phage T5. The T5

RBP, pb5, binds to the *E. coli* outer membrane protein FhuA[119–121]. T5 also has an additional adsorption apparatus via three long shaped fibers (LTFs), formed by Pb1 protein[121]. It has been shown that LTFs bind to the bacterial lipopolysaccharide (LPS) O-antigen, providing an additional binding site for the T5 phage[122,123].

In some instances, phages are able to adapt their RBP when their host receptor changes. An example system is that of the temperate bacteriophage, BPP1, infecting *Bordetella bronchiseptica*[124]. In this system, BPP1 changes its tropism by genetically altering the DNA sequence of its host receptor-binding protein, referred to as the major tropism determinant (Mtd)[125]. BPP1 needs to be able to variate the Mtd sequence to adapt to the Bordetella surface variation, which is regulated by BvgAS phosphorelay pathway responsible for the infectious cycle[126–128]. BPP1 encodes a unique diversity-generating system that allows for Mtd sequence variation and hence attachment to different host receptor molecules that are displayed throughout Bordetella infectious cycle[124,129]. The *Bordetella* surface receptor profile depends on its infection phase, therefore phage BPP1 adapts to *Bordetella* surface receptor change by varying the Mtd sequene[125]. Mtd sequence variation is possible due to transcription followed by reverse transcription and integration with site-directed, adenine-specific mutagenesis[124,130]. Chapter 4 will provide further details on BPP1 phage versatility.

1.8 Advantages of phage use for therapy over antibiotics

Some but not all antibiotics are bacteriostatic agents (e.g. tetracycline), that prevent bacteria from replication without necessarily inducing bacterial death[131–133]. The stunted microbes are more likely to evolve resistance towards the drug[131–133]. In contrast, phages are bactericidal agents which cause bacterial death, reducing the potential for evolution of resistance[132]. In addition to their bactericidal properties, phages are able to replicate at the site of infection. This allows single, instead of multiple, dosages when there is a greater amount of pathogen than expected at the site of infection. The phage ability to amplify sufficiently depending on the bacteria present is termed as 'auto dosing'[51,53,134,135]. Furthermore, phages are made up mainly of proteins and nucleic acids and hence they are inherently nontoxic[136,137]. Other phage advantages include their rapid discovery and potential for biofilm clearance[138–143]. In the case of their rapid discovery, phages are widely distributed in soil, sewage, water and animal organisms, therefore their abundance contributes towards quick and efficient isolation[144–146]. Furthermore, experimental studies have also identified that phages may both prevent biofilm

and aid eradication of biofilm forming bacteria[147]. In particular, phage depolymerases facilitate the degradation of extracellular polymeric substances (EPS), enabling phage penetration to deeper areas of the biofilm[147].

1.9 Disadvantages of phage use for therapy as compared to antibiotics

Lytic phages are preferred over temperate phages for phage therapy, due to temperate phages' association with bacterial virulence factors and toxins genes[148]. In addition, establishment of lysogeny may cause the pathogen to become immune to infection by other phages[148]. One of the current disadvantages of using phages for therapeutic purposes is the lack of understanding of the pharmokinetics and pharmacodynamics of phages in the human body[149]. This includes both phage interactions with tissues and organs as well as interactions with the body's microbial flora[150]. Moreover, it has been suggested that phage application will eventually result in the evolution of phage resistance, in a similar manner as has occurred for antibiotics[151,152]. This could happen due to multiple bacterial defence strategies that include receptor mutations, restriction modification (R/M) systems in particular DISARM, abortive infection systems that trigger cell death or metabolic inactivation upon infection (toxin-antitoxin system) and acquired immunity (CRISPR)[152–158].

Another disadvantage of phages is the presence of lipopolysaccharide (LPS, endotoxin, pyrogen) in their crude lysates produced in Gram-negative bacteria[159]. The presence of LPS and endotoxins can cause innate immunity response, as they act as an 'alarm molecule' indicating the invasion of Gram-negative bacteria[160–162]. To counter act this multiple methods have been developed for pyrogen removal from phage lysates[162,163]. As a consequence, highly purified phages can be prepared to elucidate anaphylactic responses from pyrogens such as endotoxins and LPS that are found in unpetrified phage lysates[135,136,164,165].

Another two disadvantages, narrow host range and lack of tissue specificity will be discussed in more detail in the next two subsections.

1.10 Phage host range disadvantage over antibiotics

The narrowness of a phage host range (a few strains or single species) may limit the pace at which the treatment can be applied, as suitable phage has to be found to target the bacteria causing infection, prior to phage administration[101]. This is often combated by combining multiple phages, generating phage 'cocktails', that have broader host range than single phages and at the same time are more selective when compared 'narrow-spectrum' antibiotics[166,167]. This on the other hand could also be viewed as an advantage when it comes to preserving the good bacteria found in body as the phage administered would only target the bacteria of

interest[133]. In addition, having a phage cocktail reduces bacterial resistance due to an increased number of mutations required to achieve this [168,169] . When bacterial resistances were compared after single versus phage cocktail, it was found up to 43% resistance for the single phage treated bacterial strains and 24% resistance for the cocktail treated strains[168].

1.11 Phage penetration into the bloodstream and organs

Multiple studies have shown that intravenous administration of phages results in the rapid distribution of phage into the blood stream and spread throughout the organism[170,171]. However, phage localisation in various organs, including kidney, lungs and brain, tends to be approximately 100-fold lower when compared to the blood after 3 hours of intravenous therapy administration[170,171]. Furthermore, after 12 hours of administrations, no phages can be detected in the blood, kidney, brain lung, however phage particles can still be recovered from liver, intestines and in significantly higher amounts from spleen[172]. When administered orally, it was found 10^3 PFU/ml of phage are excreted in the faeces for mice as well as human volunteers, implying that phage can cross the acid stomach barrier[173,174]. The phage clearance is facilitated by the immune system as it triggers an immunogenic response (see section 1.11)[175–178]. Greater understanding of phage host interactions while in the human body is required to precisely determine the effective phage concentration at a site of infection i.e. organ or a tissue site. In the recent studies, a precise number for phage concentration has been deduced for the minimum phage concentration required to prevent bacterial culture from spreading[179,153,180]. The required concentration for mildly growing bacterial infections is $\geq 10^7$ phages/ml, however, more actively growing infection may need higher phage concentrations[180]. In order to clear bacteria completely concentrations $\geq 10^8$ phages/ml are required[180]. Therefore, phages would need to be applied directly to the target tissue or alternatively ways of localising them in specific tissues and organs are required[181]. The human lung is exposed to airborne pathogens on the regular basis which can lead to a formation of chronic and severe infections in the airways[182]. In addition, the human lung causing infections were identified to commonly develop resistance to their common treatments[183]. In chapter 7, the lung epithelial cells were chosen as a case study to identify if phage engineering can be used to increase effective phage concentration at an organ of interest.

1.12 Phage Immunogenicity

There is increasing evidence that while in the blood, phages cause an immune response[175–178]. Immunogenicity of phages has been exploited for the evaluation of humoral (associated with the body fluids) immunity in diagnosing and observing patients suffering from immunodeficiencies[176,184]. Phage ΦX174 is able to trigger different levels of humoral response after intravenous injection to patients[184]. The use of ΦX174 is accepted as one of the standard antigens for humoral immunity determination[185]. It has been showed that antibodies pre-existing in animal serum are able to recognise phages causing initiation of complement activation[186]. It has been showed that complement binding can be overcome by modifying phage T7 capsids with carboxy-terminal arginine or lysine[186]. The amino acid residue substitution protects T7 against C-reactive protein binding and hence phage inactivation via complement mediation in rats[186]. In a further study *E. coli* phage λ and *Sa. tyhpimurium* phage P22 were examined in germ-free mice[187]. A single mutation resulting in the substitution of lysine with glutamic acid, yielded long-circulating phage phenotypes for both phages[187]. A short-circulating phage phenomenon has been described for phage T4, whereby the loss of nonessential capsid protein Hoc results in the removal from a mouse circulatory system faster than the wild type[188]. It was suggested that some phage proteins are able to regulate interaction with the immune system[188,189]. In further studies the blood plasma clearance time for phage M13 was also examined[190]. It was determined that conjugation of either succinic acid or galactose groups to phage M13 capsid proteins can lead to a reduced half-life of the phage in blood plasma[190].

1.13 The release of bacteriophage progeny

There are two known strategies of bacteriophage release from their host. The first one consists of filamentous phages (e.g., MI3, fd) that evolved a mechanism allowing phage production without killing the bacterial cells[191]. In this instance, instead of phage particles forming in the cytoplasm, they are constantly removed by secretion across the host membranes as soon as they are assembled[192,193]. All phage structural proteins are mobilised in the host inner membrane before they are assimilated into phage particles[192,194].The second strategy achieves progeny release while damaging the host membrane and cell wall eventually causing host lysis, killing the bacterial host[43]. This second strategy is employed by non-filamentous phages. The disruption of the bacterial cell wall can come in two forms, inhibition of the peptidoglycan synthesis (small ssDNA or RNA phages) or cleavage of peptidoglycan by a holin-lysin system (large dsDNA phages)[195]. Holins and lysins (lysins) are phage proteins that are expressed during the late phase of gene expression[195]. The two protein system operates by the holin forming breaks that allow the lysin to reach and cut peptidoglycan[195,196]. Phage λ, in particular, was used to study holin-lysin interactions. Upon its infection, phage λ can be induced to allow a timed lysis event[195]. Both holin (S) and lysin (R) genes, are located at the start of the late transcriptional region, which is normally activated eight minutes into the infection cycle[195]. The active lysin concentrates in the cytosol until the holin forms a membrane abrasion allowing the lysin to access the peptidoglycan while crossing the cytoplasmic membrane (Figure 1-2)[195]. In Figure 1-2, the fate of the fate of Gram-negative bacterium envelope during holin-endolysin lysis is described. Prior to hole formation lysins concentrate in the cytosol whereas holins tend to aggregate in membrane in clusters. This is then followed by holin forming a membrane break allowing lysin to attack the murein. One of the traits of the holin based systems is their dependency on the energized membrane; addition of an energy providing substance triggers induction of an early lysis[43,197]. The data available indicates that any holin will work with any lysin when holin-lysin pair is made from any Gram-positive and Gram-negative bacteria[198]. Furthermore, it has been shown that if applied externally, the lysins are capable of degrading peptidoglycan layer causing lysis of Gram-positive bacteria due to their ability to make a direct contact with peptidoglycan[199–201].

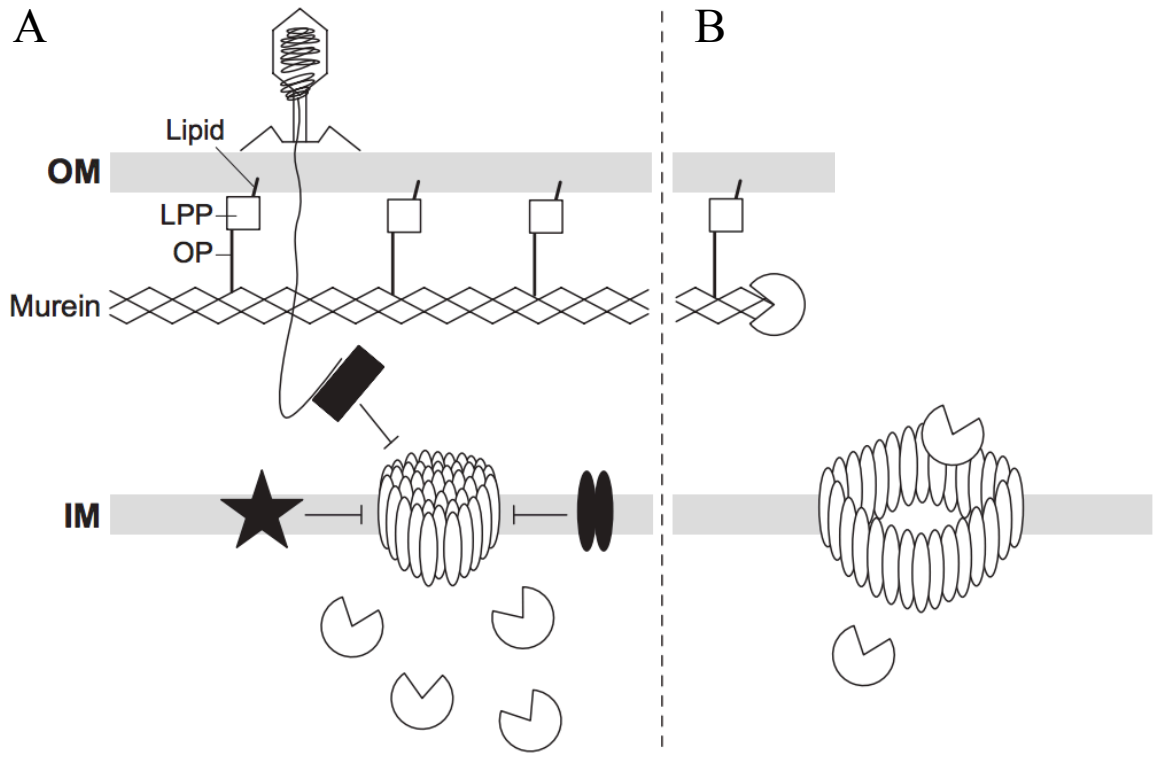


Figure 1-2. Gram-negative bacterium envelope fate during holin-endolysin lysis. The envelope is made up of inner (IM) and (OM) membranes, murein, that is attached to the OM via oligopeptide (OP), lipoprotein (LLP) and lipid chain. **(A)** Prior to hole formation. Lysins (notched white circles) concentrate in the cytosol whereas holins (white ovals) aggregate in membrane in clusters. **(B)** Hole formation. Membrane break is formed by holin facilitating lysin escape allowing them to attack the murein. The image was adopted from a previous study[195].

1.14 Bacteriophage T4

Bacteriophage T4 is a member of the Myoviridae family of the order *Caudovirales* due to its contractile tail and can infect only *E. coli*, closely related *Shigella* species as well as a non-enterobacterial vibrio species and *Yersinia*[202–206]. The main T4 receptor is the outer membrane lipopolysaccharide (LPS)[207]. The T4 genome is made up of 168 kbp and it contains 289 open reading frames[202]. T4 main structural features are the capsid, tail, baseplate, LTFs and STFs[208]. The mature phage has a 115 nm long and 85 nm wide head that encages the DNA[208].The T4 contractile tail is 92.5 nm long and 24 nm wide and is attached to a portal vertex located at the end of the head. The distal end of the tail is attached to a 27 nm height and 52 nm width hexagonal baseplate[208]. Six LTFs that recognise host receptors are attached to the baseplate[208], with six STFs located underneath the baseplate[208]. The capsid is made up of 155 hexamers of the major protein, gene product 23[209,210]. At its surface the capsid also possesses two accessory proteins, Hoc (highly antigenic outer capsid protein) and Soc (small outer capsid protein)[211]. A cascade of transcriptional events is started by T4 infection of *E. coli* via LPS[212]. Firstly, the T4 early genes are transcribed post infection using the host RNA polymerase[202]. Since the early genes' promoters have good matches for σ^{70} sequences, they do not need T4-encoded transcription factors[213]. In the case of middle T4 promoters, they require two transcriptional activators obtained from the early transcription: MotA and AsiA[214–216]. Furthermore, the late T4 promoters require T4 sigma factor, Gp55, in addition to other T4 specific activators and coactivators[212]. The induction of T4 DNA replication within infected *E. coli* is incredibly complex as it consists of the numerous circularly permuted linear copies of the phage genome that show as concatamers, during the middle and late steps of infection, launching the strand synthesis[217]. The T4 replisome has two parts, the DNA replicase as well as primosome. The DNA replicase consists of a phage encoded DNA polymerase (Gp43), the gene 32 encoded single-stranded DNA binding protein (Gp32), the genes 44 and 62 encoded ATP-dependent clamp loader complex (Gp44, Gp62) and the gene 45 sliding clamp (Gp45)[218]. ATP-dependent clamp loader complex (Gp44 and Gp62) make up an accessory protein that is necessary for gp45 loading onto DNA[219]. The Gp32 facilitates the unwinding of DNA whereas Gp43 DNA polymerase lengthen the primer of the invading strand into the next genome[220]. At the start of the elongation process consists of replication of the leading strand template where the Gp43 can repeatedly process a daughter strand. The primosome initiates the lagging strand sectional synthesis of Okazaki fragments[218]. The T4 replication complex is made up of the Gp61 primase

and Gp41 helicase, homohexamere that encompasses the lagging strand and process it in the 5' to 3' direction[221,222].

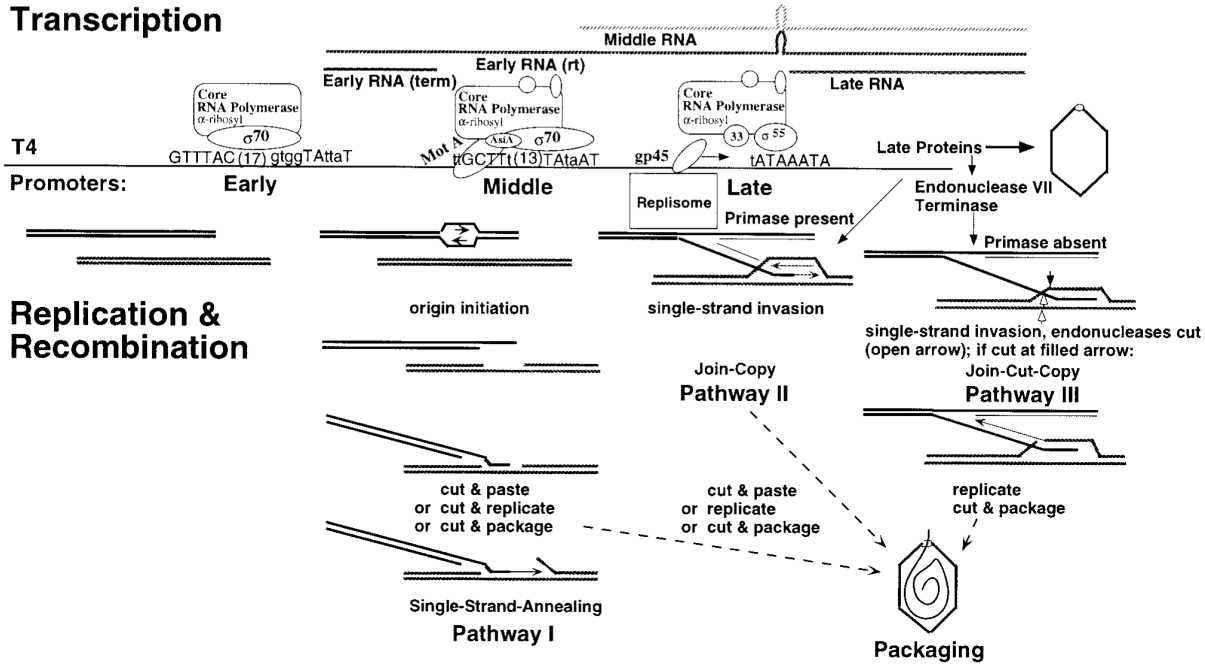


Figure 1-3. The above schematics represents the distribution of early, middle and late promoter transcription, as well as links the T4 transcriptional pattern with different mechanisms of DNA replication and recombination. The top part of the diagram presents the transcripts started from early, middle, and late promoters by host RNA polymerase. In some instances, hairpins in multiple early and middle transcripts inhibit the translation of the late genes. The bottom part of the diagram represents the pathways of DNA replication and recombination. The image was taken from the previous work by Miller et al. 2003.

1.15 Bacteriophage T7

Bacteriophage T7 is a member of the *Podoviridae* family and can infect *Escherichia coli, Shigella, Salmonella,* and selected T7 mutants can infect *Yersinia pestis*[223]. Phage T7 is double-stranded DNA (dsDNA) virus that has 39,937 bp long genome encoding for 56 genes[223]. Its main structural features consist of a 60 nm icosahedral capsid and a short tail (32 nm in length) [224,223,225]. The infection and replication cycle of phage T7 has been widely studied and is well characterised[226,227]. The first step of this process is adsorption and infection initiation[228–230]. Adsorption occurs when the tail fiber interacts with the outer membrane lipopolysaccharide (LPS)[223]. The binding is followed by the inner core injection from the capsid into the outer membrane periplasm, where it most likely degrades the peptidoglycan and forms a stable channel for DNA translocation[223]. About ~850 bp of the genome is thought to be ejected into the cell allowing the initiation of transcription by *E. coli* RNA polymerase that recognises the three early promoters A1, A2, and A3 that are located at the start of the genome[230]. The rest of the phage genome is pulled in by transcription of the T7 genome by the endogenous polymerases[230]. Gradual entry into the bacterial cell facilitates a form of gene regulation[223]. T7 possesses essential as well as non-essential genes that have been identified by making mutations at different locations in T7 genome[231]. The T7 gene transcriptome divides T7 genes into three classes[223]. The earliest genes transcribed, within the first 8 min, are referred to as Class I genes and are located at the start of the genome[232]. Class I genes constitute 9 genes, from gene *0.3* to *1.3*[231,233]. The two most important genes are gene *0.7* and gene *1*[234,235]. The product of gene *0.7* inactivates transcription of multiple host genes as well inhibits host proteins, including RNAse III and E, and *E. coli* polymerase subunits[234]. Gene *1* encodes T7 RNA polymerase that transcribes T7 class II and class III genes from T7 promoters[235–237]. Class II genes are transcribed at around 6 to 15 minutes after the initial transcription and mainly those genes that are required for T7 DNA replication[223]. One of the Class II genes is *g2* which is responsible for *E. coli* RNA polymerase inhibition[238]. *g2* expression results in the inhibition of transcription of the Class I genes as they are under control of host polymerase[238]. *g5* encodes for T7 DNA polymerase that replicates the T7 genome once it forms a complex with a *E. coli* host factor, (*trxA*)[239–241]. T7 Class III genes encode for proteins involved in DNA packaging, viron assembly and host lysis[223,242–246].

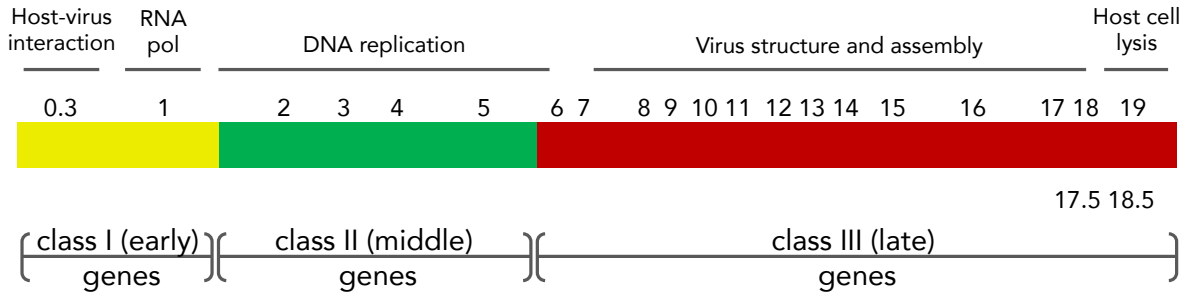


Figure 1-4. Representation of three classes of gene expression of bacteriophage T7. Class I consists of early, class II consists of middle and class III consists of late transcription genes. Class I generates host-virus interaction transcripts as well as RNA polymerase. Class II generates transcripts required for DNA replication, including DNA polymerase (*g5*). Class III generates transcripts required for virus structure and assembly including major and minor capsid proteins (*g10*), as well as host cell lysis proteins.

1.16 Phage T7 tail machinery

The T7 tail machinery is made up of the following proteins, Gp7.3 Gp8, Gp11, Gp12, and Gp17 (Figure 1-3)[247]. The most characterisation data has been gathered for Gp8 that assembles into a dodecamer and Gp17 that assembles into trimers forming the T7 tail fibers[244,248,249]. Gp11 and Gp12 assemble into the core structure of the tail[247]. The precise role of Gp7.3 remains unclear although it has been proposed that 32 subunits of this protein have role in the tail formation, either acting as a helper or contributing to the tail channel or even the tip[247,250].

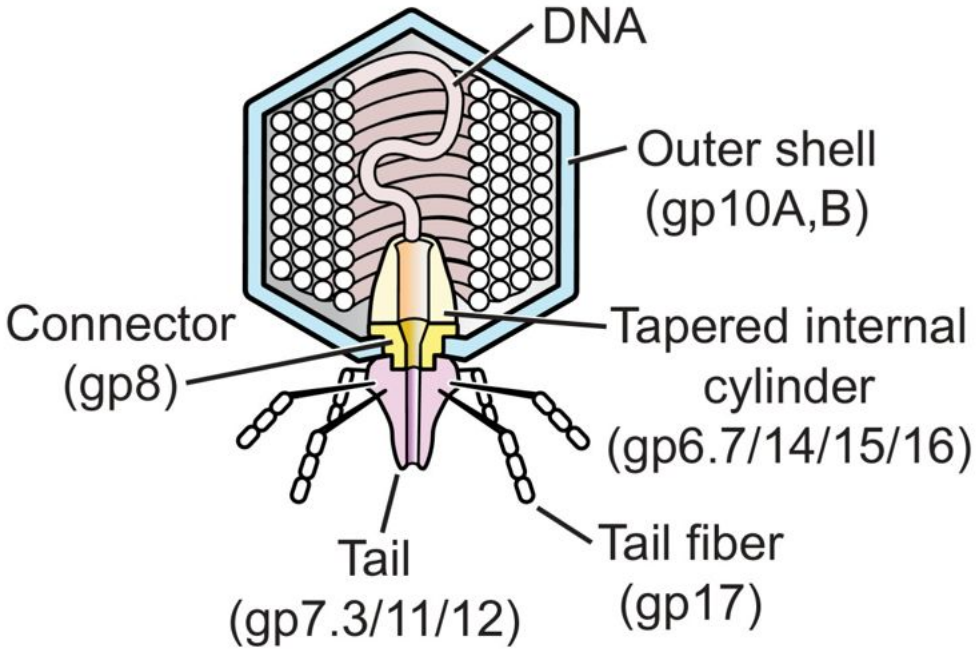


Figure 1-5. Schematic diagram of the T7 virion. The image was adapted from a previous study by Serwer et al. 2008. The outer shell is made up of Gp10A and B. The internal tapered cylinder (that is made up of Gp6.7, 14, 15 and 16) and tail (Gp7.3, 11, 12) are shown attached to the outer shell by a connector (Gp8). Tail fibers, that are trimers of Gp17, are attached to the tail.

1.17 Phage T7 DNA replication

The T7 replisome is formed of Gp4, Gp5, Gp2.5 and TrxA proteins[251,252]. First, the helicase domain of Gp4 unwinds double stranded DNA (dsDNA) resulting in single stranded DNA (ssDNA) that serves as a template for Gp5, DNA polymerase[227]. Gp5 forms a complex with its processivity factor TrxA (provided by the host bacteria) which then allows the leading DNA strand synthesis[227]. The lagging DNA strand is synthesised in fragments that are then processed to form a continuous strand[227]. The primase domain of Gp4 initiates the Okazaki fragments,

while Gp2.5, ssDNA binding protein protects ssDNA generated during replication as well as regulates the synthesis of both strands[241,253].

1.18 Thioredoxin

Thioredoxin is a 12-kDa redox-active protein expressed in various bacteria and mammals[254–256]. In the case of *E. coli,* it is a nonessential gene referred as *trxA*[254–256]. TrxA contains a universally conserved dithiol active site, as well as NADPH, flavoprotein and thioredoxin reductase (TrxR), TrxA serves as one of the main antioxidant systems in the cell[254–257].While being a non-essential protein in *E. coli*, TrxA is essential for the replication of T7, as it acts as a co-enzyme for T7 DNA polymerase (Gp5)[114,258]. TrxA binds to DNA polymerase via a thioredoxin binding domain (TBD) which is a loop of 76 flexible amino acids (See Figure 1-4)[259]. The interface between Gp5 and TrxA occurs at Cys-275 and Cys-313 of the TBD and Cys-32 and Cys-35 respectively[259].

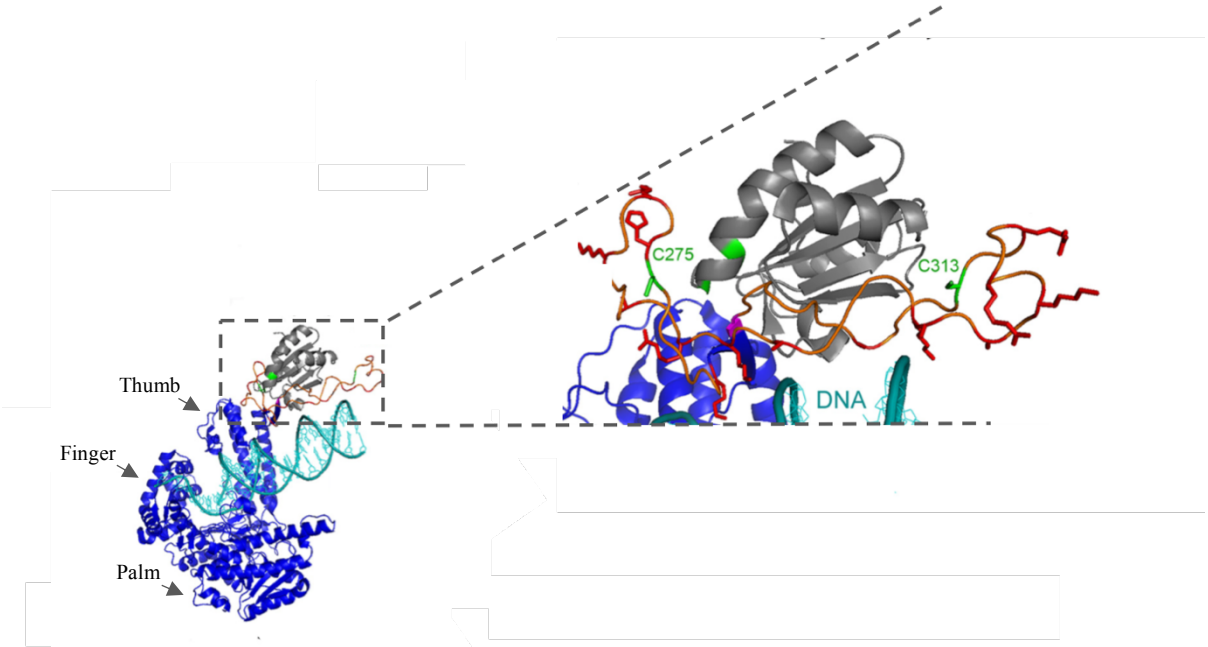


Figure 1-6. Structure of Gp5/TrxA complex on DNA. DNA, Gp5 and TrxA are presented in cyan, blue and grey respectively. Gp5 has three domains: palm, finger and thumb. TrxA is bound to Gp5 at a flexible loop of 76 amino acids long referred as TBD (shown in orange). The interference between TBD and TrxA occurs at TBD and TrxA at Cys-275 and Cys-313 of TBD and Cys-32 and Cys-35 of TrxA. The image was adapted from a previous study[259].

1.19 Phage T7 DNA polymerase processivity

TrxA binds to the thioredoxin binding domain on Gp5, in a 1:1 ratio, resulting in 20-80 fold greater DNA polymerase processivity due to reorganisation of Gp5 that allows it to better bind DNA[257,260–262]. It has been shown that upon TrxA binding to Gp5 a closed Gp5 confirmation is formed that results in greater Gp5 resistance to salt concentration changes in the environment as well as facilitation of more efficient Gp5 processing of the DNA strand[239]. The TrxA and Gp5 complex prevents from constant Gp5 dissociation from phage DNA stabilising Gp5 attachment to the DNA[239].

1.20 Lysins

Lysins, also called phage endolysins, are peptidoglycan hydrolases (PGHs) that are used by most phages to degrade peptidoglycan from within the cell[263]. Lysins fall into four main groups according to their mode of action against different covalent linkages present in the bacterial cell wall[195]. The four groups are amidase, endopeptidase (both target the oligopeptide cross-linkages), glycosylase and transglycosylase activities (both target the glycosidic linkages)[195]. It has been shown that lysins have no signal sequence and as a consequence tend to concentrate in bacterial cytosol during infection[195]. Due to the ability of lysins to disrupt the peptidoglycan layer in a very specific manner, lysins are treated as a novel class of antibacterial agents with a potential for pathogen killing circumventing effects on non-target species present in our microflora[200].

1.21 Lysin structure and the mechanism of action

The lysins that target Gram-positive and Gram-negative bacteria have distinguishing features due to the very different structure of these two bacterial groups[264]. Lysins targeting Gram-positive bacteria have two different types of functional domains that are called cell wall binding domains (CBDs) and enzymatically active domains (EADs)[201,265,266]. The EAD cleaves specific bonds of the bacterial peptidoglycan, whereas CBD facilitates the protein to substrate binding as well as tight binding to the cell wall[267]. In comparison, lysins targeting Gram-negative bacteria are generally small single-domain globular proteins and normally do not contain a CBD module[268,269]. Both Gram-positive and Gram-negative targeting lysins are not limited to two functional domains,

different domain combinations and orientations have been identified[270]. One of the examples is staphylococcal lysins that frequently have two N-terminal EADs and one C-terminal CBD[271,272].

Lysins consist of two domains—a peptidoglycan hydrolase activity is commonly located at the N terminus and can consist of one or more of a variety of glycosidases, amidases, or peptidases, while the lysin's specificity is achieved by a cell wall binding domain that recognizes cell surface features that are specific to the bacteria that it targets[273].

1.22 Lysins as antimicrobials

Phage lysins target genus or species-specific bacteria, thus are advantageous over current broad-spectrum antibiotics. Due to the fact that phages coevolved with their hosts, it is most likely that lysins adapted as well by binding and cleaving conserved cell wall targets circumventing a possibility of bacterial resistance, that would require a fundamental change in cell wall structure[201]. In general, the development of resistance against lysins has not been identified[274–276]. However, several studies suggested that intracellular resistance mechanisms including active efflux of compounds from cytosol as well as reduction in membrane permeability are employed to circumvent the action of lysins[277,278]. There are multiple examples where lysins have been employed as antimicrobials. In the first instance two pneumococcal phage lysins, differing in their target cleavage site, were examined *in vitro* against several serotypes of *Streptococcus pneumoniae,* including penicillin-resistant strains[279]. This resulted in synergistic cleavage of the bacterial strains with a potential for the prevention and elimination of pneumococcal colonization and infection. The Cpl-1 lysin was then introduced intravenously to already pneumococci bacteremic mice that resulted in 100% survival after 48 hours compared to 20% survival of the control animals[280]. It was also verified that despite lysins being immunogenic proteins, the efficacy of the Cpl-1 lysin against pneumococci colonised mice was not significantly altered due to previous intravenous exposure and antibody production[280]. In another study, the lysin PlyG which active against *B. anthracis* and other members of the *B. anthracis* cluster of *B. cereus* was identified[281]. Multiple rounds of PlyG exposure to the *B. anthracis* did not cause bacterial resistance suggesting that its peptidoglycan catalytic target cannot be easily altered by the bacterium to interfere with PlyG action[281]. PlyGBS is a recombinant streptococcal lysin that was shown to lyse group A, B, C, G and L streptococci[282]. It was originally developed for Group B streptococcal (GBS) vaginal colonization in pregnant women prophylactics with an intention

to reduce the rate of neonatal meningitis and sepsis in new-borns[282]. PlyGBS was used in mice models where a single dose of PlyGBS reduced cell titer by 1000-fold[282]. In another study lysin LysK from staphylococcal phage K showed activity against nine species of *Staphylococcus* strains, including vancomycin resistant *Staphylococcus aureus* (VRSA) and MRSA, isolated from both human and bovine sources[283–285]. Further lysin activities against target strains were identified, such as staphylococcal lysin LysH5 against *S. aureus* growing in milk, and Phi11 against mastitis-causing Coagulase-negative Staphylococcus[286,287].

1.23 Safety & Immunogenicity

Due to the proteinaceous nature of lysins they are biodegradable and noncorrosive[288]. Once systematically administered in animals they may cause a release of post lysis bacterial cellular debris such as lipoteichoic and teichoic acids, that may result in septic shock[289]. It has been showed there is an increase of proinflammatory cytokine concentrations in rats after a continuous treatment with Cpl-1 (a pneumococcal lysin)[290]. However, in another study the opposite effect was observed when applying the same enzyme every 12 hours[291]. This difference was explained by the fact that the continuous administration causes a greater generation of cellular debris leading to an increased immune response[201].

Several studies have shown that the antibodies can be potentially raised against lysins[201]. However, in animal models antibodies did not inactivate administered lysins and no side effects such as anaphylaxis were observed[292,293]. It has been reasoned that these findings are due to lysins having greater activity and response time when compared to the animal immune response[201,292,294,266].

1.24 CRISPR/Cas Systems

The abundant nature of bacteriophages has forced bacteria to develop a variety of target mechanisms to prevent phage infection, which target different steps of the phage life cycle[295–298]. Bacterial interference inhibits phage adsorption, its DNA injection as well as abortive infection[296,299,300]. It has been shown that one of bacterial defence mechanisms is an adaptive microbial immune system, consisting of clustered regularly interspaced short palindromic repeats (CRISPR), that provides an acquired defence against viruses and plasmids[301]. CRISPR is a family of DNA repeats whilst discovered in *Escherichia coli* in 1987, it was not formally described until 2000s after identification in different bacterial as well as archaeal genomes[302,303]. CRISPR loci are made up from multiple non-contiguous direct repeats that are disjointed by stretches of variable sequences, often referred as spacers[304]. Spacers represent fragments of integrated viral and plasmid sequences. Both direct repeats and spacers are positioned next to *cas* genes (CRISPR-associated)[304]. CRISPR together with Cas proteins make up the CRISPR/Cas systems. Six main *cas* genes have been determined[305–307]. CRISPR direct repeats are highly conserved for a particular organism, however, it varies across different bacterial species and their sequence range between 23 to 47 base pairs bp[301,303,308]. The CRISPR spacer regions are hypervariable, even between related strains and their sequence range between 21 to 72 bp[301,309–311].

In a nut shell CRISPR/Cas functions in the following way. Firstly, once the genetic invader enters the bacterial cell, a fragment of its DNA sequence (called protospacer or precursor, spacer) is integrated in the CRISPR loci as a spacer sequence (see Figure 1-5)[298]. The frequency of spacer acquisition in bacterial population may be affected by a variety of environmental and other factors including quorum sensing, certain catalytic enzymes (encoding adenylate cyclase)[312–315]. A promoter found in the leader sequence, transcribes the repeat-spacer region into CRISPR RNA (pre-crRNA) molecule (see Figure 1-5). The pre-crRNA is the processed by multiple maturation events that yield singular mature crRNAs composed by a fragment of a repeat and spacer regions. These crRNAs operate as guide RNAs that recruit and direct Cas proteins that help cleave target sequences (see Figure 1-5). More details about two specific CRISPR/Cas subtypes (CRISPR type I and II) are given in Chapter 3.

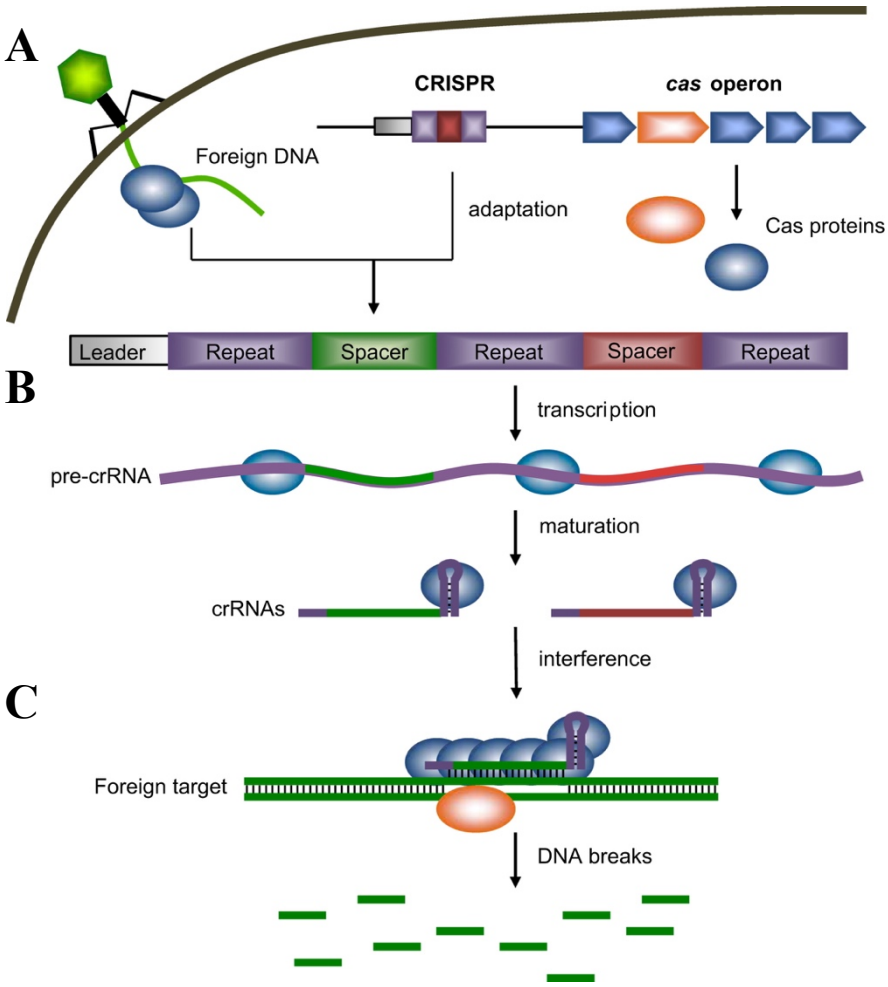


Figure 1-7. Summary of CRISPR/Cas immunity to viruses of bacteria and archaea. **(A)** Initially, phage infection is identified by Cas proteins and a short sequence of phage DNA (spacer) is incorporated at the end of the CRISPR array, that generates a new spacer sequence and a duplicated repeat. The *cas* operon encoding all Cas proteins is shown. It is located closely to the CRISPR array that is comprised of the leader sequence followed a multiple repeat/spacer units. **(B)** After the initial infection CRISPR array is transcribed from a promoter within the leader sequence generating a precursor CRISPR RNA (pre-crRNA) transcript. The pre-crRNA transcript is then processed into individual crRNAs. **(C)** The mature crRNAs target DNA sequence that is complementary to the spacer sequence in the crRNAs. In some instances, a Cas nuclease (represented in orange oval) is employed to interfere and aid in destruction of the target sequence. The figure has been adapted from the previous study[298].

Genetic engineering has exploited CRISPR/Cas systems for specific targeting and removal of desired genetic elements, including harmful genes and antibiotic resistance markers[301]. As well as acting as an immunity system, CRISPR/Cas targeting RNA may affect the transcript stability of chromosomal elements (Figure 1-6)[301].

Recent research has brought more in depth understanding about CRISPR/Cas mechanism and origin, however, multiple aspects are yet to be uncovered such as the pathway involved in spacer recognition and integration.

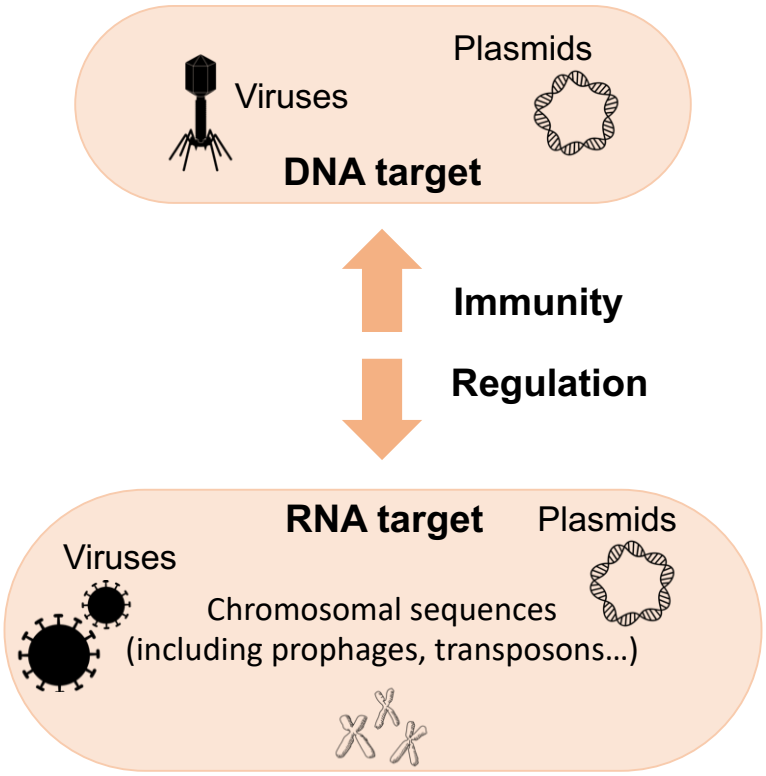


Figure 1-8. Summary of CRISPR interference. The CRISPR/Cas systems target can target DNA sequences impeding particular nucleic acid sequences (such as plasmid DNA or phage). This can be adapted to target genes of interest via genetic engineering. The CRISPR/Cas can also target RNA sequences which in turn can affect the transcript stability of chromosomal elements. The image has been adapted from the previous study by Horvath & Barrangou, 2010.

1.25 Aims and Objectives

Given the advances in biotechnology, diagnostics, molecular biology and synthetic biology wild type phage limitations can be addressed using genome manipulations bringing phage therapeutics closer to the clinical setting. The aims of this study concern first establishing the most efficient way for bacteriophage engineering followed by the use of the determined technique to address multiple limitations of wild type phages. Therefore, the aims were:

1) To determine the most efficient way of bacteriophage engineering and selection by comparing marker-based and marker-less methods (Chapter 3).
2) To examine if the host range of phage T7 can be altered by fusion of its tail fibers with those from genetically distant phage tail fibers (Chapter 4).
3) To determine if T4 can lyse an additional, to its host *E. coli*, pathogen by engineering it to deliver a cargo i.e. lysin targeting *S. aureus*, *B. subtilis* and *C. difficile* (Chapter 6).
4) To examine whether T7 engineered with tissue targeting peptides could allow increased adsorption/absorption to specific mammalian cells (Chapter 7).

Chapter 2 Methods

2.1 General

The methods presented in this chapter are used throughout this thesis. More chapter specific methods are presented in each of the chapters. This chapter is subdivided into general, bacteria specific, phage, tissue culture and protein analysis sections. In all protocols, O/N (overnight) refers to approximately 12 hours of incubation.

2.1.1 List of Strains/Phages plasmids

Strain/Plasmid/phage	Genotype	Comments/Culturing conditions	Reference
E. coli BW25113	*lacI*$^{+}$*rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR* 514 Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$ *rph-1* Δ*(araB–D)567* Δ*(rhaD–B)568* Δ*lacZ4787*(::*rrnB-3*) *hsdR514 rph-1*	LB, 37°C	Received from Jaramillo Lab, University of Warwick
E. coli MG1655	K-12 F^{-} λ^{-} *ilvG*$^{-}$ *rfb-50 rph-1*	LB, 37°C	Received from Jaramillo Lab, University of Warwick
E. coli BW25113*ΔtrxA*	F-, *Δ(araD-araB)567*, *ΔlacZ4787*(::rrnB-3), *λ*$^{-}$, *rph-1*, *ΔtrxA732::kan*, *Δ(rhaD-rhaB)568*, *hsdR514*	LB, 37°C	KEIO collection[258]
E. coli BW25113*Δcmk*	F-, *Δ(araD-araB)567*, *ΔlacZ4787*(::rrnB-3), *λ*$^{-}$, *Δcmk-734::kan*, *rph-1*, *Δ(rhaD-rhaB)568*, *hsdR514*	LB, 37°C	KEIO collection[258]
E. coli BW25113*Δwaac*	F-, *Δ(araD-araB)567*, *ΔlacZ4787*(::rrnB-3), *λ*$^{-}$, *Δwaac k-734::kan*, *rph-1*, *Δ(rhaD-rhaB)568*, *hsdR514*	LB, 37°C	KEIO collection[258]
E. coli BL21	*E. coli* B F^{-} *ompT gal dcm lon hsdS*$_{B}$(*r*$_{B}^{-}$*m*$_{B}^{-}$) [*malB*$^{+}$]$_{K-12}$(λ^{S})	LB, 37°C	Received from David Scanlan Lab, University of Warwick
Bordetella pertussis Tohama I	WT	37°C, charcoal	Received from Andrew Preston Lab, University of Bath
Bordetella bronchiseptica RB50	WT	37°C, LB and charcoal	Received from Andrew Preston Lab, University of Bath
pSMART	N/A	N/A	IDT
pWUR397	N/A	N/A	Brouns et al. 2008
pWUR400	N/A	N/A	Brouns et al. 2008
pCas9	N/A	N/A	Marraffini Lab, The Rockefeller University

pET30+	N/A	N/A	Mark van Raaij Laboratory
pMAT	N/A	N/A	IDT
pET24-gp5	N/A	contains gp5	Received from Jaramillo Lab, University of Warwick
pSEVA551	N/A	Contains codon optimised gp17	Received from Jaramillo Lab (made by Paul MacDonald), University of Warwick
Phage T7	WT	N/A	Richardson Lab, University of Harvard
T4 like phage	Genome 95% similar to the WT T4	N/A	Isolated by Pavelas Sazinas, sequence attached in the CD
Q12 T7	Amber mutation in 12th amino acid of Gp17 of WT T7	N/A	Mark van Raaij Laboratory Unpublished data

2.1.2 List of buffers

Buffer/media/Agar	Constituents	Comments
SM	50 mM Tris-HCl, pH 7.5, containing 100 mM NaCl, 10 mM MgSO4, and 0.01% gelatin	Used for phage storage
DMEM	Potassium Chloride (KCl) – 400 mg/L, Sodium Bicarbonate (NaHCO3) – 3700 mg/L, Sodium Chloride (NaCl) – 6400 mg/L.	Used for mammalian tissue maintenance.
Fetal bovine serum FBS (10%)	Bovine IgG, ≤1 mg/mL Hemoglobin, ≤20 mg/dL	Used for mammalian tissuemaintenance.
TBE x 5	54 g Tris base and 27.5 g boric acid dissolved in 900 ml of water	
SOC	2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose.	
PBS	(x1) 8 g of NaCl. 0.2 g of KCl. 1.44 g of Na_2HPO_4. 0.24 g of KH_2PO_4. Add distilled **water** to a total volume of 1 liter.	
TBS-Tween	50mM Tris, 150mM NaCl, 0.1%Tween 20	
Chromoagar Orientation selective medium (ChroMagar, UK)	N/A	Used for strain verifications
Charcoal Agar	Peptone – 10g, starch – 10g, charcoal – 10g, sodium chloride – 5g, nicotinic acid – 0.001g, agar.12g (for 1 litre)	Used for Bordetella strains maintenance.

2.1.3 Agarose gel electrophoresis

1% (w/v) agarose (14170829, Geneflow Limited, UK) gels were made and run using 1 x TBE buffer (section 2.1.2). Gels were run at 100 V with for 40 mins using Electrophoresis power supply (PowerPac^{TM}, Biorad, USA). Gels were stained with SybrSafe (S33102, Thermo Fisher Scientific, USA) used to visualize DNA bands; it was added to final concentration of 1:1000 to molten agarose gel. DNA samples were mixed with 6 x loading buffer (New England biolabs (NEB) B7024S, UK), prior to loading on to the gel. All samples were run alongside a DNA size marker (SM0313 Generuler^{TM} 1 kb DNA ladder, Thermo Fisher Scientific, UK).

2.1.4 Gel extraction

After running gel electrophoresis (see section 2.1.3) the fragment of interest was cut out from the gel and placed in a 1.5 ml Eppendorf tube (Eppendorf, Germany). 390 µl of QG buffer (28506, Qiagen, Germany) was added and dissolved in 50°C for 10 min. 130 µl of isopropanol (Fisher Scientific, USA) was added and then the entire sample was transfer to a purification column (28506, Qiagen, Germany). This was then centrifuged (5810 R, Eppendorf, Germany), for 1 min, followed by a 750 µl of PE wash (28506, Qiagen, Germany). It was then centrifuged twice (removing the overflow after the first centrifugation). This was the followed by elution of the DNA using elution buffer (28506, Qiagen, Germany).

2.1.5 PCR - Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) was carried out in 0.2 ml volume Eppendorf tubes in a thermal cycler (T100 Thermal cycler, Biorad, USA). HF Phusion Master mix (M0531S, New England Bio Labs, UK) and BSA (B9000S, New England Bio Labs, UK labs) were used for the reactions. The volume and contents of the final reactions are described below (Table 2-1). PCR amplifications were performed with cycling profile as described below (Table 2-2).

Table 2-1. PCR reaction components.

Component	25 µl Reaction	50 µl Reaction	Final Concentration
10 µM Forward Primer	1.25 µl	2.5 µl	0.5 µM
10 µM Reverse Primer	1.25 µl	2.5 µl	0.5 µM
DMSO (optional)	(0.75 µl)	(1.5 µl)	(3%)
2X Phusion Master Mix	12.5 µl	25 µl	1X
Template DNA	variable	variable	< 250 ng
Nuclease-free water	to 25 µl	to 50 µl	

Table 2-2. Thermocycler conditions for a PCR reaction.

Step	Temperature (°C)	Time	Cycles
Initial Denaturation	98	30 sec	1
	98	5-10 sec	25-35
	45-72	10-30 sec	
	72	15-30 sec per kb	
Final Extension	72	5-10 min	1
Hold	4	10 min	1

2.1.6 PCR product clean-up

Single product PCR reactions were purified using PCR purification kit (28106, Qiagen, Germany) following the manufacturer's instructions. The reaction products were eluted in 20-40µl of elution buffer (28106, Qiagen, Germany).

2.1.7 Gibson Assembly

Gibson assembly was used as described previously[316]. Fragments required for the assembly were PCR amplified using primers detailed in Appendix 2. Gel electrophoresis was followed by gel extraction (see sections 2.1.3 and 2.1.4 respectively). Each reaction was set up to a total volume of 20 µl (see the Table 2-3 below). For the assembly of three or more fragments an equilmolar ratio of fragments was used; when inserting one fragment into a vector, a ratio of 1:2 (vector to insert) was used. The reaction was incubated at 50°C for 30 min when assembling 2 fragments or 60 min when assembling 3 fragments and then incubated at 37°C for 60 min. When using the Gibson Assembly Master Mix (section 2.1.8) for electroporation, it was diluted 3-fold and 1 µl was used in a transformation reaction (Section 2.2.2). The Gibson Assembly Master Mix was stored at -20ºC.

Table 2-3. Gibson assembly reaction components.

Components	Amount /µl
Vector	X
Fragment 1	X
Fragment 2	X
Gibson Master Mix*	15
Make up with H20 to 20 µl	X

*see section X

2.1.8 Gibson Assembly Master Mix Preparation

8 ml of 5X ISO Buffer (Table 2-4) was prepared in a 15 ml falcon tube as described in Table 2-4. Gibson assembly master mix was prepared as described in Table 2-5. The 5X ISO buffer was stored at -20°C. A 60 µl aliquot of the Gibson assembly master mix was thawed and kept on ice until used.

Table 2-4. 5X ISO buffer preparation components.

Reagents	Amounts
1M Tris-HCl pH 7.5	4 ml
2 M $MgCl_2$	200 µl
100 mM dNTP mix (25 mM each: dGTP, dCTP, dATP, dTTP)	320 µl
1M DTT	400 µl
PEG-8000	2 g
100 mM NAD	400 µl
dH_20	to 8 ml of total volume

Table 2-5. Gibson assembly master mix preparation components.

Reagents	Amounts
5X ISO Buffer	320 µl
10 U/ µl T5 exonuclease	0.64 µl
2 U/ µl Phusion polymerase	20 µl
40 U/ µl Taq ligase	160 µl
dH_20	to 1.2 ml of total volume

2.1.9 Golden Gate Cloning

The Golden Gate protocol was adapted from a previous study[317]. Initial step for the golden gate protocol required primer preparation as follows. 0.5 µl of each primer (pair of annealing primers at 100 mM concentration) was mixed together with 49 µl of sterile water. The mixture of both primers was incubated at 95 °C for 5 min. The primer mixture was then diluted ten times. The Golden Gate reaction was set up as in Table 2-6. The assembly was then performed using the thermocycler with the conditions presented in Table 2-7. The samples were then left O/N at room temperature and transformed the following day using electro-transformation protocol (section 2.2.2).

Table 2-6. Golden Gate reaction components.

Reagent	Amount/ µl
vector (approx. 100ng)	1
Primer mixture	2
10X NEB T4 Buffer	1.5
100X BSA	0.15
BsaI	1
NEB T4 Ligase, 2 million cohesive end units / µl	1
Water*	to make up the total of 15 µl

Table 2-7. Golden Gate reaction thermocycler conditions.

Step	Time (min)	Temperature (°C)	Cycles
1	3	37	25
2	4	16	
3	5	50	1
4	5	80	

2.2 Bacterial methods

2.2.1 Colony PCR

A single colony was picked from a plate and resuspended in 50 µl of sterile water, by vortex for 30 sec in a 200 µl Eppendorf. For PCR, 2 µl of the resuspension was used to provide the template DNA in the PCR protocol (Section 2.1.5). The remaining resuspension was platted on appropriate agar for microorganisms and grown O/N.

2.2.2 Electro-transformation

Prior to use, the electroporation cuvettes (2 mm) (Thermo Fisher Scientific, USA) and DNA samples in Eppendorf tubes were chilled on ice for 10 min and LB-agar plates were placed in a 37°C incubator for 30 min to warm. Once the cuvettes were cold the competent cells were removed from the -80°C freezer and thawed on ice for 10 min. 1-2 µl of purified DNA for low and high plasmid copy number, was added to 50 µl of previously prepared electrocompetent cells and mixed by gentle swirling, prior to transfer into a electroporation cuvette. The electroporator (Gene Pulser, Biorad, USA) was set to 2.5 kV and the cells shocked. SOC medium (Section 2.1.2) was immediately added to the cells, prior to transferring the cells to a 1.5 ml Eppendorf and chilled on ice for 2 min. The cells were then incubated for 1 hr at 37°C, prior to platting 100 µl of cells on a pre-warmed LB plate with an appropriate antibiotic and incubated O/N.

2.2.3 Growth Curve

A single colony was picked off a plate and grown O/N in 5 ml of medium at 37 ºC whilst shaking at 200 rpm. The O/N bacterial culture was diluted 1% and grown until an OD600 of 0.2-0.3 was reached. 200 µl of each bacterial culture was loaded into 96-well plate. The OD was measured every 5 min, with 200 rpm orbital shaking in between the readings, using a plate reader (FLUOstar Omega, BMG LabTech, UK) at 37 ºC. The growth curves generated were then exported to an Excel and later analysed using Prism software.

2.2.4 Colony forming units (CFU) assay

The cells were inoculated O/N and 1:200 of the culture was diluted into an appropriate medium and grown until it reached to OD600 of 0.3-0.4 on the day of the experiment. Cells were then 10-fold serially diluted in Eppendorf tubes. 100 µl of each dilution was spread on an appropriate agar plate.

2.2.5 *E. coli* co-culture experiments with or without phage

The cells of two strains of interest (*E. coli* and wither *B. subtilis* or *S. aureus*) were separately inoculated O/N and 1:200 of the culture was diluted into an appropriate medium and grown until it reached to OD600 of 0.4 on the day of the experiment. Two of the strains were then mixed in equal ratios and grown for 4 hours with or without any phage added followed by their OD measurements every 30 mins. In the case of phage addition, T4-lysin phage was added at MOI of 0.1.

2.3 Bacteriophage Methods

2.3.1 Storage of Bacteriophage Isolates

For long term storage of bacteriophage isolates, 100-200 µl of lysate was diluted into 1 ml of SM buffer (2.1.2) and stored in cryovial (BR114831, Sigma-Aldrich, USA) at 4°C. For short term storage, phage lysates were diluted and stored in LB medium at 4°C.

2.3.2 Phage Lysate concentration

An appropriate volume of lysate was generated (see section 2.3.3). Amicon Ultra-50 (100,000 kDa MWCO, Merk Millipore) columns were used for phage concentration. The columns were filled with phage suspension and centrifuged at 4000 g for 20 min until most of the supernatant passed through the filter. This was repeated until all of the phage lysate was filtered. 0.5-1.5 ml of phage suspension was retained in the column. The media only filtrate was discarded and the concentration column was vortexed to release phage particles from the filter membrane. Phage filtrate was then collected into Eppendorf tube. The filtrate was then filtered using a 0.45um pore sized syringe filter to remove any potential bacterial contamination. Phage concentrate was subsequently stored at 4° C until further use.

2.3.3 Phage lysate preparation

O/N bacterial culture was diluted 1:200 with fresh medium; the final volume was experiment dependent, most frequently a volume of 10 ml of LB was used. When the culture reached an OD600 of 0.3, phage was added at MOI of 0.1. The culture was left to grown at 37° C, with shaking at 200 rpm until lysis occurred (liquid became clear, OD600 ≤ 0.1). 1 µl of chloroform was added and the cell debris was spun down at 4000 g for 10 min. The lysate was then filtered, transferred into a new tube and stored at 4° C until further use.

2.3.4 Phenol Chloroform-DNA extraction

0.4 ml of a phage sample was centrifuged at 13,000 g for 15 mins and the supernatant was transferred to a clean Eppendorf. This was followed by the addition of 0.4 ml of phenol and vortexing for 30 s. The supernatant was centrifuged for 10 min at 4°C, prior to extraction of the aqueous layer, which was mixed with an equal volume of phenol:chloroform (1:1) followed by vortexing for 30s and centrifugation at 13,000 g for 10 min at 4°C. Finally, the aqueous layer was extracted and mixed with an equal volume of phenol:chloroform:isoamlyalcohol (25:24:1) followed by vortexing for 30s and centrifugation for 10 min (13,000 g) at 4°C. The aqueous layer was then extracted and mixed with 1/10th volume of 7.5 M Ammonium Acetate and two volumes of ice cold ethanol were added and precipitated at -20°C O/N. A final centrifugation was carried out at 13,000 g for 30 min at 4°C. The DNA pellet was washed twice in 70% ethanol, air dried and then re-suspended in 10-30 µl of nuclease-free water.

2.3.5 Plaque assay

An overnight culture was diluted 1:200 in LB medium and grown until it reached an OD600 of 0.3-0.4 on the day of the experiment. Phage samples were 10-fold serially diluted in Eppendorf tubes. 300 µl of the culture at OD of 0.3 was added into 15 ml falcon tube followed by the addition of 100 µl of phage from the first dilution. The bacteria and phage mixture was incubated at room temperature for no longer than 10 min. 3 ml of 0.4% agar was added into the bacteria and phage containing falcon tube, which was then poured onto an appropriate agar plate (with/without antibiotics). The same procedure was repeated for each phage dilution. The plates were then incubated at 37° C O/N.

2.3.6 Spot Assay

O/N bacterial culture was diluted 1:200 into an appropriate medium and grown until it reached to OD600 of 0.3-0.4 on the day of the experiment. 300 µl of the bacterial culture were mixed with 3 ml of soft agar (0.4%) and poured onto the LB agar (1%) base layer in 90 mm diameter petri dish (if square plates were used, 1 ml of the culture was mixed with 9 ml of soft agar). The plate was left to dry for 10 -15 min. Phage dilutions were prepared in the same way as described in the plaque assay (see section 2.3.5). 5 µl of each phage dilution was then spotted on a quadrant of the plate and, after drying, the plate was incubated as required by the bacteria (Section 2.1.1).

2.3.7 One-step growth protocol

The cells were inoculated O/N and 1:200 of the culture was diluted into an appropriate medium and grown until it reached to OD600 of 0.3-0.4. 900 µl of bacteria was combined with 100 µl of phage providing a total concentration of MOI of 0.1. The mixture was incubated at 37ºC for 1 min. 10 µl of the mixture was then added to 25 ml of LB. 1 ml x 12 was aliquoted into 1.5 ml Eppendorf tubes and placed in thermomixer at 37ºC, shaking at 300rpm. Every 5 min, one of the tubes was removed from the thermomixer and mixed by vortexing. The sample was then processed either with spot or plaque assay, sections 2.3.6 and 2.3.5respectively.

2.3.8 Lysin kill spot assay

0.05% (v/v) of an O/N *E. coli* (Section 2.1.1) was inoculated into LB medium (Thermo Fisher Scientific, USA) and grown until and OD600 of 0.3-0.4 was reached. 0.3 ml of cells were combined with 3 ml of soft LB agar (0.4% w/v) and poured onto a 1 % LB agar plate and left to dry. 5 µl of T4-lysin mixture was spotted (Section 2.1.5) on to the lawn of cells.

2.3.9 Preparation of T4-lysin supernatant

The supernatant from the lysis of *E. coli* infected with phages T4 mutant phages (Table 6-2) was prepared by the infection of 10 ml of *E. coli* at an MOI of 1. Samples were taken every 3 hrs and immediately placed on ice, prior to spotting (Section 2.3.6).

2.3.10 Lysin Serial Transfer

Two cultures were prepared by inoculating 0.5 ml of O/N *E. coli* culture (Section 2.1.1) into 100 ml of LB medium and incubated at 37°C with shaking at 200 rpm, until it reached an OD600 of 0.2-0.3. Acriflavine (ACR) solution (Sigma Aldrich, USA) was added into one of the flasks at 1mg/ml concentration. Bacteriophage were added to both flasks at an MOI of 0.01. A 1 ml sample was taken every 30 min for 4 hr. Samples were centrifuged at 13,000 g for 15 min, prior to being filtered through a 0.22 µm pore size filter. 1 µl of sample was used as DNA template in PCR (2.1.5) to verify the presence of the lysin gene in both cultures (+/- ACR).

2.3.11 Homologous recombination

A culture containing homologous recombination plasmid (HR) was prepared by inoculating O/N culture at 1:200 into an appropriate medium and grown until it reached to OD600 of 0.3-0.4. Phage of interest was then added at MOI 0.01. This was then followed by incubation at 37°C at 200rpm for 3 hours if T7 phage, and 5 hours if T4 phage was used. The lysate was then processed as described in section 2.3.3.

2.3.12 Plaque PCR

A plaque of interest was stabbed with a pipette. The plaque residue collected on the pipette was then diluted in 100 µl SM buffer (section 2.1.2). 1 µl of this diluted residue was then used to perform PCR as described in 2.1.5.

2.3.13 Plaque purification

Plaque residue was collected as describe in section 2.3.12. It was then further diluted in 500 µl of SM buffer and used to make appropriate dilutions to perform plaque assays as described in section 2.3.5. The plaque collection, dilution and plaque assay were repeated three times to yield purified phage samples.

2.3.14 Bacterial lysis by phage

Bacterial culture was set up as in section 2.2.3, followed by addition of 20 µl an appropriate concentration of phage Optical density at 600 nm was measured and recorded periodically with a Microplate Reader for 8 hours.

2.3.15 Virulence assay

100 µl of O/N cultures were used to inoculate in 10 ml of LB and grown at 200 rpm at 37°C until 0.3 at OD 600 nm. Phage stocks were serially diluted from a concentration of 10^8 pfu/ml to 10 pfu/ml in a volume of 200 µl, yielding MOIs ranging from 1 to 10^{-7}. 100 ul of bacterial culture was combined with 200 µl of phage sample for each phage dilution in 96-well plates. Four wells of phage-free bacteria were included on every plate as control experiments, in addition to four media-blanks for reference. Optical density at 600 nm was measured and recorded periodically with a Microplate Reader. Measurements were taken until the control cultures reached stationary phase. Measurements were taken until the control cultures reached

stationary phase. Once all data was collected, areas underneath the optical density vs. time curves were calculated for each well, from the time of infection to the time corresponding to the onset of stationary phase in the phage-free control, using the trapezoid rule. The ratio between the area under the curve at one MOI and the area under the curve of the phage-free controls is termed the local virulence. This is obtained for each well (each MOI). Plotting the local virulence of the wells against the log of initial MOI (from -7 to 0) provides the graph for the virulence index. The virulence index is equal to the area under that curve. This protocol has been adapted from Storms ZJ, Teel MR, Mercurio K and Sauvageau D paper under review.

2.3.16 TEM prep protocol

To image phages of interest Transmission Electron Microscope (TEM) was used. 10 µl of high titre stock of phage was added to a glow discharged formvar copper grid (200 mesh), left for 2 mins, gently wiped off by allowing an edge of a filter paper to absorb access liquid, and 10 µl of water was used to wash the grid prior to being wiped off with filter paper. 10 µl of 2% uranyl acetate (w/v) stain was added to the gird and left for 30 secs, prior to its removal. The grid was air dried before imaging in a Joel 1200. The protocol adapted from Slawomir Michniewski's unpublished work.

2.4 Tissue Culture

2.4.1 Thawing A549

To thaw the frozen vial was placed in a 37°C water bath for approximately 2 mins. The vial was then removed from the water bath decontaminate by spraying with 70% ethanol. The vial contents were transferred to a 15mL centrifuge tube containing 5mL DMEM (Thermo Fisher Scientific, USA) and spun down at 1000 rpm for 5 mins using centrifuge (5810 R, Eppendorf, Germany), at room temperature. The supernatant was then removed and the cell pellet was re-suspended in 1mL DMEM. The hemocytometer was used to count the cells. The cells were then diluted to achieve a density of 5 x 10^5 cells/mL DMEM and transferred to T75 flask and incubated at 37°C.

2.4.2 Splitting of A549 cells

Prior to sub-culturing the cells, the fresh DMEM culture medium was placed in a 37°C water bath or incubator for at least 30 min. The medium in a T75 flask was removed using a pitpette and aspirator. 5 mL of sterile PBS (see section 2.1.2) was added into the flask followed by tipping the flask gently a few times. This was then followed by aspiration of the PBS. 2 ml of trypsin EDTA (see section 2.1.2) was added and incubated in 37°C incubator for 2 min. This was followed by addition of 8 ml of culture DMEM medium using this cell suspension, a required volume of cells into new flasks was pipetted at a required split ratio.

2.4.3 Adsorption/absorption assay

The following was carried out on 80-90% confluent cells, seeded two days prior to the experiment in a six well plate. The O/N medium in each well was replaced with 1 ml of DMEM containing the desired phage concentration. The plate was incubated either for one or four hours. The media was then removed and followed by the addition of 0.5 ml of PBS with gentle pipetting up and down. The 0.5 ml of PBS was collected from each well in 1.5 ml Eppendorf tubes. The PBS wash was repeated one more time. 1 ml of 0.5 % saponin (Sigma Aldrich, USA) was added to each well and the plate was incubated for 15 min at 37ºC. After the incubation, the saponin in each well was pipetted up and down followed by sample collection in 1.5 ml Eppendorf tubes. The samples collected after PBS and saponin additions were then processed either with spot or plaque assay, section 2.3.6 and 2.3.5 respectively.

2.5 Protein Analysis

2.5.1 Pertactin IPTG induction and phage-killing growth assays

Strains with the pertactin constructs of interest (see section 2.1.1) were grown O/N at 37°C. The cells were then re-inoculated into fresh LB medium and grown to at least 0.2 at $OD600_{nm}$. In each 96-well plate well 100 µl of cells were added to 100µl of media containing a given IPTG concentration and the appropriate antibiotics. During the phage killing experiments, the medium also contained the appropriate phage at an MOI of ~ 0.1. Growth was assayed by measuring absorbance at OD600 (FLUOstar Omega, BMG LabTech, UK) over a period of 5 h.

2.5.2 Preparation of protein samples

Strains with the construct of interest were grow O/N at 30°C. 200 µl of O/N culture was re-inoculated into 20 ml of pre-warmed LB medium and grown to between 0.2 and 0.6 OD600, then induced with 1 mM IPTG. 5 ml samples were taken prior to induction, then at 1 h and at 3 h post-induction. These were pelleted by centrifugation for 10 min at 5,000 g and stored at -80 °C before fractionation. The cell pellets were then separated as described by Charles et al. 1994 and summarised in 2.5.3[318].

2.5.3 Cell pellet separation

Cells were disrupted using 3 x 30 s sonification bursts followed by 30 s periods of incubation on ice. Unbroken cells, large cellular debris and insoluble materials were pelleted by low speed centrifugation at 3362 g for 10 min. Inner and outer membrane enriched fractions were collected by high speed centrifugation at 100000g for 45 min. The supernatant containing extracellular, cytosolic and other soluble components was also harvested. The membrane pellet was resuspended in ENV (10 mM sodium phosphate, pH 7.2) buffer containing 1% sodium (w/v) sodium lauroyl sarcosinate (Sarkosyl) and incubated at room temperature with frequent vortexing to aid the solubilization of the inner membrane. The insoluble outer-membrane fraction was re-pelleted by ultracentrifugation at 100000g for 90 min. The supernatant was the removed and the pellet resuspended in ENV buffer.

2.5.4 SDS-PAGE, Western Blotting and Membrane transfer

SDS-PAGE was carried out using 8-16% Mini-PROTEAN TGX Precast Protein gels (BioRad, USA). 2x Laemmli sample buffer (Biorad, USA) was mixed with pelleted samples at a ratio of 100 µl per 1 OD 600 and liquid samples at 1 ratio of 1:1. The samples were then boiled for 10 min, before loading 10 µl into the gel. The Colour Prestained Protein Standard, Broad Range (NEB, UK) was used to determine the size of the protein bands. The gel was run at 120 V for 20 min using power pack (Biorad, USA) then 180V for 50 min. The polypeptides were then transferred to a PVDF membrane (Biorad, USA) using the Trans-Blot Turbo Transfer System (BioRad, USA) according to the manufacturer's instructions.

2.5.5 Western Blotting

Rabbit anti-pertussis pertactin polyclonal antibodies (Alpha Diagnostic International, USA) were pre-incubated O/N at a concentration of 1:2500 on a gel containing only *E. coli* BL21 cells, together with blocking buffer consisting of TBS-Tween (TBST; Section 2.1.2) + 10% milk (Marvel, UK). The post-transfer membrane (Biorad, USA) was blocked for 1h in the blocking buffer, prior to addition of the primary antibody and O/N incubation. The next day the membrane was washed in TBST three times within 1h, and a goat anti-rabbit HRP-conjugated secondary antibody (Alpha Diagnostic International, USA) was incubated on the membrane for 1h at a concentration of 1:2500 in blocking buffer. It was then washed in TBST as above and incubated O/N in TBST. Finally, it was then visualised using the Western ECL Clarity Kit (BioRad, USA) in the ImageQuant LAS 4000, according to the manufacturer's instructions.

Chapter 3 Optimisation of Bacteriophage Engineering methods

3.1 Introduction

3.1.1 Current methods for bacteriophage engineering

A variety of methods have been used for the genetic modification of bacteriophage genomes. This includes the utilization of *Saccharomyces cerevisiae* and a Yeast Artificial Chromosome (YAC). A phage genome can be cloned into a YAC, in one piece or as PCR fragments assembled through gap-repair cloning, where phage DNA overlapping fragments are first mixed together with yeast/*E. coli* shuttle backbone allowing this DNA to be integrated into yeast spheroplast chromosome after transformation[319,320]. In particular, this technology enabled the genetic engineering of phages that naturally target *E. coli* to target pathogenic *Yersinia* and *Klebsiella* bacteria, and conversely, *Klebsiella* phage scaffolds to target *E. coli* by modular swapping of phage tail components allowing host range modifications[321]. One of the most prominent tools is homologous recombination which has originated as an *in vivo* analytical method used for bacterial gene manipulation[322]. Furthermore BRED (bacteriophage recombineering with electroporated DNA) has been used to genetically modify mycobacteriophages and coliphages[323,324]. BRED exploits the mycobacterial recombineering system, in which expression of the RecE/RecT-like proteins of mycobacteriophage Che9c confers high levels of homologous recombination and facilitates simple allelic exchange using a linear DNA substrate[325]. In BRED, phage DNA template and a targeting substrate are co-electroporated into *Mycobacterium smegmatis* cells that have been induced for recombineering functions[323]. BRED was used for the construction of unmarked deletions of both essential and non-essential genes, mutations, foreign gene insertions and gene tags additions[323,324]. The downside of BRED as a method is that it is limited to only temperate phage engineering and hence it will not be used for lytic phage engineering in this study[326].

Homologous recombination could potentially be used for engineering of the lytic phage genomes. If an exogenous gene is introduce in between homologous arms into a plasmid in *E. coli* that are then infected with wild-type phage T7, some of the phage progeny will undergo homologous recombination. However, this method does not provide a selection mechanism for recombinant phages, since there is no reliable way to couple the possession of a selection marker and the survival of the lytic phage. Using known antibiotic resistance genes when selecting for bacterial recombinants, is not possible as phage infection causes cell lysis[327]. In the absence of selection, either high frequency recombination systems are required or an examination of a large number of plaques created by potentially recombinant phages[323,326].

Therefore, a reliable method of selection is required for isolation of genetically engineered lytic phages using homologous recombination.

3.1.2 Marker-based selection

A selective marker can be used to select for recombinant phage after homologous recombination[114]. An antibiotic marker for a lytic phage like T7 cannot be used because it does not confer a selective advantage to the phage. However, a host co-factor gene that is essential to the chosen phage for replication and infection can be employed instead. Such examples in *E. coli* are thioredoxin and cytidine monophosphate kinase, encoded by *trxA* and *cmk* genes respectively, which are the only known co-factors essential for the replication of the phage T7, but are non-essential for *E. coli* growth[328]. The homologous recombination of a heterologous fragment with the genome of wild type T7 requires a plasmid containing the *trxA* or/and the *cmk* gene flanked by short T7 homologous sequences. Selection for recombinant phages is performed using *cmk/trxA* as a marker, since only phages that contain *cmk* or *trxA* will propagate in *cmk/trxA* deficient *E. coli* cell lines[1].

3.1.3 CRSIPR/Cas (marker-less) selection

An alternative to marker-based selection is a marker-less method that is based on CRISPR-Cas systems. As previously described in Chapter 1 (section 1.24) CRISPR is a type of adaptive immune response system in prokaryotes[300]. During homologous recombination, it is possible to exploit a bacterial immune system against phages to eliminate genome copies that have not undergone homologous recombination. The endogenous CRISPR system of some bacteria is appropriate for this because of its reprogrammability. In one report, the innate CRISPR-Cas Type II-A system of *Streptococcus thermophilus* was shown to successfully genetically modify the genome of the virulent phage 2972, by the introduction of mutations, large deletions and even the addition of a new gene by means of selective pressure to increase recombination efficiencies and allow more effective modified genome efficiency[329]. Another report used the I-E CRISPR-Cas system from *E. coli* combined with homologous recombination to target a non-essential gene[327]. Therefore, homologous recombination can be used to change the wild-type phage genome sequence followed by CRISPR targeting the wild-type sequence, favouring recombinant bacteriophage generation, acting as selection mechanism. T7 genes can either be essential or non-essential for T7 replication, thus non-essential gene sequences can potentially

avoid CRISPR targeting via CRISPR escape mutants (CEM) while not affecting phage vitality[327]. These escape mutants for non-essential genes can result in synonymous or non-synonymous mutations. This would not be the case for the essential genes as non-synonymous mutations are less likely to be tolerated by the phage and result in decreased phage infectivity.

As type I-E CRISPR-Cas is a Type-I CRISPR, in this study it will be referred as CRISPR type I; and as CRISPR/Cas9 in a Type-II CRISPR, it will be referred as CRISPR type II. An overview of each CRISPR can be found in Figure 3-1.

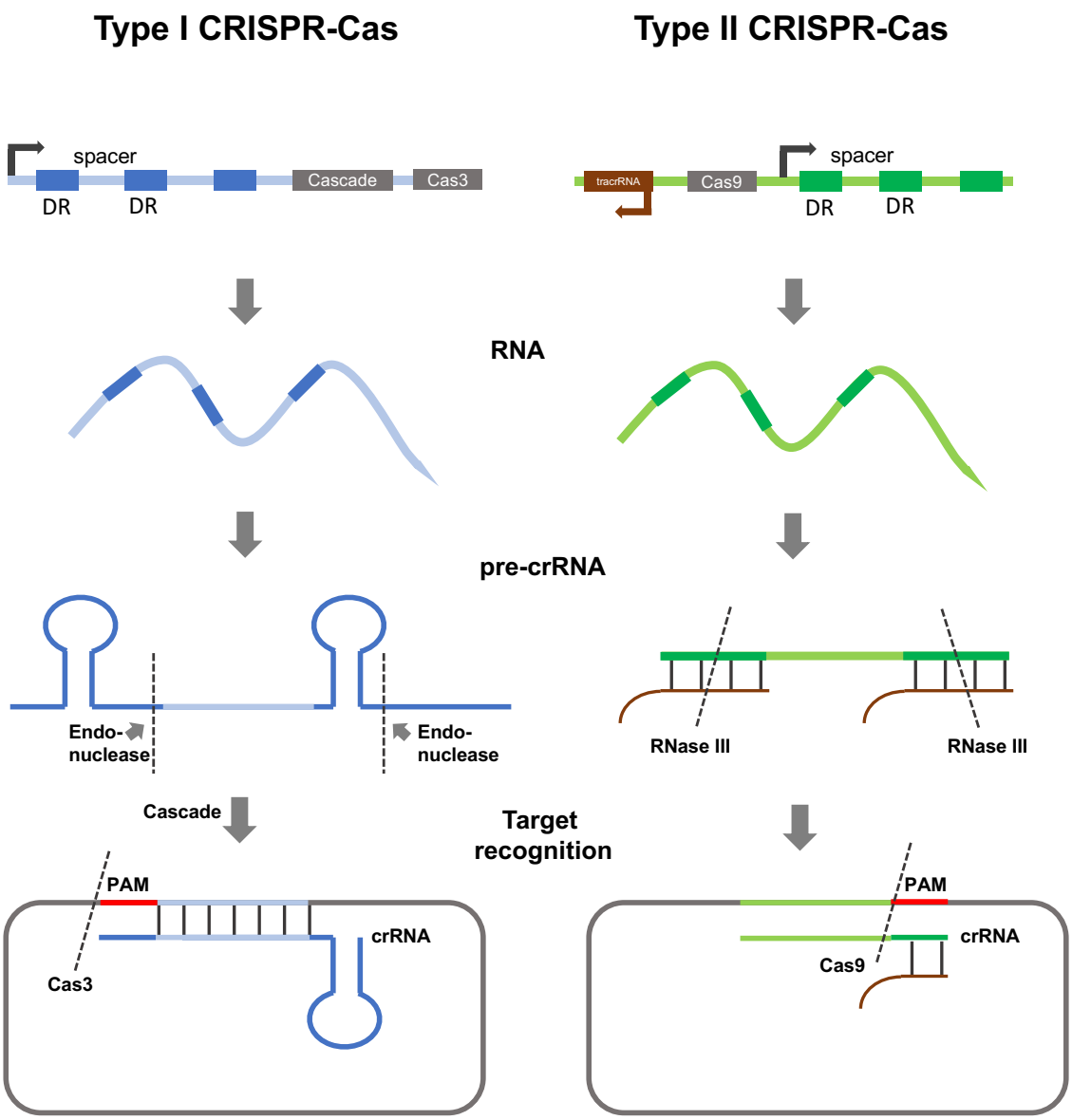


Figure 3-1. CRISPR arrays, comprised of direct repeats (DRs; dark blue and dark green) and spacer regions (light blue and light green) are first transcribed into a single large pre-crRNA. The resulting pre-crRNA transcript is then cleaved into individual crRNAs by the Cas6 endonuclease (one of Cascade enzymes, CRISPR Type I) or the ubiquitous RNase III enzyme (CRISPR Type II). Mature crRNAs are guided to invading nucleic acids through homology between crRNAs and the corresponding invader protospacer sequence. Type I and II interference mechanisms require recognition of one of multiple protospacer adjacent motif (PAM) sequences, which collectively make up the consensus PAM element (red). The image adapted from the previous study[483].

3.1.4 CRISPR-type I

The CRISPR type I system used in this study originated from *E. coli* K12 W3110 (BW25113)[330]. This system is comprised of three components: a five protein complex that consists of CasA, CasB, CasC, CasD, and CasE, often referred to as cascade, a Cas3 enzyme, and DNA repeat-spacer (a repeated sequence called 'repeats' and a specific target sequence called 'spacers')[305,330]. Once the DNA of repeat-spacer region is transcribed into RNA, the repeat regions form secondary structures (hairpins). This single large RNA transcript is now referred to as pre-crRNA[305,330]. It is cleaved by one of the cascade components, endonuclease, into individual crRNAs[305,330]. The remaining cascade enzymes aid in finding the target DNA sequence by first recognising the protospacer adjacent motif (PAM) and by a homology between cr-RNA and the target sequence[331]. Cascade then recruits Cas3 enzyme that cleaves and degrades the target DNA[331]. It has been shown that the I-E type CRISPR-Cas PAM motif for *E. coli* is AWG, and that out all of the possibilities AGG generated the highest targeting effieciency[332]. The other two essential parts for the crRNA are 3' and 5' handles (which are well defined repeat sequences before and after the spacer) that are on both sites of the pre-crRNA and later allow crRNA binding to Cascade subunits[306].

3.1.5 CRISPR type II

The type II system used in this study originated from *Streptococcus pyogenes*[333]. This system consists of four components: a Cas9 enzyme, DNA repeat-spacer, trans-activating CRISPR RNA (*tracr*RNA), ubiquitous RNAase II enzyme[334]. Once the DNA of the repeat-spacer region is transcribed into a single large RNA, the *tracr*RNA binds to repeat regions forming pre-crRNA. This pre-crRNA is then cleaved into individual crRNAs by ubiquitous RNAase II enzyme equivalent to Cascade in CRISPR type I. The Cas9 handle is formed of two stems in each side of crRNA. crRNA recruits and forms a complex with Cas9 protein. Cas9:crRNA complex finds the target DNA sequence by first recognising PAM and by a homology between crRNA and the target sequence. Cas9 then cleaves both strands of the dsDNA target[333].

3.1.6 Aims

Both marker-based and marker-less methods can be used for selection of engineered phage selection. In the first case the desired modification together with the selection marker (host factor) could be inserted using homologous recombination and then selected for in *E. coli* Δ*trxA* strains**.** The marker-less selection could be achieved by first introducing the desired genomic change using homologous recombination followed by CRISPR targeting wild-type sequence and allowing selection for engineered phage. Therefore, both methods must reduce the wild-type phage background.

The main aim of the work presented in this chapter is to compare the efficiencies of marker-based and marker-less methods in reducing wild-type progeny i.e. the selection efficiency for engineered bacteriophages. For the verification of both methods *g17* region, coding for T7 tail fibers will be used as a target region.

The aim of this chapter was divided into the following sub-aims:

1) To test CRISPR type I efficiency for selection of phage T7 mutants;
2) To test CRISPR type II efficiency for selection of phage T7 mutants;
3) To test the efficiency of marker-based selection using *trxA and cmk* for the selection of phage T7 mutants.

3.2 Methods

3.2.1 CRISPR type I gRNA delivery vector component/vector design

In a previous study the bacteriophage T7 genome was engineered using CRISPR type I, where all of the three components (Cas3, Cascade and gRNA) were placed on separate vectors to increase orthogonality of the system[327]. Therefore, all of the three components were required to replicate the system in this study. The enzyme cascade and Cas3 containing vectors, pWUR400 and pWUR397 respectively were obtained from the Udi Qimron Lab (Tel Aviv University)[335]. However, the third component of the system, a vector containing a gRNA to target specific sites of interest had to be designed and synthesized as part of this work. The layout of the construct design is presented in Figure 3-2 where only the designed fragment is presented. It was synthesized as a synthetic fragment and placed in pSMART amp vector (by IDT (Appendix 1 – plasmid map, Appendix 3 – sequence) and was used as a negative non-targeting gRNA construct, referred to as scrambled RNA (RNA (scr)) in the following chapters.

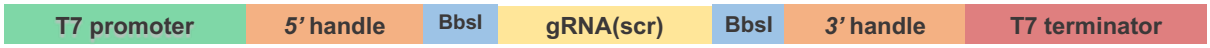


Figure 3-2. Schematic representation of the sequence inserted into pSMART gRNA delivery vector synthesized by IDT. of the gRNA delivery vector. The total length of the construct is 130 bp. It consists of a T7 promoter two handle regions, non-targeting gRNA(scr) with two surrounding BbsI cut sites and T7 terminator.

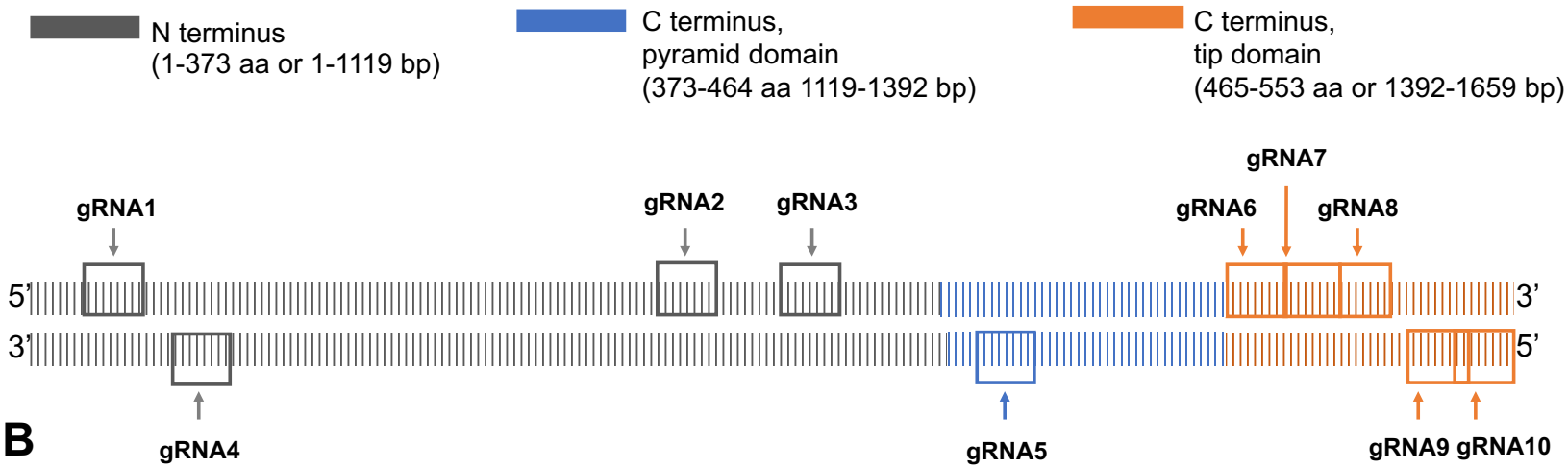
The main components to include in the design were: a) T7 RNA polymerase promoter at the 5'of the construct; b) 5' handle (responsible for Cas3 and Cascade binding); c) gRNA for a target gene sequence; d) restriction enzyme sites flanking the gRNA region; e) 3' handle (responsible for Cas3 and Cascade binding); f) T7 terminator. The T7 promoter was chosen to drive high expression of the gRNA construct upon T7 entry into a cell. A 5' handle repeat sequence (5' GAGTTCCCCGCGCCAGCGGGG 3') was chosen followed by eight well-defined bases of the repeat sequence (5' ATAAACCG 3')[330].

The Bbs*I* restriction recognition site was used to flank the target/spacer sequence for the Golden Gate Assembly (Section 2.1.9) to later replace gRNA(scr) and generate pSMART vectors with multiple target sequences. The 3' handle was constructed by adding 21 nucleotides (5'GAGTTCCCCGCGCCAGCGGGG 3') as described by other studies[331, 336].

3.2.2 CRISPR type I gRNA design

As described in the section 3.1.1 the gRNA delivery vector contains all of the requisite components for the target sequence transcription, processing and identification; however, it does not contain the actual target sequences i.e. gRNAs. Each of the CRISPR type I gRNAs for *g17* were designed manually taking into consideration PAM requirements for the system. Previous research has identified that the PAM that fits most efficiently is AGG[332]. The main principle for the design is to first identify the AGG sequence; the 32bp following AGG comprise the target sequence. All AGG motifs were identified in *g17*. Ten out of 37 identified potential gRNAs were chosen for gRNA design: four targeting the 5` region and six targeting the 3` region of the gene respectively (see Figure 3-3). They were designed and synthesized as a complementary single stranded DNA fragments (See Appendix 2 for the primers).

A



B

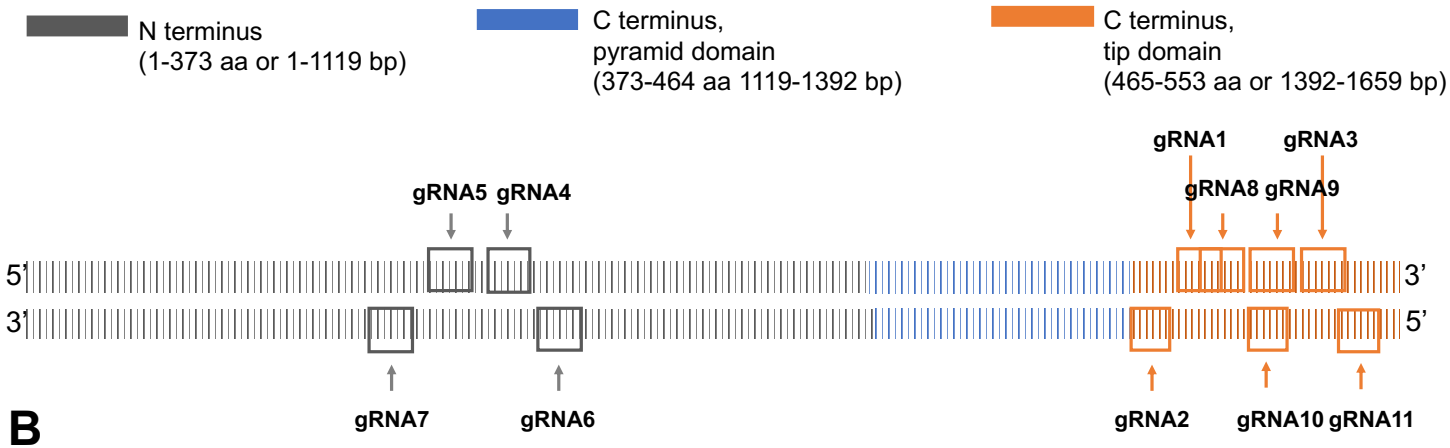
RNA no	Coding/non-coding strand	Sequence *5' to 3'* on *g17*	bp in *g17*
1	coding	TCCTTACGATTAATACAGACTATCGCTTTGCT	128 - 159
2	coding	AAATATTCACGCTAACGGGCGCCTTTACATGA	864-895
3	coding	CGGTAACATCCAGTTAGTAGTAAACGGACAGA	1032-1063
4	non-coding	TTACTCGACGTAACTCGATGGTCGTGTAGCCA	204 - 235
5	non-coding	TACAGTCATTGTTGTTATCTGACCCTCTACCA	1194-1225
6	coding	CGTGGACTCAGGTGTGGTCTGGTAGTGCTGGC	1406-1437
7	coding	TGTGGTCTGGTAGTGCTGGCGGTGGGGTAAGT	1418-1449
8	coding	ATCTCCGCTTCCGCAATATCTGGATTAAGTGT	1466-1497
9	Non-coding	CTATGAAGTAGATTCCATCGGGGCCAGTACGG	1521-1552
10	Non-coding	TACTGAACGACTGTCTGCAATATTCTTGAATC	1601-1632

Figure 3-3. Representation of gRNAs designed for CRISPR type I**. (A)** distribution of gRNAs across *g17*. **(B)** Summary of the position, sequence as well as location on the *g17* for each gRNA.

3.2.3 gRNA design for CRISPR type II

In contrast to CRISPR type I, here the CRISPR components (gRNA, *tracr*RNA and Cas9) required for the system were placed on one vector pCas9-Marafini (see Appendix A), received from Marafini lab already containing gRNA (scr) that will be used as a negative control for the experiments, as it does not target any region in T7. DNA 2.0 (Autumn) was used to design gRNAs against *g17*. The following criteria were selected when using the software: NGG was entered for the desired PAM and species off target verification was based on *E. coli*. Each of the gRNAs generated by the software were appended with Bsa*I* enzyme restriction sites. A total of 11 gRNAs was designed and then synthesised as a complementary single stranded DNA fragments (See Figure 3-4).

A



B

gRNA n	Coding/non-coding strand	Sequence *5' to 3' on g17*	bp in *g17*
gRNA1	coding	AAGTGTGACTGTTTCACAGG	1446-1465
gRNA2	non-coding	AGGCGTGGACTCAGGTGTGG	1403-1422
gRNA3	coding	AGTGTGCCAACAACTCTTGG	1493-1512
gRNA4	coding	TTCCGCTGCGCATCAATCTG	684-703
gRNA5	coding	ACGCTACGAACACAAAGCAG	575-594
gRNA6	non-coding	CAGCATCCGCTAACTCTGCTC	728-747
gRNA7	non-coding	TACAGTTCCGTAATGAGGCT	506-525
gRNA8	coding	GGTAAGTGTGACTGTTTCAC	1443-1462
gRNA9	coding	GATCTCCGCTTCCGCAATAT	1466-1485
gRNA10	non-coding	CTTCCGCAATATCTGGATTA	1473-1492
gRNA11	non-coding	AATACACTCCAACGGTCTCG	1581-1600

Figure 3-4. Representation of gRNAs designed for CRISPR type II. **(A)** distribution of gRNAs across *g17*. **(B)** Summary of the position, sequence as well as location on the *g17* for each gRNA.

3.3 Results

3.3.1 Determination of CRISPR type I efficiency as a method for engineered phage T7 selection

A total of ten vectors containing DNA fragments designed in 3.2.2 was constructed by first annealing the single stranded fragment pairs and then cloned in pAG1 (Section 3.2.1) vector using Golden Gate cloning protocol (Section 2.1.9). The resulting vectors are presented in Table 3-1.

Table 3-1. Summary of CRISPR type I gRNA containing vectors.

Name	Components	Vector	Origin
pAG_1	T7-promoter-5'handle-PAM-BbSI site-(scr)gRNA- BbS*I* site-3'handle-T7 terminator	pSMART-AMP	Designed in this study, IDT synthesised, Appendix 1
pAG_2	gRNA1	pAG1	Generated in this study
pAG_3	gRNA2	pAG1	Generated in this study
pAG_4	gRNA3	pAG1	Generated in this study
pAG_5	gRNA4	pAG1	Generated in this study
pAG_6	gRNA5	pAG1	Generated in this study
pAG_7	gRNA6	pAG1	Generated in this study
pAG_8	gRNA7	pAG1	Generated in this study
pAG_9	gRNA8	pAG1	Generated in this study
pAG_10	gRNA9	pAG1	Generated in this study
pAG_11	gRNA10	pAG1	Generated in this study

The efficiency of each gRNA was determined by EOP (Section2.3.5) with respect to a reference *E. coli* BW25113 strain containing a non-targeting (scrambled) gRNA (Figure 3-5). It was found that gRNA7 was the most effective as it reduced T7 progeny by ~90-fold. The majority of gRNAs (gRNA2-gRNA5 and gRNA8) showed a reduction in progeny by ~10-fold. gRNA1, grNA6, gRNA9 and gRNA10 showed a reduction of less than 10-fold.

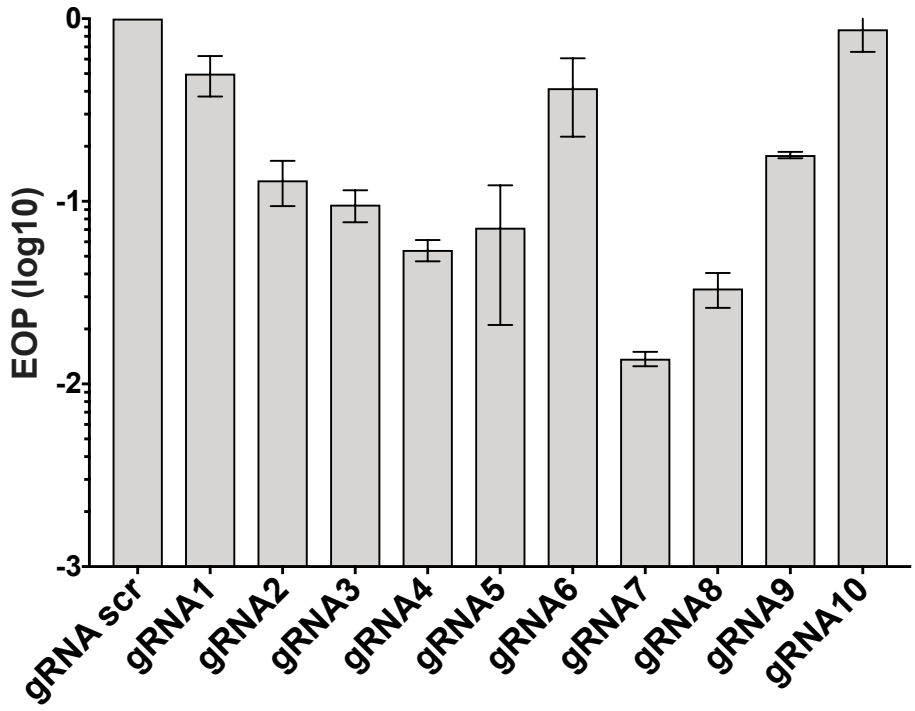


Figure 3-5. Efficiency of plating for T7 against *E. coli* strain BW25113 containing CRISPR type I gRNAs; the T7 efficiency of plating was determined with respect to a reference *E. coli* strain BW25113/pAG1. EOP data are presented as the mean of three independent experiments, n=3.

3.3.2 Determination of CRISPR type II efficiency as a method for engineered phage T7 selection

A total of eleven vectors containing DNA fragments designed in 3.2.3 was constructed by first annealing the single stranded fragment pairs and then cloned in pCas9-Marafini vector using Golden Gate cloning protocol (Section 2.1.9). The resulting vectors are presented in Table 3-2.

Table 3-2. Summary of CRISPR type II gRNA containing vectors made in this study.

Name	Components	Vector	Origin
pAG_12	gRNA1	pCas9 (Marafini) (SF_X)	Generated in this study
pAG_13	gRNA2	pCas9 (Marafini)	Generated in this study
pAG_14	gRNA3	pCas9 (Marafini)	Generated in this study
pAG_15	gRNA4	pCas9 (Marafini)	Generated in this study
pAG_16	gRNA5	pCas9 (Marafini)	Generated in this study
pAG_17	gRNA6	pCas9 (Marafini)	Generated in this study
pAG_18	gRNA7	pCas9 (Marafini)	Generated in this study
pAG_19	gRNA8	pCas9 (Marafini)	Generated in this study
pAG_20	gRNA9	pCas9 (Marafini)	Generated in this study
pAG_21	gRNA10	pCas9 (Marafini)	Generated in this study
pAG_22	gRNA11	pCas9 (Marafini)	Generated in this study

These eleven gRNAs targeting *g17* of T7, were then tested for their ability to cut the T7 genome and thus reduce T7 progeny (Figure 3-6). The efficiency of each gRNA was determined by EOP (Section 2.3.5), with respect to a reference *E. coli* BW25113 containing a non-targeting gRNA (scr). It was found that gRNA2 was the most effective resulting in a 1000-fold reduction compared to the control. The majority of remaining gRNAs (gRNA5, gRNA7-gRNA13) showed a reduction in progeny numbers by less than 10-fold. gRNA3, gRNA4 and gRNA6 showed a reduction more than 10-fold but less than 100-fold. Determination that CRISPR type II can reduce the T7 progeny suggested that it has a potential for being used as a selection method for engineered phage selection. CRISPR type II reduced the numbers of wild type phage by 1000-fold which surpasses the result seen for CRISPR type I (Section 3.3.1).

The next step was to determine the T7 reduction i.e. marker-based method selection potential for the two host factors *cmk* and *trxA*.

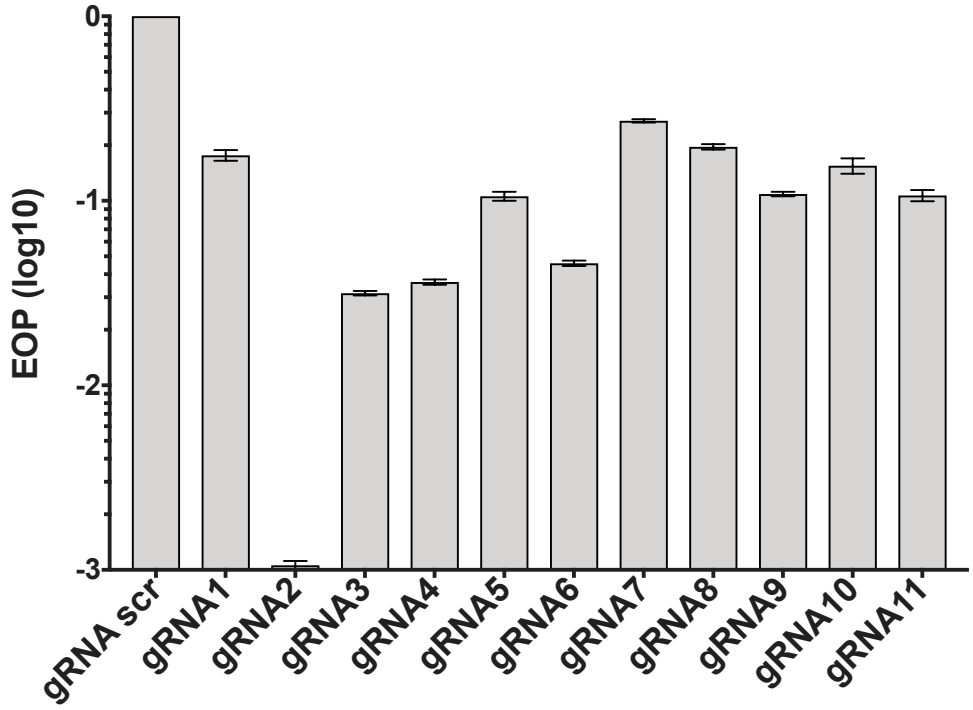


Figure 3-6. Efficiency of plating for T7 against *E. coli* strain BW25113 containing CRISPR type II gRNAs; the T7 efficiency of plating was determined with respect to a reference *E. coli* strain BW25113/pAG1. EOP data were $\log_{10}$ transformed and are presented as the mean of three independent experiments, n=3.

3.3.3 Determination of marker-based (*cmk* and *trxA*) selection efficiency as a method for engineering phage T7 selection

To determine the extent to which both of the host factors reduce T7 progeny, plaque assays were carried out on *E coli* strains that lacked *cmk* and *trxA* (Section 2.3.5). EOP on both strains were compared to the control strain *E. coli* BW25113. In the case of *E. coli* Δ*cmk* cells, the number of progeny phage was reduced by approximately 10^5 fold (Figure 3-7), whereas with *E. coli* Δ*trxA* cells no progeny was detected (detection limit < 1) and is presented as zero. Thus, utilising *trxA* is the most efficient marker-based method for the selection of engineered phages.

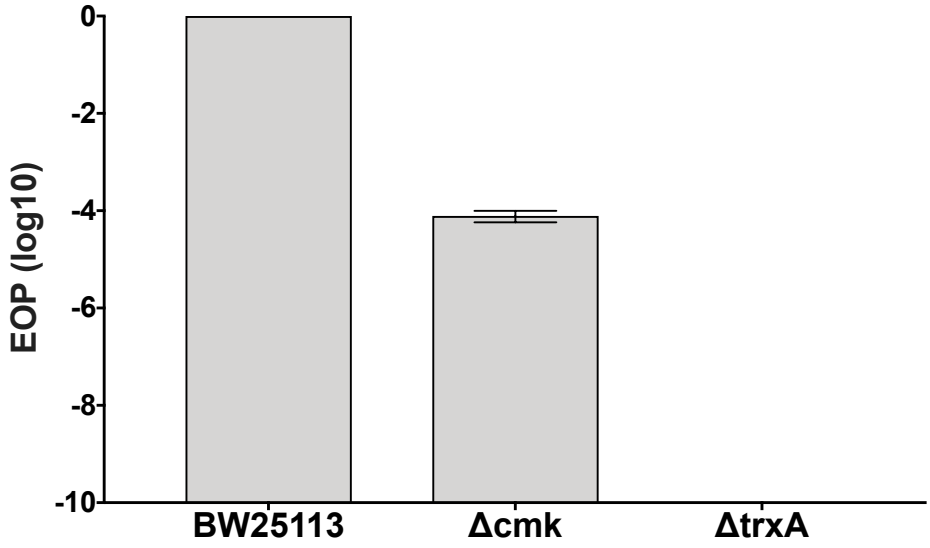


Figure 3-7. Efficiency of plating for T7 on *E. coli* BW25113*, E. coli Δcmk* and *E. coli ΔtrxA* strains was determined with respect to a reference *E. coli* strain BW25113. EOP data were $\log_{10}$ transformed and are presented as the mean of three independent experiments, n=3. An asterisk indicates that the EOP was below the detection limit which is < 1.

In addition, a T7 one-step growth assay (Section 2.3.7) was performed on *E. coli* BW25113, *E. coli Δcmk* and *E. coli ΔtrxA* strains to further show that there is no phage replication taking place when *trxA* is knocked out and very little when *cmk* is knockout (Figure 3-8).

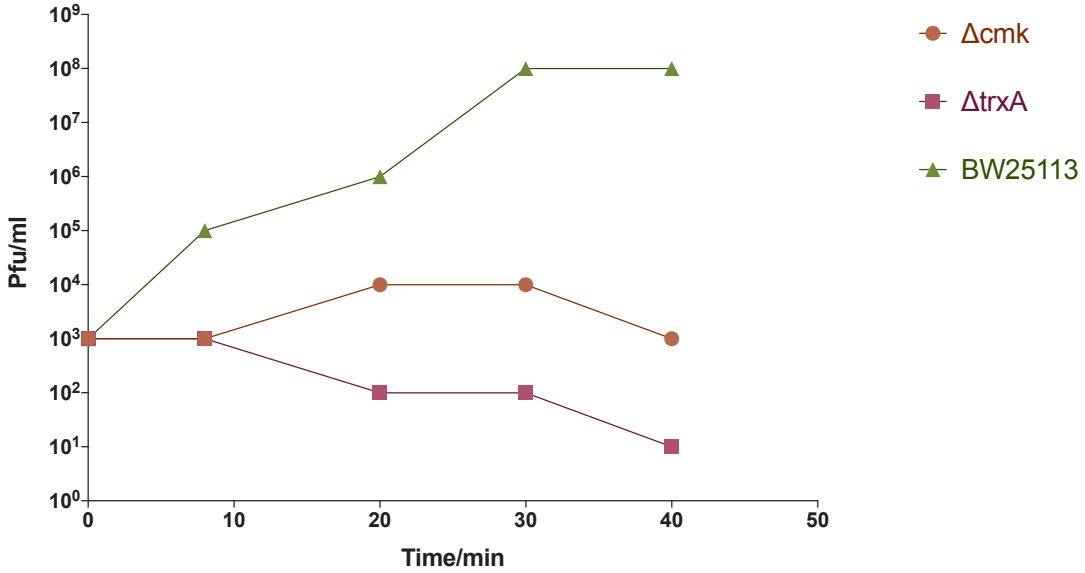


Figure 3-8. Burst size assay for T7 phage on *E. coli* BW25113, *E. coli Δcmk* and *E. coli ΔtrxA* strains was carried out over 40 minutes. The assay was carried out at MOI 0.1.

3.3.4 Comparison of marker-less and marker-based selection efficiencies

The EOP of CRISPR type II BW25113/gRNA2 (the most efficient gRNA) and *E. coli* Δ*trxA* cells was directly compared (Figure 3-9). The marker-less method reduced T7 progeny numbers by ~1000-fold whereas the marker-based method gave undetectable amount of progeny (it is presented as zero in the Figure 3-9). Therefore, the marker-based method is the most efficient method for selection of phages and hence it will be used as the main engineering method.

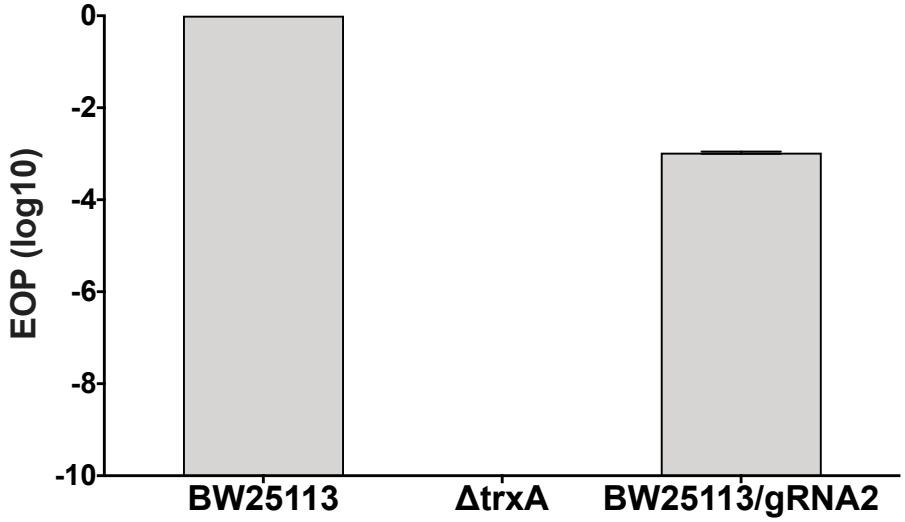


Figure 3-9. Efficiency of plating for T7 on *E. coli BW25113*, *E. coli ΔtrxA* and *E. coli ΔtrxA*/pAG3 strains and was determined with respect to a reference *E. coli* strain BW25113; hence the value of 1 (0) for BW25113 strain; EOP data were $\log_{10}$ transformed and are presented as the mean of three independent experiments. An asterisk indicates that the EOP was below the detection limit (detection limit <1).

3.4 Discussion

3.4.1 CRISPR type I

The maximum reduction in wild-type T7, and hence the potential T7 mutant selection efficiency, was found to be 100-fold greater than that of a control for the CRISPR type I, when targeting *g17*. Interestingly, in the previous work by Kiro et al. (2014) showed a 1 x 10^4 fold decrease in wild-type T7 numbers when targeting *g1.7* using the same CRISPR type I system suggesting a 100 times greater efficiency when compared to the results for *g17* (Figure 3.3.1-1)[327]. In contrast to *g17* targeted in this chapter, the 1.7 gene, encoding a nucleotide kinase, is not essential for phage growth under standard laboratory conditions[337]. In principle, when a gene is non-essential it is perceived to be less conserved and more prone to adaptation when compared to an essential gene[258,338]. Therefore, a non-essential gene such as *g1.7* being more tolerant to mutations would be better suited to avoid CRISPR targeting by mutating the region of interest and would result in lower CRISPR targeting efficiency when compared to *g17*. However, the opposite is the case, when the result generated in this chapter for *g17* is compared with the data to that of g1.7, therefore potential explanations will be discussed next.
A reason for the steep difference between the CRISPR type I efficiency numbers in this study and the work of Kiro et al. (2014) is likely caused by the recombinant phage selection protocol. In their study an additional dual pre-induction step of the selection strain was carried out before the addition of recombinant phage.

CRISPR pre-induction prior to phage entry into the cell is likely to be very important factor to allow maximum efficiency when targeting a lytic phage. Phage infection is very rapid (T7 is 17 min) and hence CRISPR component transcription and translation as well as complex formation are time consuming steps; therefore, having all of these components ready upon the lytic phage arrival would allow more time for the target sequence screening and hence more genomic cuts made[226]. If the system is solely dependent on the phage to provide the T7 RNA polymerase, which is one of the early genes (*g1*) of the T7 genome, for the initiation of CRISPR, the entire process is then delayed and then, most likely, results in lowered efficiency. To induce the CRISPR, Kiro et al. (2014) used the following. Firstly, all three vectors which make up the CRISPR type I are under control of T7*lac* promoter which means that addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) would free the promoter region by binding to the lacI (that normally acts as an inhibitor for T7lac promoter regulated regions). Secondly,

they used *E. coli* BL21-AI strain which is an arabinose-inducible strain because it has the tightly regulated arabinose-inducible araBAD promoter upstream of the T7 RNA polymerase gene. Therefore, when both araBAD promoter pre-induction is combined with a T7 promoter-based vector (which all of the three CRISPR components were) it produces tight regulation for the T7 promoter. Thus, an active CRISPR was present at the time of T7 infection, whereas in the system here the system would not become active until after phage infection. Since it takes approximately six minutes after infection at 30°C (and potentially even less time at 37°C), for T7 encoded RNA polymerase to start transcription of the middle genes it is beneficial to have the CRISPR complexes expressed and assembled prior to phage entry[230].

3.4.2 CRISPR type II

The maximum reduction in wild-type T7, and hence the potential T7 mutant selection efficiency, was found to be 1000-fold greater than that of a control for the CRISPR type II, when targeting *g17*. In a previous study a type II CRISPR of *Streptococcus thermophilus* DGCC7710 (not identical to the one used in this study, which is from *Streptococcos pyogenes*) has been used as a tool to select mutations in a specific region of a virulent phage[329]. The same study also replaced orf33 in the phage 2972 genome with LlaDCHIA (methyltransferase gene of the type II restriction/modification (R/M) system). In contrast to this study they found that 1 x 10^6 fold efficiency was achieved when type II CRISPR was used to aid the homologous recombination[329]. As well as in the case of Kiro et al. (2014), a non-essential gene, *g1.7*, was targeted. This higher difference in CRISPR efficiency (1000-fold) could be due to the following reasons.

Firstly, CRISPR type II used in this study originated from *S. pyogenes*. Jiang et al. 2013 generated this plasmid by amplifying Cas9, trcRNA and repeat regions, as well as accompanying native promoters for each, directly from *S. pyogenes* and inserting into pZE21-MCS1 vector[333]. The *S. pyogenes* promoters when in *E. coli* drive the expression constitutively, however, the strength of the promoter and hence the level of the expression has not been empirically compared between *E. coli* and *S. pyogenes*. The lack of an inducible promoter with the type II CRISPR system, means that the expression cannot be tuned in any way. The reduced effective strength of *S. pyogenes* promoters in *E. coli* could be one of the potential reasons for reduction in CRISPR type II efficiency found in this study. In addition, in the previously mentioned study by Martel & Moineau (2014), they performed the CRISPR targeting in *S.*

thermophilus, the strain that harbours its own CRISPR system[329]. Therefore, only a vector containing a gRNA targeting the phage needed to be introduced in order for the system to work.

Secondly, Martel & Moineau (2014) engineered bacteriophage 2972 rather than T7. The two phages in question have very different latent periods: 13 and 34-40 mins for phage T7 and 2972 respectively[339,340]. It is noticeable that for phage with a longer latent period, type II CRIPSR was a more efficient method for the selection of mutants; this could be due to the fact that longer latent period time would allow more time for CRISPR assembly as well as target recognition. In addition both phages have a different GC content, 40.15% and 48.4% for phage 2972 and T7 respectively, which can potentially affect CRISPR efficiency[341,224].

In addition, it seems that the CRISPR of *S. thermophilus* offers a few advantages over the CRISPR *S. pyogenes*, such as higher targeting efficiency and sequence specificity. This is due to the fact that *S. pyogenes* CRISPR system uses NGG as its PAM region, whereas *S. themophilus* is able to detect longer PAM allowing greater target specificity and reducing of targets[342].

3.4.3 Potential Explanation for CRISPR inefficiency

Interestingly, the most efficient gRNA for each CRISPR type (gRNA7 for type I and gRNA2 for type II) in this study was in the region of 1402-1448 nt of *g17* (they had a 5 bp overlap), suggesting a potential location/site preference for CRISPR targeting. For CRISPR type II there has been more analysis carried out to understand the factors contributing towards CRISPR efficiency. It has been suggested that coding vs non-coding strand targeting could potentially make a difference for efficiency[343]. In a study carried out by Qi et al. 2013, the non-coding strand targeting was shown to be more effective when a gene region is targeted; however when a promoter region is targeted, both coding and non-coding are effective[343]. It was shown that the most effective CRISPR type I gRNA7 targets the coding strand, whereas for CRISPR type II the most efficient gRNA2 type targets the non-coding strand. The next efficient gRNA for CRISPR type I and type II was gRNA8 and gRNA3 respectively; both of which are coding strands. The next in line for efficiency for CRISPR type I was gRNA8 targeting coding strand and for CRISPR type II gRNA6 targeting non-coding strand. Even though, non-coding targeting appears to be more efficient, the efficiency of the gRNAs

cannot be directly compared given the different CRISPR types as well as limited number of gRNAs investigated in this study.

A previous large scale study that evaluated the cleavage activities of 218 gRNAs targeting mouse Neuro2A cell line proposed that the relationship between GC% content and cleavage efficiency is likely to be a non-linear one, with sgRNAs having GC% of 40–60 are optimal[344]. The GC content of gRNA2 and gRNA7 were 59% and 65% respectively, which is within or just above the previously identified optimal range of GC% of 40–60 Figure 3-10. Interestingly, when the GC% is compared between all of the gRNAs it is seems that the most effective gRNAs tend to have higher GC content.

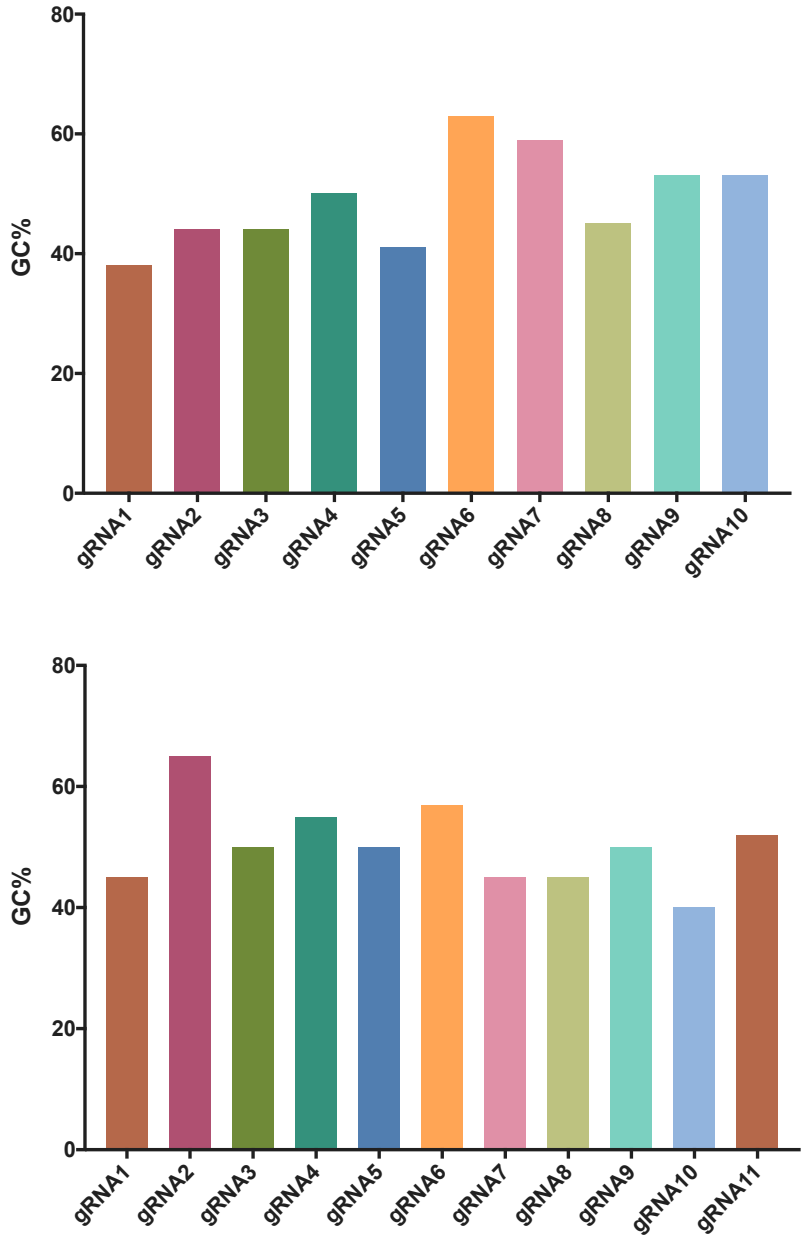


Figure 3-10. Summary of GC content for gRNAs used in this study. **(A)** gRNAs for CRISPR Type I. **(B)** gRNAs for CRISPR Type II.

With the use of CRISPR to engineer phages, it is important to remember that phages have evolved anti-CRISPR systems. Both Martel, B. & Moineau, S 2014 and Kiro et al. 2014 showed that phage mutants can bypass the interference activity (CEM, CRISPR-Escape Mutants) by point mutations in the protospacer (PS) sequence or adjacent motif[329,327,339]. The frequency of escape mutants varies from 1 x 10^{-5} - 1 x 10^{-7}. Differences in the frequency of escape mutants could potentially affect CRISPR efficiency and hence explain the differences observed; the frequencies of escape mutants were not identified in this study.

Potential approaches to improve CRISPR efficiency for bacteriophage engineering could be prolonging phage lysis as well as reducing the number of phage genome escape mutants. In the case of the phage lysis, as seen in section 3.4.1, it may occur faster than CRISPR protein assembly and provide with insufficient time for CRISPR proteins to perform an edit. In addition to CRISPR pre-induction in bacterial target host, one could employ physical factors i.e. temperature to extend phage lysis. For example, lowering the temperature would slow down phage infection giving more time for CRISPR to alter phage genome. In the case of escape mutant formation (as discussed in above) one could design multiple gRNAs targeting nearby region lowering the probability of successful escape mutant generation.

3.4.4 Marker-based selection of T7 mutants

A previous comparison between the efficiency of phage particle formation in the absence of *trxA* or *cmk* was carried out by Qimron in 2008[328]. In the study 100 pfu of T7 were plated on *ΔtrxA* and *Δcmk* cells and their EOP was worked out in respect to *E. coli* K-12 BW25113. The EOP value for *trxA*-based selection was <0.01, indicating a 100-fold reduction, which means that no viable progeny was observed, as only one hundred phages were plated. The EOP value for *cmk* selection was 0.05, indicating the recovery of 5 out of 100 phage particles and showing a 20-fold reduction. The data here for the selection using *trxA,* is in agreement with the values presented by Qimron et al. 2006[328]. However, the EOP value for *cmk* presented in this work showed a reduction of progeny by ~1 x10^5 fold, far exceeding previous reports[328]. This is could be due to the fact that in the previous study by Qimron, U. et al. 2006 only 100 T7 phages were used to enumerate the reduction by the lack of the host gene i.e. *cmk*; whereas in this study 10^{10} phages were used providing with a more realistic representation of the T7 progeny reduction[328].

Chapter 4 Host range expansion via tail fibre modification

4.1 Introduction

4.1.1 Expansion of phage host range via tail fibre swapping

Although phages with similar tail fibres are usually similar genetically, there have been instances reported in which genetically unrelated phages share similar tail fibres[345]. For example, the *E. coli*-infecting P1 and P2 phages are unrelated genetically, despite sharing similar tail fibres and having been isolated from the same bacterial strain[346]. This disconnect between genetic similarity and tail fibre phenotype can be explained via tail fibre transfer, in which phages attain new host ranges by acquiring a portion of a tail fibre gene from another phage. In one study it was suggested that the similarities in the tail fibre genes of distantly related phages such as P2, P1, Mu, X, and T4 provide evidence that phages have undergone tail fibre exchange repeatedly[115]. Natural tail fibre domain swapping is most likely an evolutionary mechanism adopted by phages to enhance their survival[115]. Recently, advances in phage genome engineering allow for host range expansion via genetically engineered tail fibre manipulation; specifically, other studies have successfully expanded the host ranges of both T7 and T3 phages[321]. In this chapter, in order to test if tail fibre swapping is possible for distantly related phages, the tail fibres of *E. coli*-infecting T7 were swapped with those of Bordetella-infecting BPP-1.

4.1.2 Structure of the phage T7 tail fibre protein (Gp17)

The function of the T7 tail fibre protein, Gp17, is to attach the phage to the lipopolysaccharide (LPS) of *E. coli*[347]. Even though it is clear that LPS is the main T7 receptor, it has been shown that porins such as OmpA and OmpF affect phage adsorption and infection, indicating that they take part in the initial virus-host interaction[348]. Each of the six T7 tail fibres consists of three copies of the Gp17 protein[244]. These homotrimers, and especially the C-terminus of Gp17, are responsible for the initial recognition and binding to the target host[347,249]. The crystal structure of the C-terminus of Gp17 has been determined, where the terminal domain is divided into two parts: a globular 'tip' domain (Ala465 - Glu553) and pyramid domain (Gly371 – Glu533) Figure 4-1[249,349].

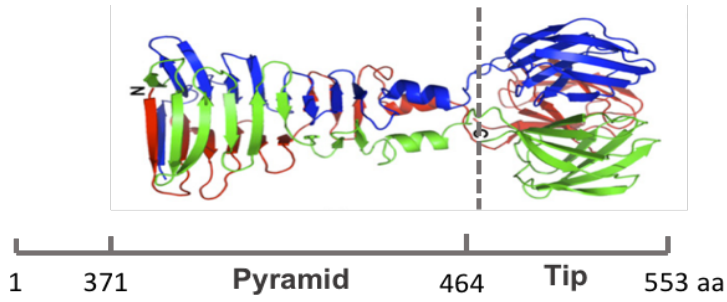


Figure 4-1. Representation of the structure of Gp17 (only C-terminus shown, as the structure for the N-terminus has not been determined). The C-terminus domain consists of pyramid (372-464aa) and tip (464-553aa) domains. The image was adapted from a previous study by Garcia-Doval, Carmel Van Raaij, Mark J. 2012.

4.1.3 Structure of the phage BPP-1 tail fibre protein (Mtd)

The tail fibre of BPP-1, Mtd, binds to pertactin on the *Bordetella pertussis* cell surface and uses it for initial attachment[130,350,351]. Pertactin is an outer membrane autotransporter produced during the Bvg+ (virulent) phase by all members of the *Bordetella bronchiseptica* cluster, which includes *B. pertussis, B. parapertussis*, and *B. bronchiseptica*[318,351,352]. The Mtd protein is divided into three domains: a N-terminal β-prism (1-48 aa), an sandwich domain formed by a β-sandwich (56-170 aa), and a C-terminal lectin C domain (170-381 aa) Figure 4-2[350]. In addition, the β-prism is connected to the sandwich domain by a short 3_{10} – helix (49–54 aa) Figure 4.1.3-1[350]. Though the functions of the β-prism and sandwich domains are not clear, they may reinforce overall trimeric assembly and therefore have an indirect role in stabilizing the backbone of the variable region. The function of the lectin C domain is to directly bind to pertactin[350].

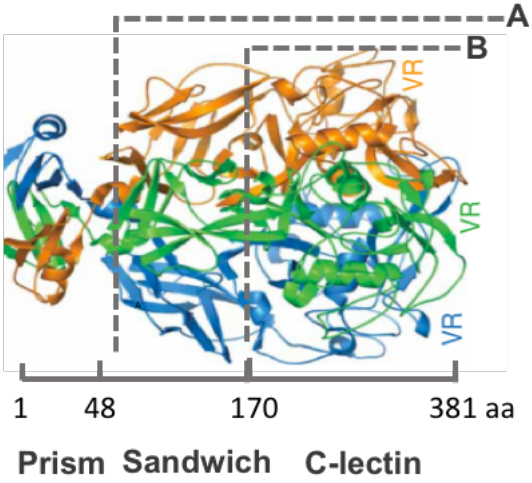


Figure 4-2. Representation of Mtd structure. Mtd has three domains: prism (1-48aa), sandwich (54-170 aa) and C-type lectin (170-382 aa). In addition, a short helix (49–54 aa) connects the β-prism to the sandwich domain. The image was adapted from a previous study by S. McMahan et al. 2005.

4.1.4 Antibody-like versatility of the Mtd C-lectin domain

The C-terminal domain of Mtd is composed of a C-lectin type domain that is able to tolerate massive sequence variation and is frequently compared to that of the immunoglobulin fold[350]. The C-lectin fold of Mtd is related to that of the macrophage mannose receptor as well as intimin[353,354,355]. The key structural features of the C-lectin fold in Mtd are two-stranded antiparallel β-sheets formed by the domain's N and C termini (β1β5), which are then connected by two α-helices and a three-stranded antiparallel β-sheet (β2β3b4). The β2β3β4 sheet in Mtd contains an additional three-residue strand, β4, and a short 3_{10}-helix. The C-lectin domain allows the bacteriophage to bind to a vast diversity of *B. pertussis* surfaces, as theoretically the C-terminus of Mtd can be composed of 10^{12} different polypeptide sequences. This high sequence variation is due to a 134 bp long tropism-determining region called variable region 1, located at the 3' end of the Mtd locus.

4.1.5 BPP-1 and its variants BMP-1 and BIP-1

Phage BPP-1 infects *B. pertussis* and *B. bronchiseptica* via pertactin, an outer membrane autotransporter protein that is only expressed in Bvg^+ phase *Bordetella* spp[129]. BPP-1 can give rise to two classes of tropic variants at a frequency of approximately 10^{-6}. One of the classes, BMP-1 (Bvg minus-tropic phage), has acquired a tropism for Bvg^- phase bacteria while the second class, designated BIP-1 (Bvg indiscriminate phage), can infect both Bvg^+ and Bvg^- phase *B. bronchiseptica* with equal efficiency[129].

4.1.6 Aims

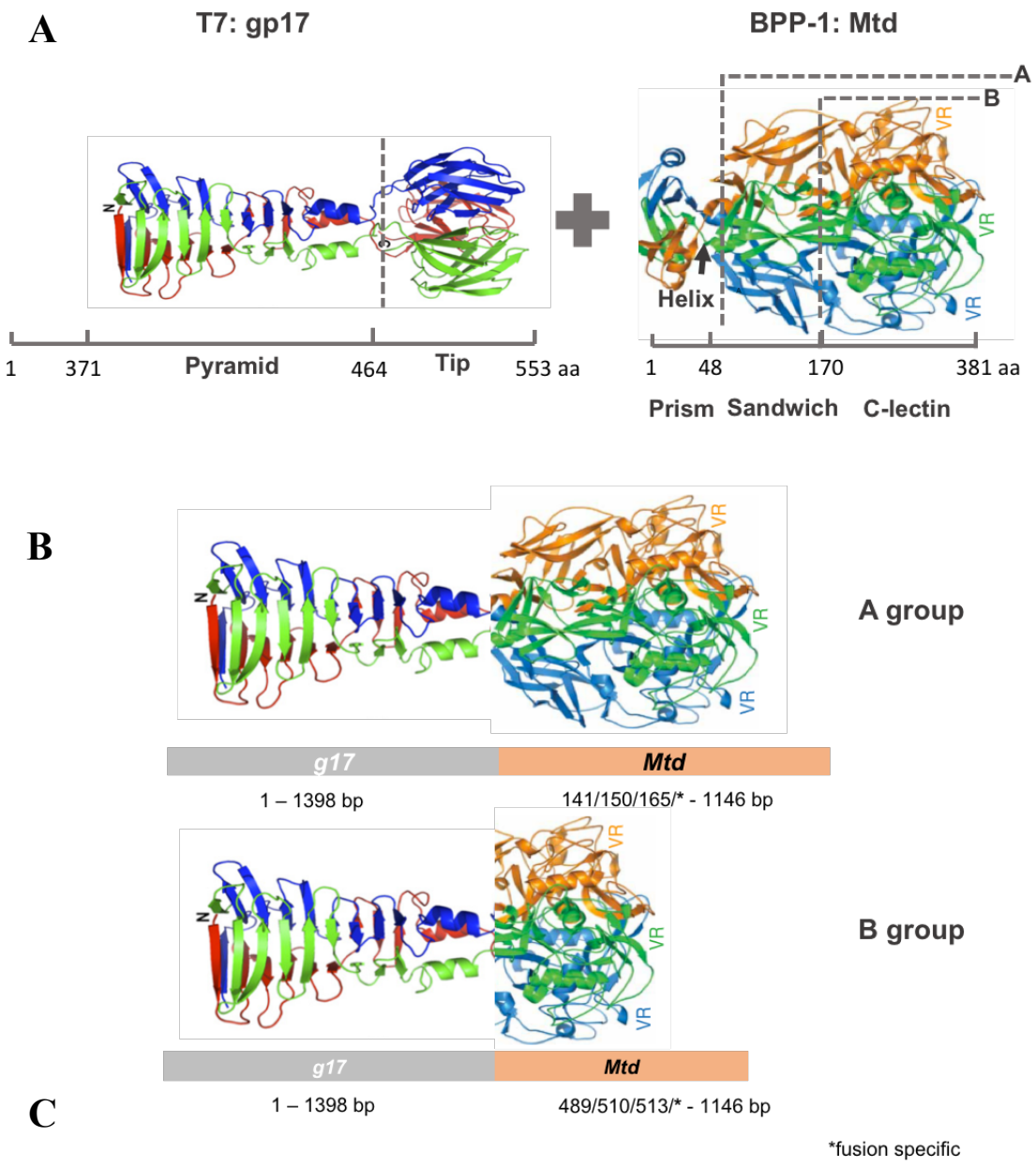
The main aim of the chapter is to determine if the host range of T7 can be altered when tail fibre domains from a genetically distant host are appended to it. This aim is divided into the following sub-aims:

1. To design Gp17 and Mtd tail fibre fusions
2. To establish a method for *in trans* complementation with tail fibre fusions, where tail fibre defective phage T7 infect cells expressing mutant tail fibres, allowing tail fibre complementation
3. To perform *in vitro* testing of tail fibre fusions for tail fibre folding and trimerization analysis
4. To execute *in vivo* testing where the genes for the fusion tail fibres were integrated into the genome of the phage T7, generating T7 chimeras
5. To test the host range of chimeric phage against *B. pertussis* and *B. bronchiseptica.*

4.2 Results

4.2.1 Design of T7 and BPP-1 tail fibre fusions

When creating tail fibre fusions, the tail fibre encoding gene (Gp17) from phage T7 was used as the main scaffold (Figure 4-3). To that scaffold, the Mtd C-terminus domains of the BPP-1 phage tail fibre were fused. The C-terminus domain was chosen because it is the key site for binding to the host surface. To assess which potential fusions of Gp17 and Mtd were likely to result in a functional tail fibre, the precise locations for fusions had to be chosen on both Gp17 and Mtd proteins.



CH:T7/BPP1	Fusion Group	Gp17/*g17* aa/bp	Mtd/*Mtd* aa/bp
1 (T7-CH1)	B	437 - 466/1-1398	171 – 382/513 – 1146
2 (T7-CH2)	A	437 - 466/1-1398	55 – 382/165 – 1146
3 (T7-CH3)	B	437 - 466/1-1398	163 – 382/489 – 1146
4 (T7-CH4)	B	437 - 466/1-1398	170 – 382/510 – 1146
5 (T7-CH5)	A	437 - 466/1-1398	47 – 382/141 – 1146
6 (T7-CH6)	A	437 - 466/1-1398	50 – 382/150 – 1146

Figure 4-3. T7 and BPP-1 phage tail fibre fusion designs: **(A)** Representation of Gp17 and MtD domains. Only the C-terminus of Gp17 shown, as the structure for the N-terminus has not been determined. The Gp17 C-terminus domain consists of pyramid (372-464 aa) and tip (464-553 aa) domains. MtD has three domains: prism (1-48 aa), sandwich (54-170aa) and C-type lectin (170-382 aa) domains. Prism and sandwich domains are connected via flexible helix (49 – 54 aa). **(B)** g17 and Mtd fusion representations at DNA level. Fusion group A consists of g17 (1-1338 bp) fused to Mtd (141/150/165-1146 bp). Fusion group B consists of g17 (1-1338 bp) fused to Mtd (489/519/513-1146 bp). **(C)** Summary of all fusion designs for both A and B groups.

To ensure that the chosen Gp17 scaffold would provide stability and facilitate trimerization for the fusion protein the Gp17 C-terminal pyramid domain was chosen. It forms stable trimers (the calculated dissociation energy of the pyramid domain trimer is 90 kcal/mol whereas the calculated dissociation energy of the tip domain trimer is 17 kcal/mol), suggesting that it enables Gp17 C-terminal trimerization[356]. Furthermore, it was also found that the short loop (Val464-Lys466) between the pyramid and tip domains is flexible[356]. Therefore, it was decided to fuse the Mtd domain directly after the 464-466 region loop as this would most likely help to retain flexibility between Gp17 and Mtd domains.

To determine which C-terminal parts of Mtd to fuse to Gp17 the following was considered. Since the Mtd C-lectin domain (C-t): 170 – 381 directly binds to the pertactin of *B. pertussis*, it is therefore essential for host recognition and must be included in the fusion with Gp17[125]. In an effort to ensure C-lectin domain stability and trimerization, the sandwich domain was also included in the fusion, as its extensive interaction with the C-lectin domain indicates it may be linked to C-lectin domain stability. To ensure a fusion tail fibre length variation for both *in vivo* and *in vitro* work, two types of fusions were designed. For the first set of fusions, named type A fusions, the pyramid domain of T7 tail fibre (Gp17) was fused with both the sandwich domain and the C-lectin domain of Mtd starting at amino acids 47, 50 and 55 (Figure 4-3). For type B fusions a Gp17 pyramid domain was fused only with the C-lectin domain before amino acids 163, 170 and 171 (Figure 4-3).

The fusions described above were adapted for three lines of work: 1) *in trans* complementation with tail fibre fusion experiments, where tail fibre defective T7 infects cells expressing mutant tail fibres, allowing tail fibre complementation. 2) *in vitro* experiments, for tail fibre folding and trimerization analysis. 3) *in vivo* experiments where the genes for the fusion tail fibres were integrated into the genome of the phage T7, generating T7 chimeras.

4.2.2 Experimental setup for *in trans* complementation of tail fibre mutants

The production of engineered phage genomes, while possible, is still a time-consuming task[357]. The ability to assess if tail fibre fusions will produce functional tail fibres, without having to produce recombinant phage, is a valuable tool in the development of synthetic phage-based antimicrobials. A system to test this requires a defective tail fibre (non-infective) phage and a strain with a vector that expresses a non-defective tail fibre. Upon defective phage infection in the strain, a non-defective vector tail fibre would be integrated into the virion, ensuring further phage infectivity. Such system was referred to as *in trans* tail fibre complementation. For this system phage Q12, that is a T7 derivative, was used. Phage Q12 has an amber stop codon (Amber mutation) in the 12th amino acid of *g17* of T7 and was only infective in amber suppressor *E. coli* strains such as R11 (Section 2.1.1)[358–360]. As an initial step to verify the *in trans* tail fibre complementation system, it was confirmed using EOP (efficiency of plating) that the Q12 cannot produce viable phage on a non-amber suppressor strain (MG1655) and can on an amber suppressor strain (R11) with a 1 x 10^4 increase in viable phage (Figure 4-4) (Section 2.3.5).

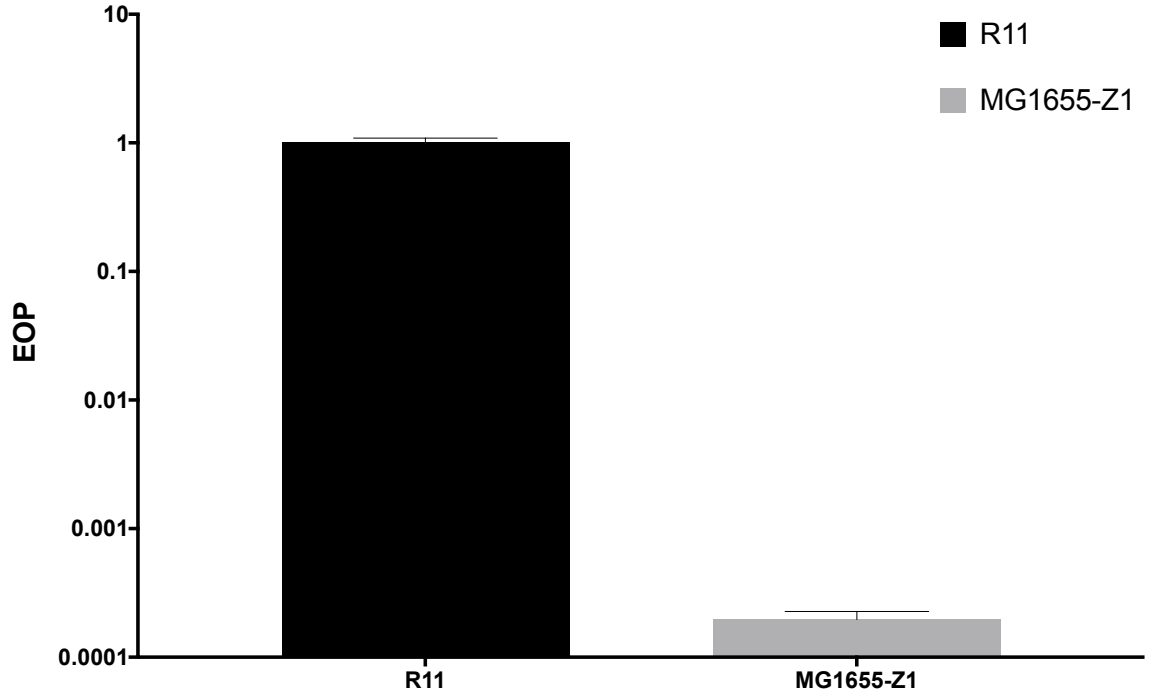
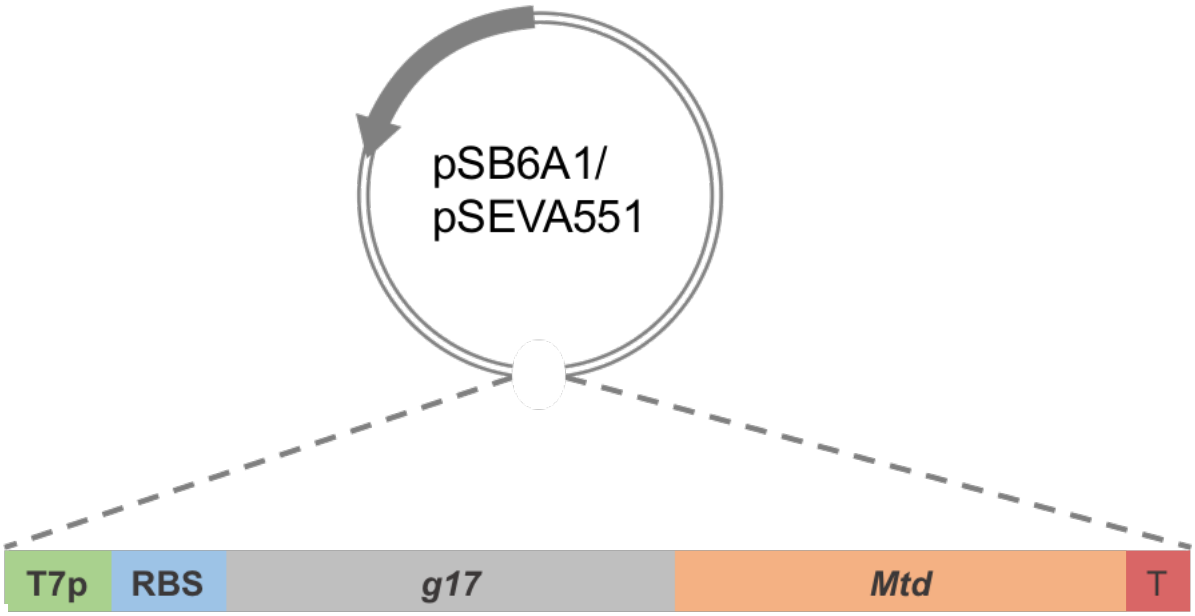


Figure 4-4. EOP of Q12 against R11 (amber suppressor) MG1655-Z1 strains. R11 as well as MG1655-Z1 were infected with Q12 phage; phage enumerated were compared against the R11 strain. Error bars show standard deviation of three biological repeats.

4.2.2.1 Construction of tail fibre mutants for in trans complementation

To test if viable Q12 could be propagated in MG1655 cells when wild type *g17* was provided *in trans*, a vector expressing wild type *g17* was constructed. If successful *in trans* tail fibre complementation platform would allow the function of tail fibre fusions to be tested. T7 *g17* was amplified with overhangs and inserted into pSB6A1 in front of a T7 promoter and B0034 RBS using Gibson cloning (Section 2.1.7) creating the vector pAG_25. To generate the *in trans* complementation vector pAG_26, with a fusion tail fibre coding gene, T7 *g17* was removed from pAG_25 and replaced with one of the fusions designed as described in section 4.2.1 (Figure 4-5). As a control for the pAG_26 vector, pAG_27 was created. In this instance, T7 *g17* was removed from pSEVA-551 vector (Appendix 1), containing codon optimised *g17* that has been codon optimized for T7, and replaced with one of the fusions designed as described in section 4.2.1 (Figure 4-5).

A



B

Name	Gp17/aa// *g17/bp*	Fusion Group	Mtd/aa// *Mtd*/bp	Vector
pAG_25	1-553// 1-1659	N/A	N/A	pSB6A1
pAG_26	1-466*// 1-1398	B	171 – 382// 513 – 1146	pSB6A1
pAG_27	1-466**// 1-1398	B	171 – 382// 513 – 1146	pSEVA551

Figure 4-5. Summary of the constructs made for the in trans experiments. **(A)** Partial vector sequence representing the order of the main elements of the in trans fusion plasmids. Main components are the T7 promoter, RBS (B0034), the g17 sequence up to the end of the pyramid domain (until 1398 bp, or 466 aa), and the Mtd sequence of the C-lectin domain (513 – 1146 bp, 171 – 382 aa). **(B)** Breakdown of the three plasmids made. pAG_25 contains the full sequence of WT g17 instead of Mtd fusion, and will be used as a control for the in trans experiments. T – terminator; *-wild type Gp17; **- codon optimized Gp17.

4.2.2.2 Phage Q12 complementation with wild type Gp17 and Gp17: Mtd fusion tail fibres

To test if Q12 would recover its infectivity when provided the wild type *g17 in trans*, phage enumeration assays were performed using *E. coli* Z1/pAG_25. The efficiency of plating (EOP) phage Q12 on MG1655/pAG_25 was compared to that of plating on R11 and the negative control MG1655 Z1 (Figure 4-6). The EOP using MG1655/ pAG_25 was equal to that of R11 and was 1 x 10^4 higher than the negative control, confirming *in trans* complementation of the defective tail fibre. However, complementation with both variants of fusion tail fibre using pAG_26 and pAG_27 resulted in approximately the same EOP value as the negative control.

This was expected given that these chimeric fusion tail fibers had their tip domain swapped for Mtd domain making the tail fiber no longer suitable for *E. coli* infection.

As *in trans* complementation required a wild-type copy of *g17* to be expressed, there was the potential for *g17* to recombine with Q12 (unpublished data by Mark van Raaij's laboratory). Therefore, a codon optimised version of *g17* (pSEVA551) was also tested that is 78.9% similar to *g17*, and therefore less likely to recombine. The codon optimised Gp17 complemented the mutation in Q12 as seen by the same EOP when using R11 cells (Figure 4-6). It was therefore decided to use the codon optimised *g17* in further chimeric tail fibre generation.

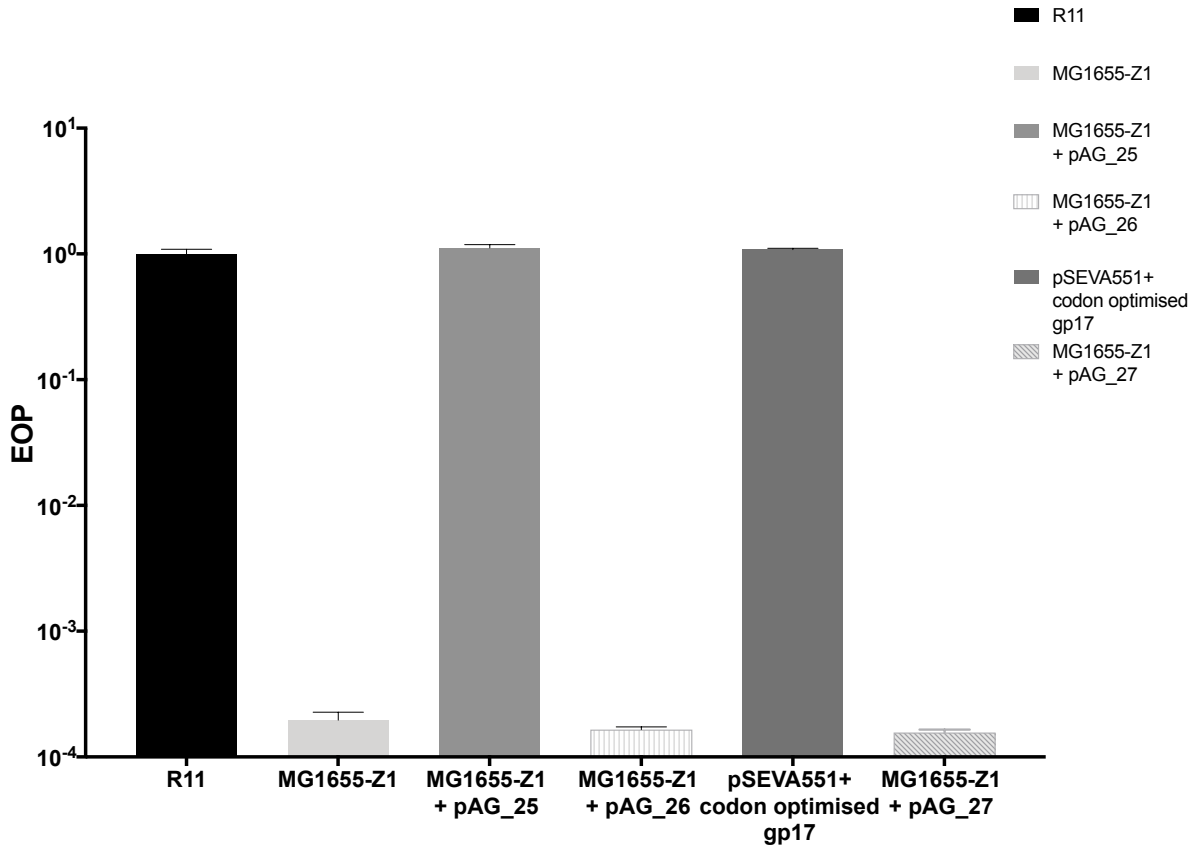


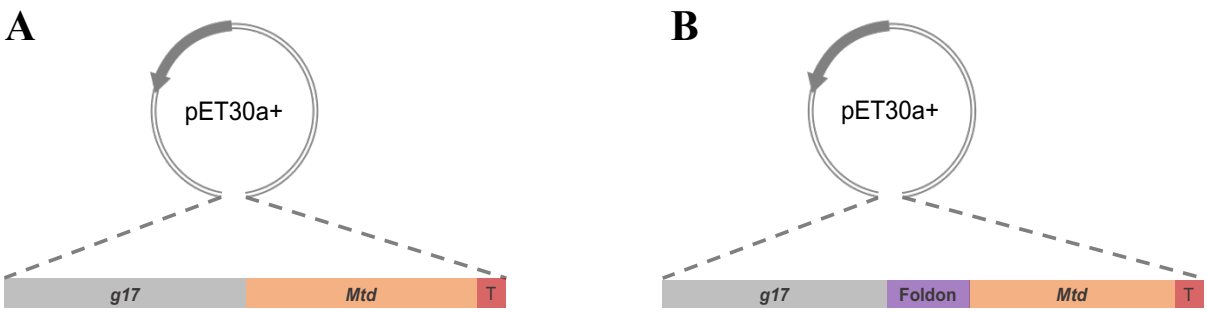
Figure 4-6. *In trans* complementation with wild type Gp17 as well as fusion tail fibre. The T7 tail fibre gene g17 was expressed in trans in *E. coli* MG1655-Z1/pAG-25 and infected with phage Q12. In addition, a T7 and Mtd tail fibre Mtd fusion gene was expressed in trans in *E. coli* MG1655-Z1/ pAG_26 and infected with Q12. As an additional control, codon optimised tail fibre gene was expressed in *E. coli* MG1655-Z1/pAG-27 and infected with Q12. MG1655-Z1 and the amber suppressor strain R11 were used as negative and positive controls respectively. Efficiency of plating relative to R11 was calculated. Error bars show standard deviation of three biological repeats.

Given that the fusion tail fibre did not sustain the complementation, most likely due to incorrect folding, it was decided that the functionality of each fusion tail fibre needs to be assessed *in vitro* prior to placement in the T7 genome. Therefore, the next step was to place all of the fusions into an expression vector followed by *in vitro* trimerization and solubility assays.

4.2.3 Construction of fusion tail fibre chimeras in an expression vector for *in vitro* tail fibre analysis

It is known that Gp17 proteins can form trimers, either when both N-terminal and C-terminal domains are present or when only C-terminal residues (pyramid and tip domains) are present[349,249]. Therefore, the trimerization and solubility of tail fibres are good indicators of correct folding and infectivity[249,361]. Thus, it was decided to test tail fibre fusions designed in this study for potential trimerization and solubility. To achieve this, tail fibre fusions designed in Section 4.3 were cloned in pET30a (+) under control of a T7/lac promoter to allow tuneable expression by IPTG, using Gibson cloning (Section 2.1.7). The partial vector representation for the tail fibre fusion expression is showed in Figure 4-7.

Previous work has shown that a T4 foldon domain can aid the folding of adenovirus tail fibres when used in a fusion protein[362,363]. To increase the likelihood of tail fibres trimerising and producing functional fibres, a second strategy of incorporating a foldon domain into fusion proteins was carried out. All of group A and B tail fibre fusions designed in Section 4.2.1 were redesigned to include a foldon domain that sits between the Gp17 pyramid domain and the Mtd fusion of interest (Figure 4-7). Constructs were made using Gibson cloning (Section 2.1.7) and confirmation of the plasmids by Sanger sequencing (see Appendix 4).



C

Name	Gp17/aa	Mtd/aa	Foldon	Vector
pAG_28	371 – 467	171 - 382	yes	pET30a+
pAG_29	371 – 467	171 - 382	no	pET30a+
pAG_30	371 – 467	55 - 382	yes	pET30a+
pAG_31	371 – 467	55 - 382	no	pET30a+
pAG_32	371 – 467	163 - 382	yes	pET30a+
pAG_33	371 – 467	163 - 382	no	pET30a+
pAG_34	371 – 467	170 - 382	yes	pET30a+
pAG_35	371 – 467	170 - 382	no	pET30a+
pAG_36	371 – 467	47 - 382	yes	pET30a+
pAG_37	371 – 467	47 - 382	no	pET30a+
pAG_38	371 – 467	50 - 382	yes	pET30a+
pAG_39	371 – 467	50 - 382	no	pET30a+
pAG_40	371 – 553*	N/A	yes	pET30a+
pAG_41	371 – 553**	N/A	yes	pET30a+

Figure 4-7. Summary of constructs made for the *in vitro* tail fibre assays, for both A and B group Gp17 and Mtd fusions. **(A)** Vector insertion sequence for the no foldon *in vitro* fusions. It consists of g17 (aa, bp) and Mtd (sequence dependent on if it is a type A or type B fusion). **(B)** Vector insertion sequence for the foldon containing *in vitro* fusions. It consists of g17 (aa, bp), foldon, and Mtd (sequence dependent on if it is a type A or type B fusion). **(C)** Summary of all fusions made in this study for the *in vitro* study.

4.2.3.1 In vitro trimerization and solubility assays of tail fibre fusions

To assess the solubility and the ability of tail fibre fusions to trimerize correctly, *in vitro* expression assays (Section 2.5.3) were carried out. The constructs made in section 4.2.3 were sent to Mark van Raaij's research group where SDS PAGE as well as native gel were carried out for *in vitro* tail fibre analysis (Figure 4-8, Figure 4-9).

To verify chimeric protein solubility SDS PAGE assay was carried out on all samples (only some of the samples are shown in Figure 4-8). Both wild-type *g17* and a codon optimised *g17* fused to a foldon domain produced soluble protein as Gp17 monomers were seen in both supernatant as well as the pellet. To verify possibility of trimerization, native PAGE gels (only two of the samples are shown in Figure 4-9) were performed at 0, 1, 2, 3 and 16 hours post IPTG induction (Table 4-1). Trimerisation occurred for all of the tail fibre fusions with the exception of one (pAG_29), however, the trimers were only observed in the insoluble fraction (Table 4-1). No monomers were observed for any of the fusions in the soluble fraction, suggesting the fusions were insoluble Table 4.2-1.

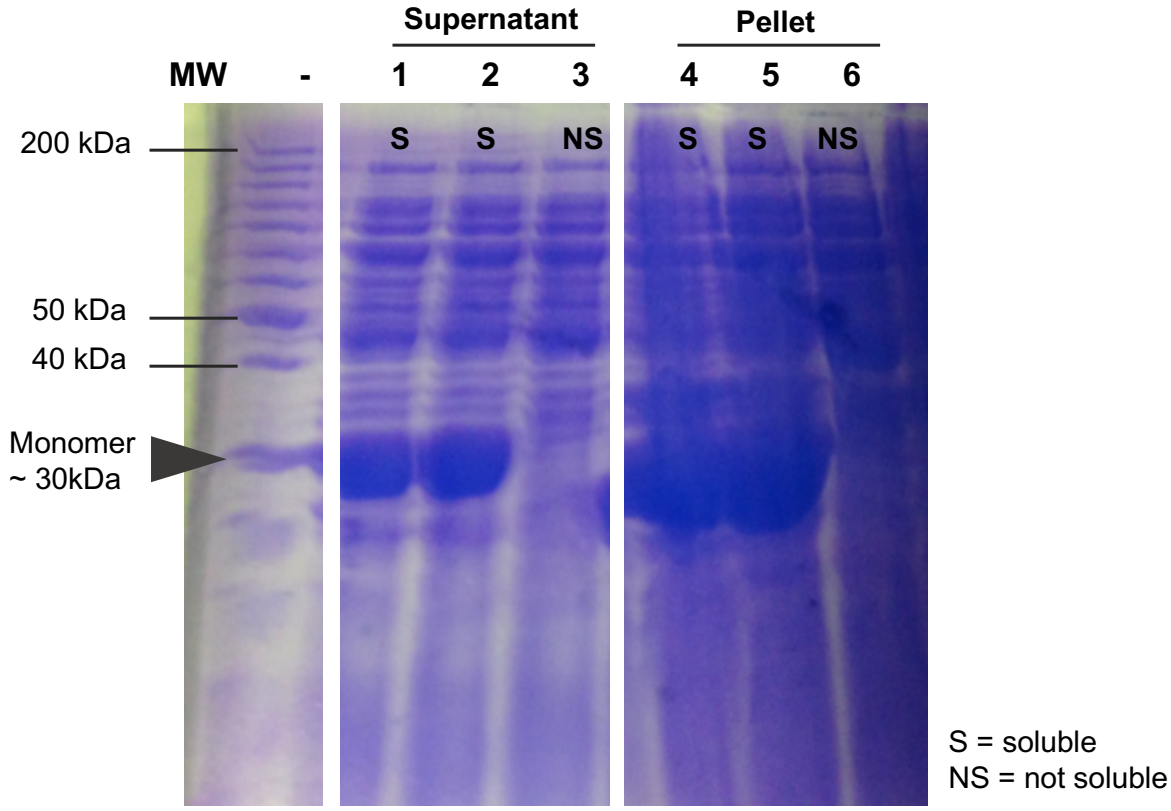


Figure 4-8. SDS-PAGE polyacrylamide gel from attempt one. Lanes and their contents are labelled 1-6. Lane 1 corresponds to the non-boiled supernatant fraction of IPTG induced pAG_40. Lane 2 corresponds to the non-boiled supernatant fraction of IPTG induced pAG_41. Lane 3 corresponds to the non-boiled supernatant fraction of IPTG induced pAG_29. Lane 4 corresponds to the non-boiled pallet fraction of IPTG induced pAG_40. Lane 5 corresponds to the non-boiled pallet fraction of IPTG induced pAG_41. Lane 6 corresponds to the non-boiled pallet fraction of IPTG induced pAG_29. The figure has been generated by Mark van Raaij's research group.

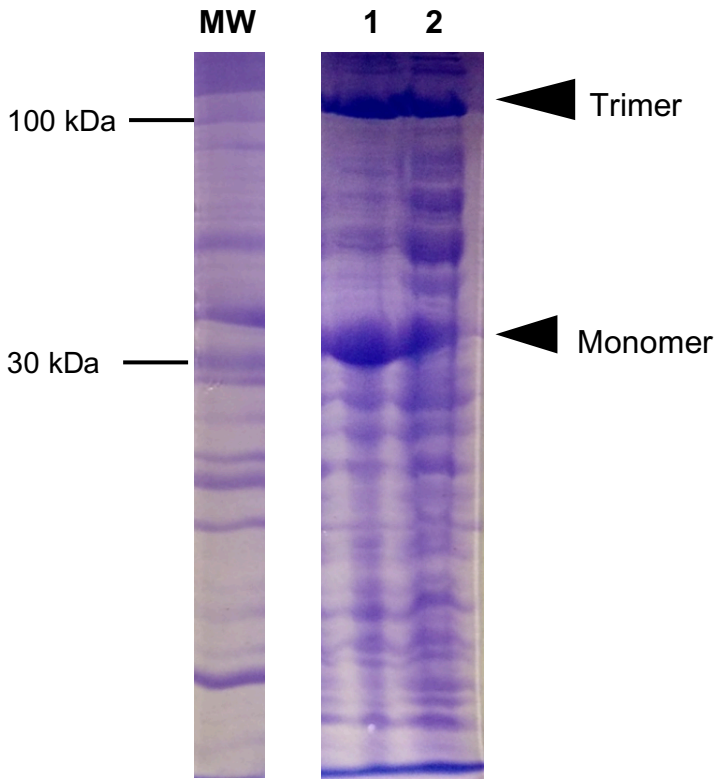


Figure 4-9. Representation of a sample of Native PAGE gel**.** Lane 1 corresponds to the non-boiled supernatant fraction of IPTG induced pAG_40. Lane 2 corresponds to the non-boiled supernatant fraction of IPTG induced pAG_41. In both cases the sample was collected 16 hours post induction.

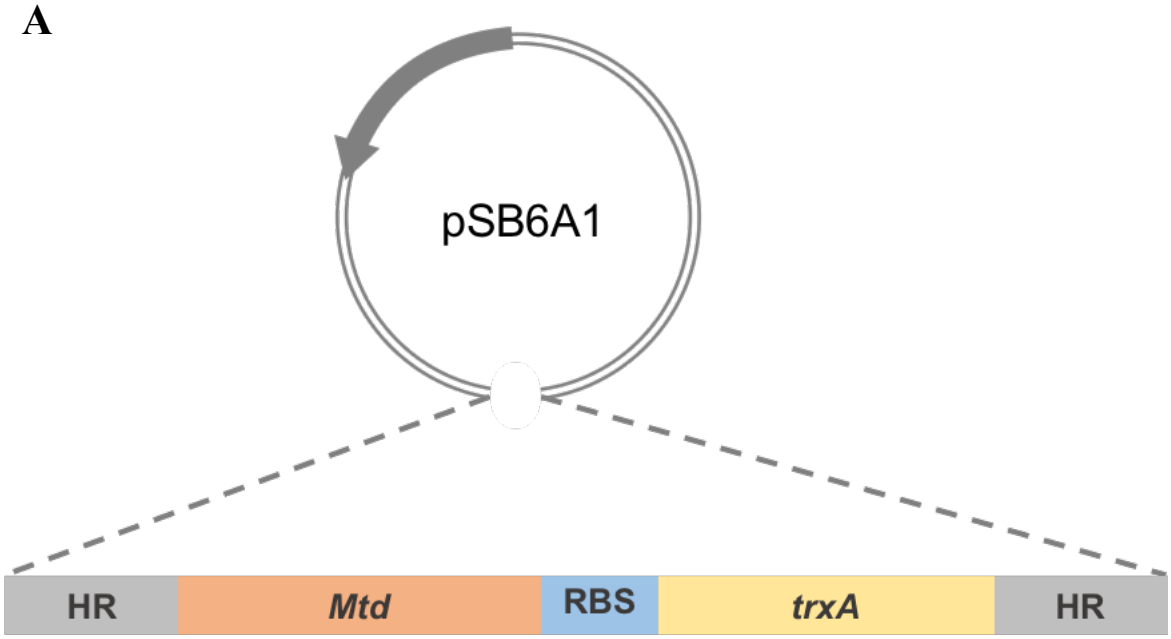
Table 4-1. Summary of tail fibre fusion *in vitro* expression assays.

Name	Conditions				Outcome
	Pellet		Supernatant		
	Boiled	Non-boiled	Boiled	Non-boiled	
pAG_40	Monomer	Trimer	Monomer	Trimer	Protein expression detected at time 3 hours post IPTG induction; trimerization not detected since there was no trimer corresponding band neither in the pallet or the supernatant; a monomer band found in the supernatant (as well as the pallet) indicating that this fusion is soluble.
pAG_41	Monomer	Trimer	Monomer	Trimer	Protein expression detected at time 3 hours post IPTG induction; trimerization not detected since there was no trimer corresponding band neither in the pallet or the supernatant; a monomer band found in the supernatant (as well as the pallet) indicating that this fusion is soluble.
pAG_28	Monomer	Trimer	No	No	Protein expression detected at time 1, 2, 3, and 16 hours post IPTG induction; the sample can trimerize since in the pellet two bands are seen at the expected MW for the monomer (when boiled) and trimer (when not boiled), but it's insoluble since these bands are not seen in the supernatant.
pAG_29	No	No	No	No	Protein expression was not detected at time 3 hours post IPTG induction; the sample does not trimerize and is insoluble since there were no trimer or monomer corresponding bands neither in the pallet or the supernatant.
pAG_30	Monomer	Trimer	No	No	Protein expression detected at time 1, 2, 3, and 16 hours post IPTG induction; the sample can trimerize since in the pellet two bands are seen at the expected MW for the monomer (when boiled) and trimer (when not boiled), but it's insoluble since these bands are not seen in the supernatant.
pAG_31	Monomer	Trimer	No	No	Protein expression detected at time 1, 2, 3, and 16 hours post IPTG induction; the sample can trimerize since in the pellet two bands are seen at the expected MW for the monomer (when boiled) and trimer (when not boiled), but it's insoluble since these bands are not seen in the supernatant.
pAG_32	Monomer	Trimer	No	No	Protein expression detected at time 1, 2, 3, and 16 hours post IPTG induction; the sample can trimerize since in the pellet two bands are seen at the expected MW for the monomer (when boiled) and trimer (when not boiled), but it's insoluble since these bands are not seen in the supernatant.
pAG_33	Monomer	Trimer	No	No	Protein expression detected at time 1, 2, 3, and 16 hours post IPTG induction; the sample can trimerize since in the pellet two bands are seen at the expected MW for the monomer (when boiled) and trimer (when not boiled), but it's insoluble since these bands are not seen in the supernatant.
pAG_36	Monomer	Trimer	No	No	Protein expression detected at time 1, 2, 3, and 16 hours post IPTG induction; the sample can trimerize since in the pellet two bands are seen at the expected MW for the monomer (when boiled) and trimer (when not boiled), but it's insoluble since these bands are not seen in the supernatant.
pAG_37	Monomer	Trimer	No	No	Protein expression detected at time 1, 2, 3, and 16 hours post IPTG induction; the sample can trimerize since in the pellet two bands are seen at the expected MW for the monomer (when boiled) and trimer (when not boiled), but it's insoluble since these bands are not seen in the supernatant.

4.2.4 Construction of tail fibre chimeras for homologous recombination

The results of the *in vitro* folding assays suggested that the tail fibre fusions were probably not able to form functional proteins as they were insoluble and therefore were unlikely to fold correctly. To investigate if tail fibres fusions may work *in vivo*, the same tail fibre gene fusions were engineered into the genome of phage T7. Genome engineering was carried out using the previously developed homologous recombination and marker-based selection method on *E. coli ΔtrxA* (Sections 2.3.11 and 3.3.3). In order to perform homologous recombination, first plasmids had to be generated that contain recombination arms, gene fusions and the *trxA* selective marker Figure 4-10. All of the constructs were cloned in pSB6A1. Each of the constructs generated had homologous recombination arm (HR1) (sequence of 99 bp at the end of *g17* genome region coding for Gp17 pyramid domain), *Mtd* sequence (depending on the fusion of interest) full *trxA* and HR2, which consisted of 99 bp after the *g17* Figure 4-10. A total of seven different variants, three for each no-foldon A and B fusion group, as well as a complete *g17* knockout (Figure 4-10) were generated using Gibson assembly cloning (Section 2.1.7).

A



B

Name	CH:T7/BPP1	Fusion Group	HR 1: *g17*/bp	*Mtd*/bp	RBS, *trxA*	HR 2	Vector
pAG_42	1	B	1309 – 1398	513 – 1146	full sequence	90 bp after *g17*	pSB6A1
pAG_43	2	A	1309 – 1398	165 – 1146	full sequence	90 bp after *g17*	pSB6A1
pAG_44	3	B	1309 – 1398	489 – 1146	full sequence	90 bp after *g17*	pSB6A1
pAG_45	4	B	1309 – 1398	510 – 1146	full sequence	90 bp after *g17*	pSB6A1
pAG_46	5	A	1309 – 1398	141 – 1146	full sequence	90 bp after *g17*	pSB6A1
pAG_47	6	A	1309 – 1398	150 – 1146	full sequence	90 bp after *g17*	pSB6A1
pAG_48	N/A	N/A	90 bp upstream *g17*	N/A	full sequence	90 bp after *g17*	pAJ57

Figure 4-10. Summary of constructs made for homologous recombination, for both A and B group Gp17 and Mtd fusions. **(A)** Partial vector sequence representing the components of the homologous recombination vector (Mtd fusion). From left to right, the components are: homology arm 1 (HR 1) which consists of g17 (1309-1398), the Mtd sequence of interest (depends on the fusion), a stop codon (not shown here), RBS (B0034), full trxA gene, stop codon, terminator, and then homology arm 2 (HR 2) that is 89 bp after the g17 stop codon. **(B)** Constructs made for homologous recombination.

4.2.4.1 Marker-based phage mutant selection with and without in trans complementation

Prior to making all T7 tail fibre chimeras it was decided to test the applicability of marker-based method selection on just one tail fibre chimera. Two sub-methods were devised to verify this. Both methods had the same initial step of phage T7 infection of *E. coli* BW25113, which contained the homologous recombination plasmid pAG_42 (Figure 4-11). The resulting progeny consisted of both T7 as well chimeric phages (Figure 4-11). The first method consisted of simply plating the above phage mixture on *E. coli* Δ*trxA* for selection (Figure 4-11). The data collected up to this point suggested that the tail fibre fusions would be non-functional. If this is the case then there would be no plaques generated on the *E. coli* Δ*trxA* because the recombinant phage with the fusion tail fibres would not infect the cells due to the non-functional tail fibres. Therefore, a second selection method was devised. The phage mixture after the homologous recombination was plated on *E. coli* Δ*trxA/*pSEVA551 (Figure 4-11). In the case of tail fibres not being functional, having *g17 in trans* complementation would enable the chimeric phage infectivity. At a genomic level these chimeric phages would retain the tail fibre fusions via the selection of *trxA.* However, the tail fibre phenotype would be a mixture of wild type and tail fibre fusions. This should retain the phage infectivity and generation of plaques on the *E. coli* Δ*trxA/*pSEVA551 even if the tail fibre fusions were non-functional.

Homologous recombination for T7 was carried out using vector pAG_42 and the resultant progeny was plated on *E. coli* Δ*trxA/*pSEVA551 or *E. coli* Δ*trxA* (Figure 4-11). Both selections resulted in plaque formation (Figure 4-12). Plaques for the *E. coli* Δ*trxA* only selection generated half the size plaques when compared to *E. coli* Δ*trxA/*pSEVA551 (Figure 4-12). Selection on *E. coli* Δ*trxA* produced progeny at a titre of 1 x 10^2 PFU/ml, which was lower than the 1 x 10^8 PFU/ml when selecting on *E. coli* Δ*trxA/*pSEVA551 (Figure 4-12).

Plaques from both methods were tested by PCR (Section 2.3.12) to determine if recombinant phage were present using primers AG063 and AG064 that bind outside the *g17* region. A difference of ~1 kb should be observed between wild type *g17* and a chimeric *g17.* Selection on *E. coli* Δ*trxA* only resulted in two bands: one for the recombinant phage and one for the wild type phage, which suggested that there was wild type phage contamination in all of the plaques (Figure 4-12). This was not the case for the plaques selected using *E. coli* Δ*trxA/*pSEVA551. The plaques from the first method were plaque purified ten times, however, the two bands persisted at each purification step, indicating wild type contamination. The in

trans method was chosen to generate further chimeric phages (Table 4-2), where each of the chimeric phages were purified using *E. coli ΔtrxA* complemented with *g17* expressing vector (pSEVA551).

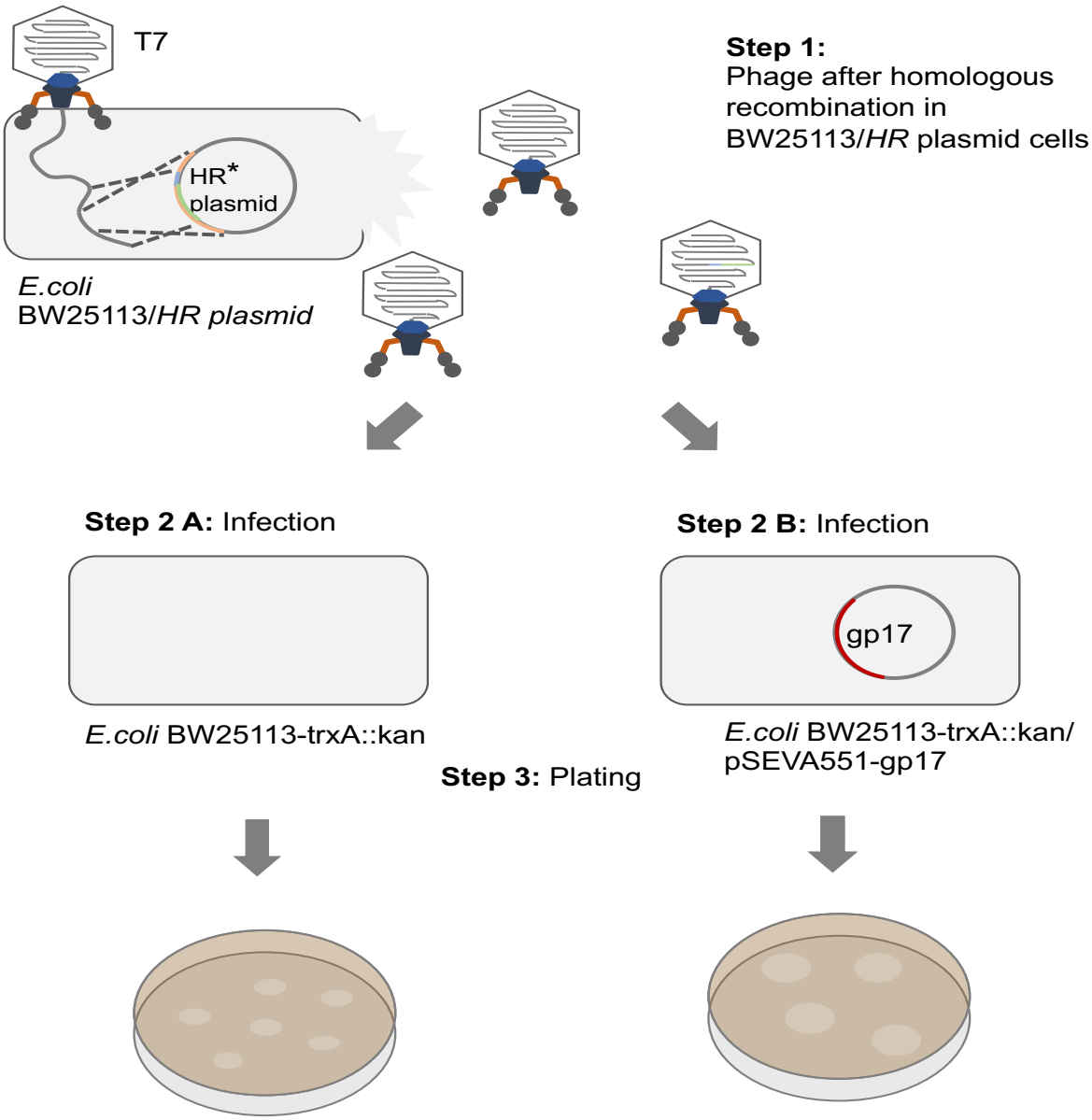


Figure 4-11. Selection for recombinant phage. **Step 1** - infection of phage T7 in *E. coli* BW25113 cells containing a tail fiber gene homologous recombination (HR) plasmid. The infection yields a mixture of phage, the wild type T7 as well as homologously recombined T7 (chimeric T7). **Step 2** - selection of chimeric phage by infection of the phage mixture in the Step 1 into **(A)** *E. coli ΔtrxA* cells or **(B)** *E. coli ΔtrxA*/pSEVA551 cells. **Step 3** – plating the cells and phage from Step 2 on an agar plate.

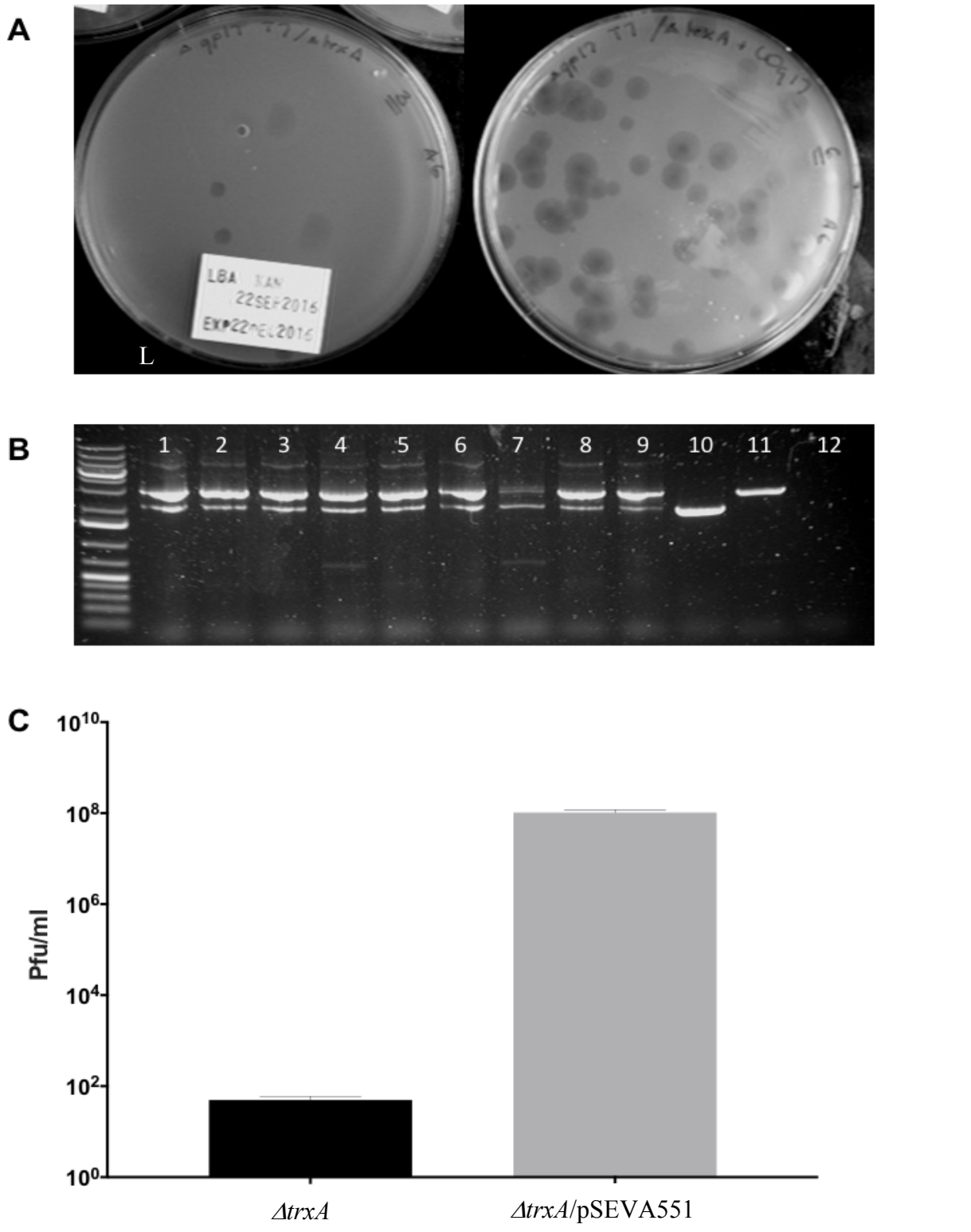


Figure 4-12. Phage enumeration after T7 homologous recombination in *E. coli* BW25113/pAG_42 and its later genomic verification. **(A)** Enumeration of phage plated after homologous recombination of T7 on *E. coli ΔtrxA* and *E. coli ΔtrxA/* pSEVA551. **(B)** Chimeric phage verification using PCR for *g17* region (AG063 and AG064 primers), WT T7 product is ~2 kb and chimeric of ~3 kb. 'L' lane represents GeneRuler 1 kb Plus DNA, samples/lanes 1-9 *E. coli ΔtrxA* plaques. Lane/sample 10 product of T7 *g17* (+ve control). Lane/sample 11 is a plaque generated on *E. coli ΔtrxA/*pSEVA551. **(C)** Phage enumeration of the phage mixture after homologous recombination plated against *E. coli* Δ*trxA* and *E. coli* Δ*trxA/*pSEVA551.

Table 4-2. Tail fibre chimeric phages were generated using trxA *in trans* method, where each of the chimeric phages bellow were purified using *E. coli ΔtrxA* complemented with *g17* expressing vector (pSEVA551).

Name	Fusion Group	HR 1/aa	Mtd*/aa	Rbs/trxA	HR 2**/aa
phAG_1 (T7-CH1)	B	437-466	171 - 382	full sequence	end of Gp17
phAG_2 (T7-CH2)	A	437-466	55 - 382	full sequence	end of Gp17
phAG_3 (T7-CH3)	B	437-466	163 - 382	full sequence	end of Gp17
phAG_4 (T7-CH4)	B	437-466	170 - 382	full sequence	end of Gp17
phAG_5 (T7-CH5)	A	437-466	47 - 382	full sequence	end of Gp17
phAG_6 (T7-CH6)	A	437-466	50 - 382	full sequence	end of Gp17
phAG_7 (T7-CH7)	N/A	upstream *g17*	N/A	full sequence	end of Gp17

4.2.4.2 Pertactin expression in E. coli for chimeric phage testing

In order to verify if the chimeric phages were able to infect, they were tested on *B. pertussis* and *B. bronseptica*. However, the strains were not initially available, so prior to this testing an alternative characterisation system was sought. It was decided to express the receptor protein pertactin in *E. coli*. If successful this would then be followed by assays with chimeric phage challenging (infecting) pertactin expressing *E. coli*, to determine if pertactin presence allows binding by the designed tail fibre fusions. Pertactin containing expression vector, pSL*prn*32 (Section 2.1.1), was kindly provided by Angus Bonetti and was used in this study.

To confirm that pertactin is expressed and localised in the outer membrane cellular layer, pertactin containing cultures were separated by ultracentrifugation post induction followed by Western blotting (Sections 2.5.1, 2.5.2 and 2.5.4). In Figure 4-13, lanes 1-3 show the OM fraction of pSL*prn*32 expressing cells. No pertactin equivalent band (P 69) was seen prior to induction, while by 1 hour a weak band was visible. Lower weight bands were also seen, which may be due to pertactin degradation by-products.

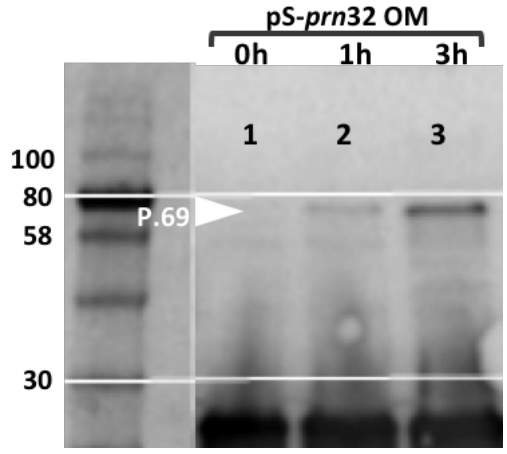


Figure 4-13. SDS-PAGE and Western Blotting of the pS-prn32 in BL21. 'OM' is the outer membrane, fractionated by ultracentrifugation and selective dissolution in sarkosyl. The location of P.69 was determined by overlaying the protein molecular weight marker on the membrane image which showed the corresponding band to be 58-80kDa.

4.2.4.3 Infection of pertactin expressing E. coli with chimeric phage

Once the pertactin expression in *E. coli* was confirmed it was possible to test whether the chimeric phage was able to infect using pertactin as a receptor. To determine this, growth assays were conducted in which different types of host cells were subjected to either phage, IPTG induction, or both.

The *in trans* method without complementation was used to generate chimeric phages. As seen in 4.2.4.1 this method results in chimeric and T7 phages. In order to test whether phage killing was due to pertactin-mediated entry or T7 entry via LPS, *E. coli ΔwaaC* cells were used to express pertactin. This strain has a kanamycin gene inserted into the coding region of its *E. coli ΔwaaC* gene, the product of which facilitates the biosynthesis of LPS. Consequently, it is unable to produce wild-type LPS and is resistant to T7 infection[328]. The use of *E. coli ΔwaaC* was to prevent any wild-type T7 infection and any infection had to be as a result of T7 chimeric phages. The pL*prn*34 construct was transformed into *E. coli ΔwaaC* as well as *E. coli* BW25113. Growth assays were conducted with and without phage (Section 2.2.3). No lysis was observed in any of the LPS-deficient *E. coli ΔwaaC* cells, whether or not they bore the pertactin-expression construct or were induced (Figure 4-14). Conversely, lysis was observed in the *E. coli* BW25113 control. However, this was likely due to wild type T7 contamination as the phage used in this experiment was *E. coli* Δ*trxA* purified without *in trans* complementation. Therefore, there was likely no lysis generated due to chimeric phage.

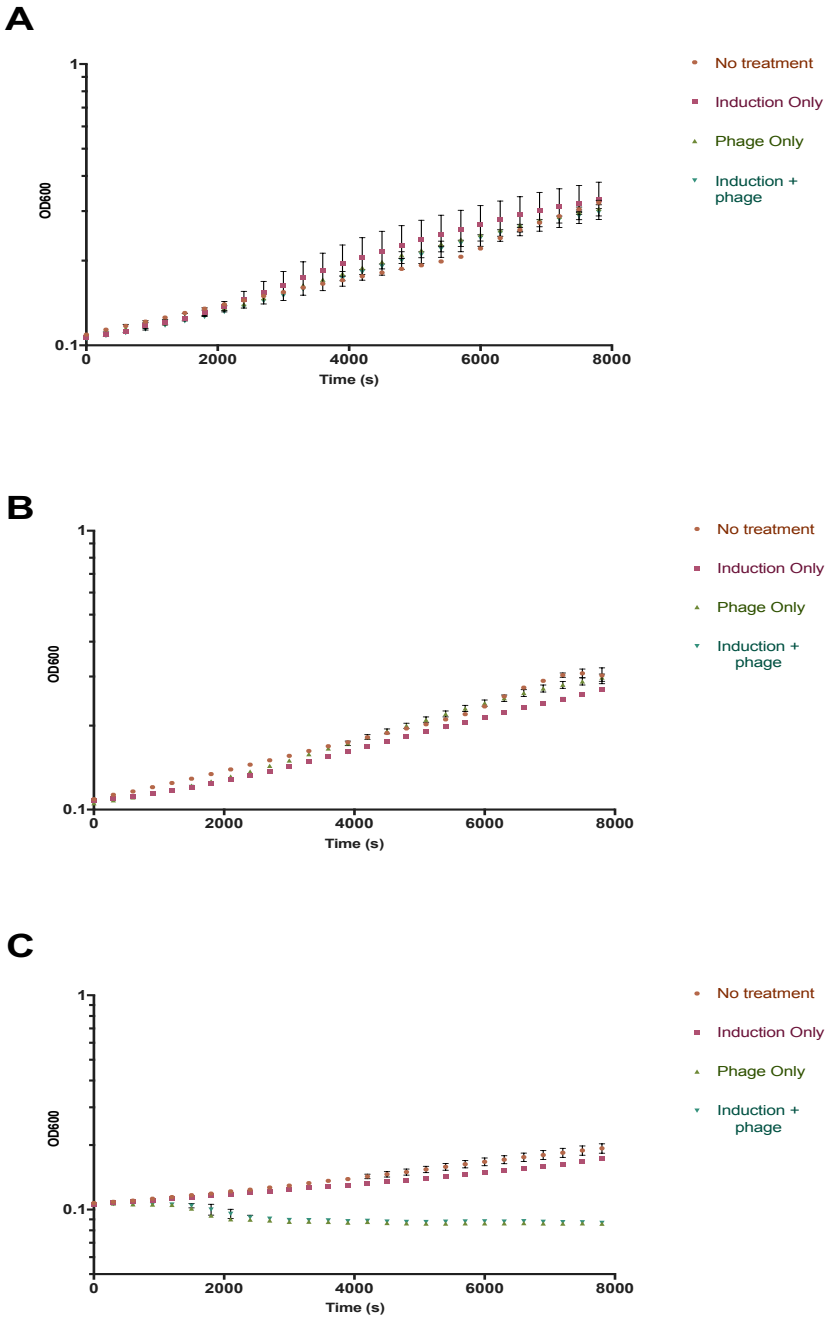


Figure 4-14. *E. coli ΔwaaC*/pSLprn32 challenge: growth of *E. coli* strains bearing pertactin expression constructs was assayed by OD600 and under the following conditions: no treatment, induction with 0.05mM IPTG, challenge with chimeric phage at an MOI of 0.1, and both IPTG induction and challenge with phage. **(A)** *E. coli ΔwaaC* cells. **(B)** *E. coli ΔwaaC* cells containing pertactin, pLprn34, *in trans*. **(C)** *E. coli* BW25113 cells. N=3.

4.2.4.4 Infectivity of T7 chimeric and BPP-1 variant phages on E. coli, B. pertussis and B. bronchiseptica

For all of the characterisation experiments, chimeric phage T7 were generated using the *E. coli* Δ*trxA* without *in trans* tail fibre complementation. This means that the samples were wild type T7 contaminates as identified in section 4.9; however, this was not an issue since wild type T7 does not infect either of the *Bordetella* strains and it will also be tested separately as a negative control. BPP-1, BMP-1, and BIP-1 phages were kindly sent by the research group of Dr Jeff F. Miller, based at the University of California, Los Angeles.

4.2.4.5 Host range of chimeric T7 phages

E. coli BW25113 was challenged with chimeric phages through a spot assay (Section 2.3.6). No spots were generated indicating that the chimeric tail fibres do not allow infection of *E. coli* Table 4-3. Phages BPP-1, BMP-1, and BIP-1 were also tested on *E. coli* and as expected, there was no indication of lysis Table 4-3.

Table 4-3. Summary of phages tested against *E. coli* BW25113.

Phage	Lysis/no lysis
phAG_1	no lysis
phAG_2	no lysis
phAG_3	no lysis
phAG_4	no lysis
phAG_5	no lysis
phAG_6	no lysis
phAG_7	no lysis
BPP-1	no lysis
BMP-1	no lysis
BIP-1	no lysis

4.2.4.6 Infection assays on B. bronchiseptica

In order to quantify the potential of chimeric phages to infect *B. bronchiseptica* it was necessary to test that T7 does not infect *B. bronchiseptica*. A dilution series of T7 was spotted onto lawns of *B. bronchiseptica.* Surprisingly it was found that T7 produced lysis on *B. bronchiseptica* (Figure 4-15). This was replicated three times for the two biological replicates of *B. bronchiseptica* with three different stocks of T7, all of which resulted in lysis. To ensure that the phage T7 stocks used were not contaminated the plaque from Figure 4-15 was used to perform PCR (amplification of *g17*) that confirmed the presence of T7; no TEM was used to further ensure the presence of T7. The *B. bronchiseptica* was taken directly from a frozen glycerol stock, created from the original stock sent by A. Preston (University of Bath); its genome was not verified using sequencing. To verify that T7 was causing the lysis, PCR was performed on all of the zones of lysis of *B. bronchiseptica* using T7 *g17* specific primers (Appendix 2) and then sequenced which confirmed the presence of T7 in the zones of clearing. To determine if the T7 on *B. bronchiseptica* underwent replication the following work was carried out. First, the amount of phage that theoretically was present in the 10^{-2} dilution zone of lysis was calculated, using the assumption that no replication was taking place. The stock concentration of the phage used was 4 x 10^8 PFU/ml. After spotting 10 ul of each of the phage dilutions 10^0, 10^{-1}, and 10^{-2}, PFUs were generated of 4 x 10^6, 4 x 10^5, and 4 x 10^4, respectively. A 10 ul pipette (0.033 cm in diameter) was used to stab the 10^{-2} dilution spot (1.66 cm in diameter), allowing ~1/50th of the phages from the spot to be transferred to 110 ul of medium. This resulted in a phage concentration of 8 x 10^3 PFU/ml, calculated as shown below:

4 x 10^4/50 = 8 x 10^2 PFU or 8 x 10^2 PFU/110 ul or **~8 x 10^3 PFU/ml**

After carrying out a plaque assay (see section 2.3.5) the empirical value for the phage concentration was found to be ~2.125 x 10^5 PFU/ml (Figure 4-14), which is ~27-fold greater than the theoretical phage concentration. This suggests that the amplification of phage was taking place to some extent.

BPP-1 and its variants generated no lysis on *B. bronchiseptica* indicating that there might have been an issue either with phage storage (the phage was stored at 4C for seven months prior to the characterisation assays) or with the phage stock received.

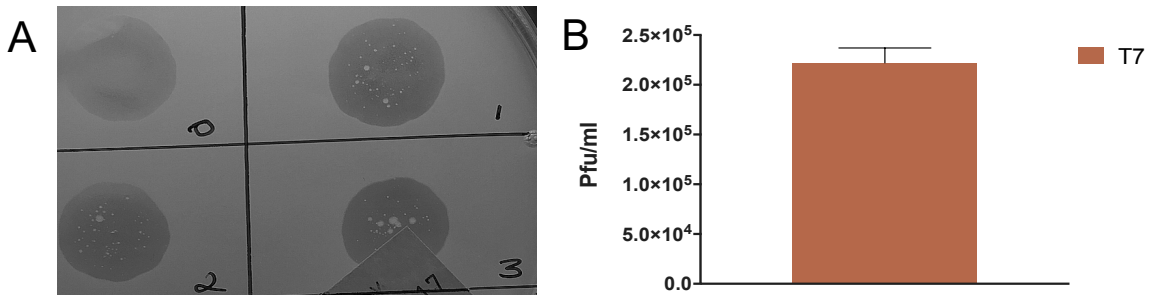


Figure 4-15. (A) T7 against B. bronchiseptica spot assay plated on LB agar. The T7 stock concentration used for the assay was 10^8 PFU/ml. **(B)** Phage enumeration. 2 ul of phage was collected from the second dilution of T7 against B. bronchiseptica spot assay and then used to re-infect *E. coli* BW25113.

4.2.4.7 Infection assays on B. pertussis

Given the characterisation data up to this point, it is important to note that the positive controls, BPP-1 and its variants, have not worked on *B. bronchiseptica.* T7, T7 chimeric phages (ph_AG1, phAG_2 and phAG_3) as well as BPP-1 and its variants were assayed on *B. pertussis* (received from A. Preston, no sequencing performed) using a spot assay (see section 2.3.5). It was found that T7, T7-ph_AG1, T7-phAG_2, T7-phAG_3 as well as BPP-1 caused zones of clearing, suggesting they infect *B. pertussis* Figure 4-16.

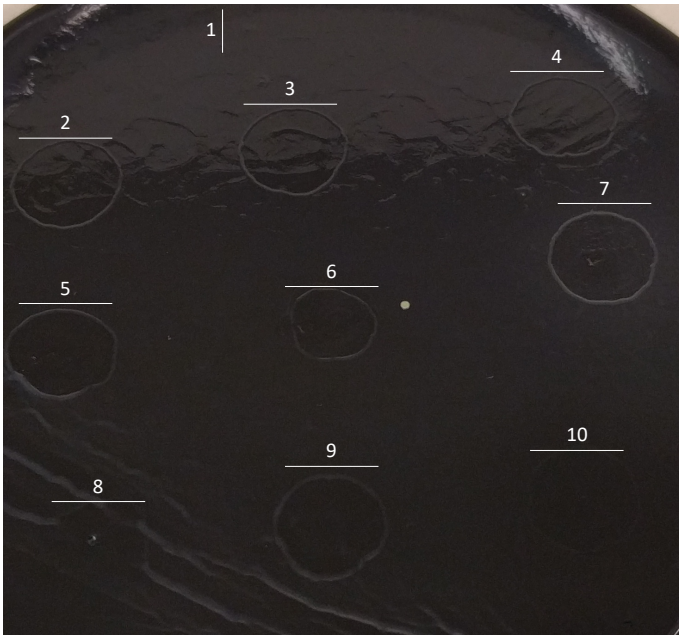


Figure 4-16. Spot assay of T7, T7 tail fibre chimeric as well as BPP-1, BPP-2, BPP-3 phages against B. pertussis. Lane 1 shows the negative only LB control, lanes 2-4 represent phAG_1, phAG_2 and phAG_3 chimeric bacteriophages, lanes 5-7 represent BPP-1-3 phage variants, and lanes 8-10 represent T7 three biological controls. The spot assay was carried out on charcoal agar.

4.3 Discussion

4.3.1 Fusion tail fibre design and insolubility

In vitro tail fibre analysis (section 4.2.3.1) showed that the fusion tail fibres are not soluble and, in some instances, unable to form trimers. This is a strong indication that the fusion tail fibres either no longer form trimers due to instability, or that the Gp17 and Mtd domains fused, resulting in misfolded insoluble fusion tail fibre aggregates[249,361]. One potential cause of Mtd instability might be that only keeping two (sandwich and tip domains) out of three Mtd domains (prism, sandwich and tip domains) for the fusions was not sufficient to retain tail fibre stability and trimerization. The 'prism' domain, which wasn't included in either A or B fusions, might be required for stability as it has a tight trimeric organisation when compared to the sandwich domain that branches away from the prism domain trimer axis. Therefore, potentially leaving the prism domain in when designing protein tail fibre fusions could aid the trimerization. However, leaving in the prism domain would also increase fusion tail fibre length. In the previous study it was suggested that upon tail fibre binding to the target bacterium receptor, the tail fibre reaches a confirmation that allows opening of the internal tail channel, resulting in DNA injection and capsid protein release[364]. Significant length alterations of the tail fibre could potentially interfere with the formation of the correct confirmation. Furthermore, even though the tip domain has a potentially independent fold from that of other Gp17 domains when the trimer is formed, it was also suggested that the tip domain monomer formation could initiate the trimer formation of the entire tail fiber tail fiber[356]. This could happen by the tip domain initially forming a monomer with a spontaneous formation of a monomeric carboxyl-terminal β-barrel[356]. The interaction of three β-barrels could then lead to the trimer formation[356]. This suggests that by altering the tip of the tail fber may interfere with its ability to form trimers.

4.3.2 Continuous contamination of chimeric phage with T7

When *E. coli* Δ*trxA* without *in trans* complementation was used for phage selection, wild type T7 was found after multiple rounds of plaque purification. The chimeric phage should make up the majority of the phage particles since it is able to undergo replication due to genomic insertion of *trxA*, while the T7 phage lacks *trxA* and is therefore is incapable of replication[257,260,365]. If the T7 persistence was to be due to contamination, multiple plaque purification steps should have removed it due to the contaminant being diluted out[366].

Furthermore, it was found that on its own, the phage T7 titre was reduced from 4 x 10^8 PFU/ml to <1 PFU/ml when plated on *E. coli* Δ*trxA*, whereas in this case the titre of 4 x 10^8 PFU/ml was reduced to 10^2 PFU/ml. This suggests that there are additional factors contributing to the persistence of T7. When comparing different selection methods, it was found that using *E. coli* Δ*trxA* resulted in visibly smaller plaques when compared to using *E. coli Δtrx/*pSEVA551*,* indicating a somewhat reduced phage burst size[367,102]. It is unlikely that this is due to homologous recombination between the wild type and the chimeric phage, as both wild type and chimeric variants were detected in plaques (Figure 4-12). In addition, the *in vivo* characterisation suggested that the tail fibre fusions are most likely not folded properly and therefore not functional. A potential explanation for this result could be a co-infection between the wild type and the chimeric phage, where one would provide correct tail fibres and the other provide the *trxA* gene required for replication. However, it is currently thought that T7 excludes coinfection[368].

4.3.3 Phage testing against *B. pertussis* and *B. bronchiseptica*

T7, T7 tail fiber chimeric mutants and BPP variants yielded zones of clearance while there was no clearance in the negative control sample of LB only, suggesting that all phage types generated lysis. Attempts to test phage variant infectivity of *B. pertussis* were complicated by poor *B. pertussis* lawn formation during all assay attempts. Additional attempts by A. Preston's lab (University of Bath, unpublished data) to grow *B. pertussis* also resulted in similar poor lawn formation. Therefore, the results for the phage infectivity against *B. pertussis* are inconclusive.

In the case of phage testing against *B. bronchiseptica* it was found that the T7 negative control not only generated bacterial lysis but also achieved phage amplification. Given the differences between the *E. coli* and *B. bronchiseptica* LPS surfaces, this result was surprising, if not unprecedented[369]. There are several reports of broad host range phage[107,370,371]. However, T7 has only been reported to infect *E. coli*, rough *S. typhimurium* and some *Shigella* strains[223,372]. A possible explanation for the observed *B. bronchiseptica* lysis by T7 could be due to the lysis from without phenomenon, where multiple phage particles adhering to the bacterial cell result in cell death without any phage particles having replicated[41]. However, this possibility can be ruled out since it was calculated that after *B. bronchiseptica* lysis by T7 the resulting phage

increased by 100-fold, suggesting that it was not lysis from without due to observed phage amplification.

Chapter 5 Bacteriophage T7 Co-infection

5.1 Introduction

In chapter 4, wild type T7 contamination was detected after numerous chimeric phage purification steps, suggesting co-infection between the two phages. Current literature suggests this would not be the case due to a widely accepted interpretation of T7 superinfection exclusion principle, where the entry of the first phage into a bacterial cell inhibits the entry of a secondary phage[368,373]. Although this conventional view of T7 infection lacks an elucidation of the precise exclusion mechanism, multiple conflicting studies have sought to interpret the mechanistic features causing this exclusion.

5.1.1 Superinfection exclusion of T7 homologous phages

There are multiple studies supporting the superinfection exclusion phenomenon. In one study a dual strategy was used to determine if the secondary attachment after primary infection is excluded from replication[373]. First, the conservative transfer of ^{32}P labelled parental DNA was used to show that un-replicated parental DNA molecules are not linearly related to the number of phage adsorbed per cell, and that only a limited number of phages can replicate their DNA in each cell[373]. Secondly, the primary and secondary infections were examined using two different T7 amber mutants. The results revealed that if the amber mutants were added simultaneously the productive infection is equal to that of the number of bacteria infected. However, if the secondary amber mutant phage is added seven or more minutes post infection the number of productive infections is not influenced at all; it can no longer complement the primary infecting phage, supporting the T7 superinfection exclusion principle[373]. The same study also suggested that addition of chloramphenicol can reduce the exclusion principle, implying that the superinfection exclusion takes place after the initial protein synthesis of the primary phage[373]. Moreover, in a further study the chloramphenicol method for superinfection exclusion prevention was used to show that UV-irradiated wild-type T7 phage markers can be rescued by coinfection with a T7 amber mutant[374]. Superinfection exclusion inhibition was also demonstrated using alkylation of primary T7 phage by methyl methane sulfonate as it interfered with superinfection exclusion[375]. This interference could be potentially correlated with a delay in the synthesis of phage-specific proteins. Therefore, it has been suggested that protein synthesis directed by the primary infecting phage is key for efficient exclusion of superinfecting secondary phage particles. In contrast, it has been suggested that

chloramphenicol treatment does not prevent superinfection exclusion and that the phenomenon does not depend on protein exclusion[368].

5.1.2 Superinfection exclusion of heterologous phages

Mixed infection of *E. coli* with phages T1 and T2 results in the exclusive development of T2[376]. Similarly, mixed infection by T3 and T7, closely related coliphages, usually gives rise to the exclusive development of either T3 or T7, with both phages generated only rarely[368]. In some instances two phages that share the same host cell receptor but are genetically different maybe able to co-infect[377]. As demonstrated by infection of *E. coli* cells with ΦX-174 at an MOI of 10 is insufficient to inhibit phage T7 progeny production, suggesting ΦX-174 does not interfere with T7 adsorption[377].

Heterologous superinfection has been systematically investigated by making early gene mutants of T7 that would interfere with their function[368]. These T7 mutants were then used as primary phages followed by infection with T3 as a secondary phage[368]. The most effective of the early genes in preventing superinfection appeared to be *g1* and *g0.3*[368]. In the case of *g1*, it was not clear whether the exclusion was due to the action of *g1* coding phage RNA polymerase itself or if it was the result of the genes expressed by it[368,378]. In addition, to explain the superinfection exclusion between T7 and T3 phages, it was suggested that binding of the heterologous RNA polymerase enzyme to the DNA of co-infecting phage particles may interfere with the transcription of this DNA by the homologous enzyme[379].

5.1.3 Aims

The aim of the chapter is to determine if co-infection is possible between T7 and recombinant phage mutants. To meet this aim the chapter has been divided into following sub-aims:

1) Further investigation into two phage co-infection in solid medium;
2) Investigation into two phage co- in liquid medium.

5.2 Results

5.2.1 Investigation of chimeric and wild type phage infection on solid medium

To investigate the interaction between T7 and chimeric phages, T7-CH1, T7-*g17::trxA*, and T7 phages were used. T7-CH1 as described in Figure 4-3 and Table 4-2, is a tail fiber chimeric phage, where the tip domain of T7 Gp17 was replaced with the C-lectin domain of Bpp-1 Mtd. The gene sequence of the C-lectin domain was inserted together with *trxA* gene that confers selection advantage for engineered phage. T7-*g17::trxA* phage had the *g17* tail fiber gene completely replaced by *trxA* gene. Each of the phages was amplified, purified (Section 2.3.3 and 2.3.13) and the stock for each was diluted to approximately 1 x 10^8 PFU/ml. Either T7-CH1 or T7-*g17::trxA* was mixed with T7 in 1:1, 1:0.1, 1:0.01 and 1:0.001 ratios (the number of total phage was kept the same (10^8 PFU/ml) throughout all ratios and used to perform plaque assays on *E. coli* Δ*trxA* (section 2.3.5). For the T7-CH1 and T7 mixture, the 1:1 ratio resulted in the highest number of progeny of approximately 10^4 PFU/ml. When one of the phages (either T7 or T7-CH1) was reduced by 10-fold the PFU reduced by approximately 5-fold (Figure 5-1). An even greater decrease was seen when one of the phages was reduced 100 or 1000-fold, approximately 10 and 100-fold respectively. A similar trend is seen for the ratios between T7-*g17::trxA* and T7 phages. This result supports suggest the phages are dependent on each for replication.

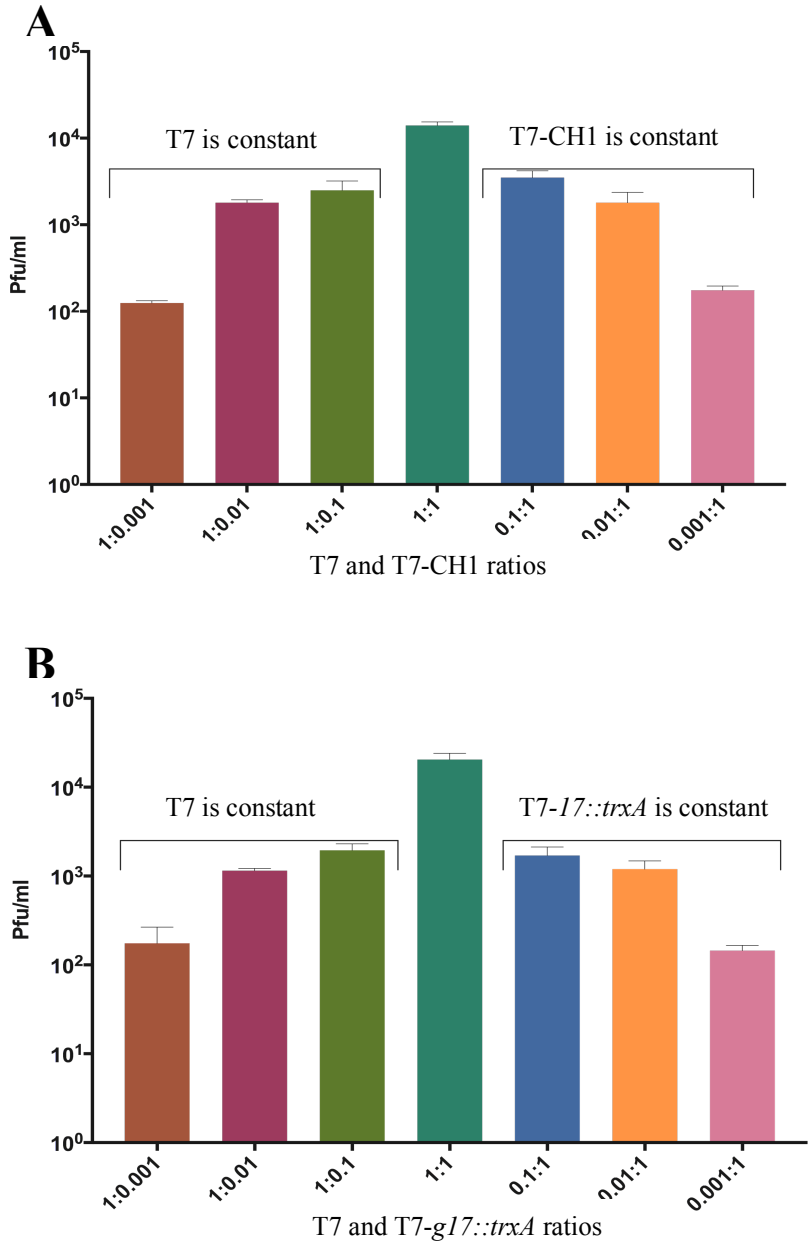


Figure 5-1. Phage enumeration for different chimeric and T7 phage ratios. T7 and chimeric phage stocks were diluted to approximately 10^8 PFU/ml and mixed in the following ratios: 1:0.001, 1:0.01, 1:0.1, 1:1 for each phage. Plaque assays were performed, for each of the phage ratios; phages were then enumerated. **(A)** Phage enumeration for T7 and T7-CH1 ratios. **(B)** Phage enumeration for T7 and T7-*g17::trxA* ratios. The PFU/ml data are presented as the mean of three independent experiments, n=3.

5.2.2 Investigation into chimeric and T7 phage interaction in liquid medium

In order to deduce if the two phages are co-dependent on each other for replication in liquid medium, the following experimental set-up was used. Bacterial growth/phage lysis assays (Section 2.3.14) at MOIs of 0.001, 1 and 3 with the following conditions: *E. coli* Δ*trxA* cells only, *E. coli Δtrx*A cells with phage T7, *E. coli* Δ*trxA* cells with the chimeric phage T7-CH1 and *E. coli* Δ*trxA* cells containing both T7 and T7-CH1 phages (Figure 5-2). No lysis occurred when either T7 or T7-CH1 was added to *E. coli* Δ*trxA* cells at an MOI of 0.001 or 1, but did occur for T7 at an MOI of 3. At this higher MOI some lysis occurred at ~40 minutes post infection, but *E. coli* growth was quickly resumed at 50 minutes. When both T7 and T7-CH1 were added to *E. coli* Δ*trxA* cells lysis occurred for all samples. This result in Figure 5-2 potentially occurred due to the fact that T7 was not able to replicate in *E. coli* Δ*trxA*, whereas T7-CH1 was unable to adsorb due to the altered tail fiber. To further verify that the replication only occurred when both T7 and T7-CH1 phages were present, the resulting phages after the MOI 0.001 growth/lysis assay (Figure 5-2) were enumerated using spot assay (section 2.3.6). Both T7 and T7-CH1 only samples resulted in 10^4 PFU/ml whereas the sample that contained both phages generated 10^7 PFU/ml (Figure 5-3) The original amount of phage added was approximately 10^5 PFU/ml, therefore combining both phages results in replication.

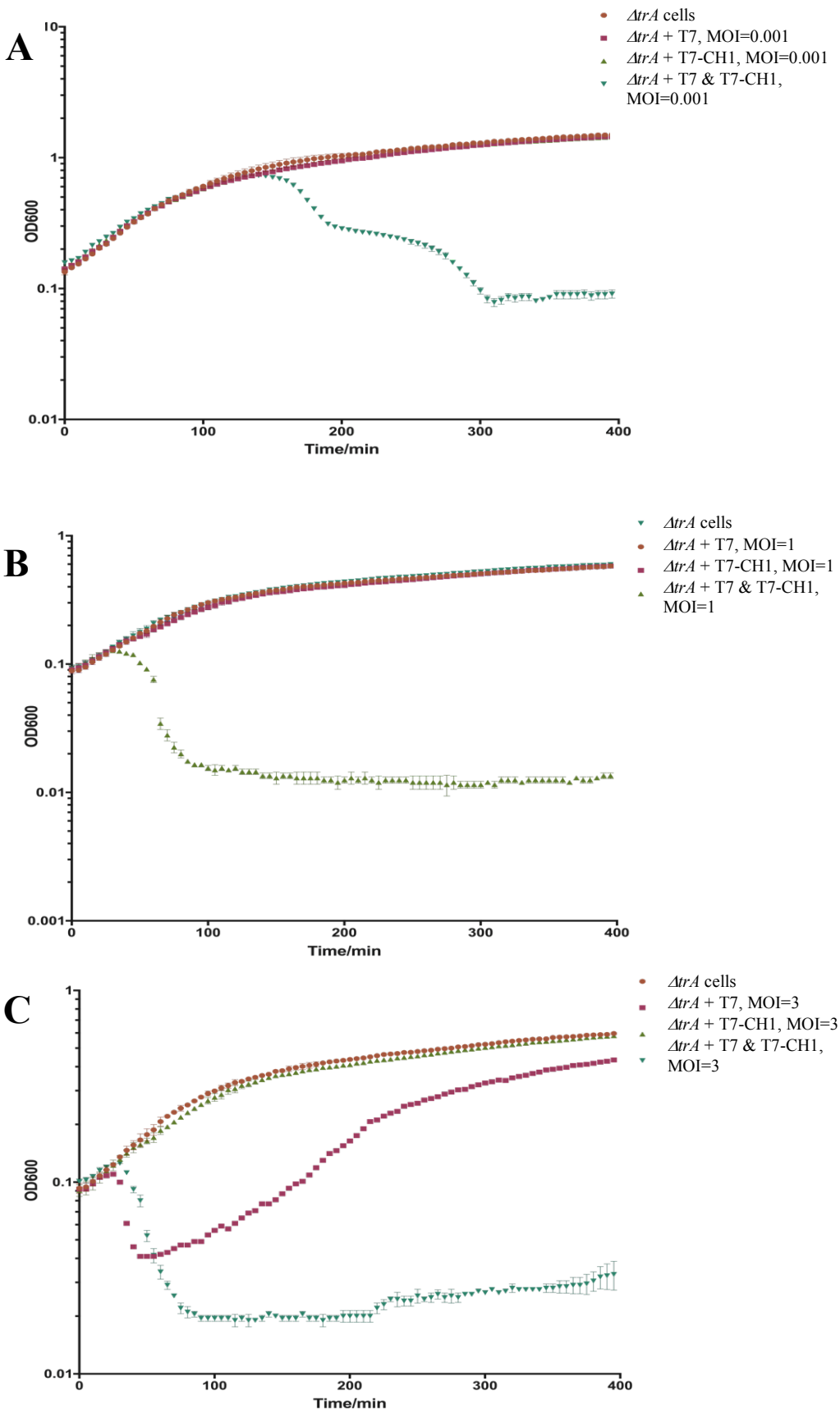


Figure 5-2. Bacterial growth/phage lysis assays. Each of the assays had the following conditions: *E. coli* Δ*trxA* cells only, *E. coli* Δ*trxA* cells with wild type T7 (unable to replicate) added, *E. coli* Δ*trxA* cells with T7-CH1 (unable to adsorb) added, and *E. coli* Δ*trxA* cells with wild type T7 and T7-CH1 added. **(A)** Phages were added at an MOI of 0.001 (n=3 for each condition except for *E. coli* Δ*trxA* cells with wild type T7 and T7-CH1, which were n=2 due to lack of lysis in the third biological replicate). **(B)** Phages were added at an MOI of 1 (n=3 for each condition). **(C)** Phages were added at an MOI of 3 (n=3 for each condition).

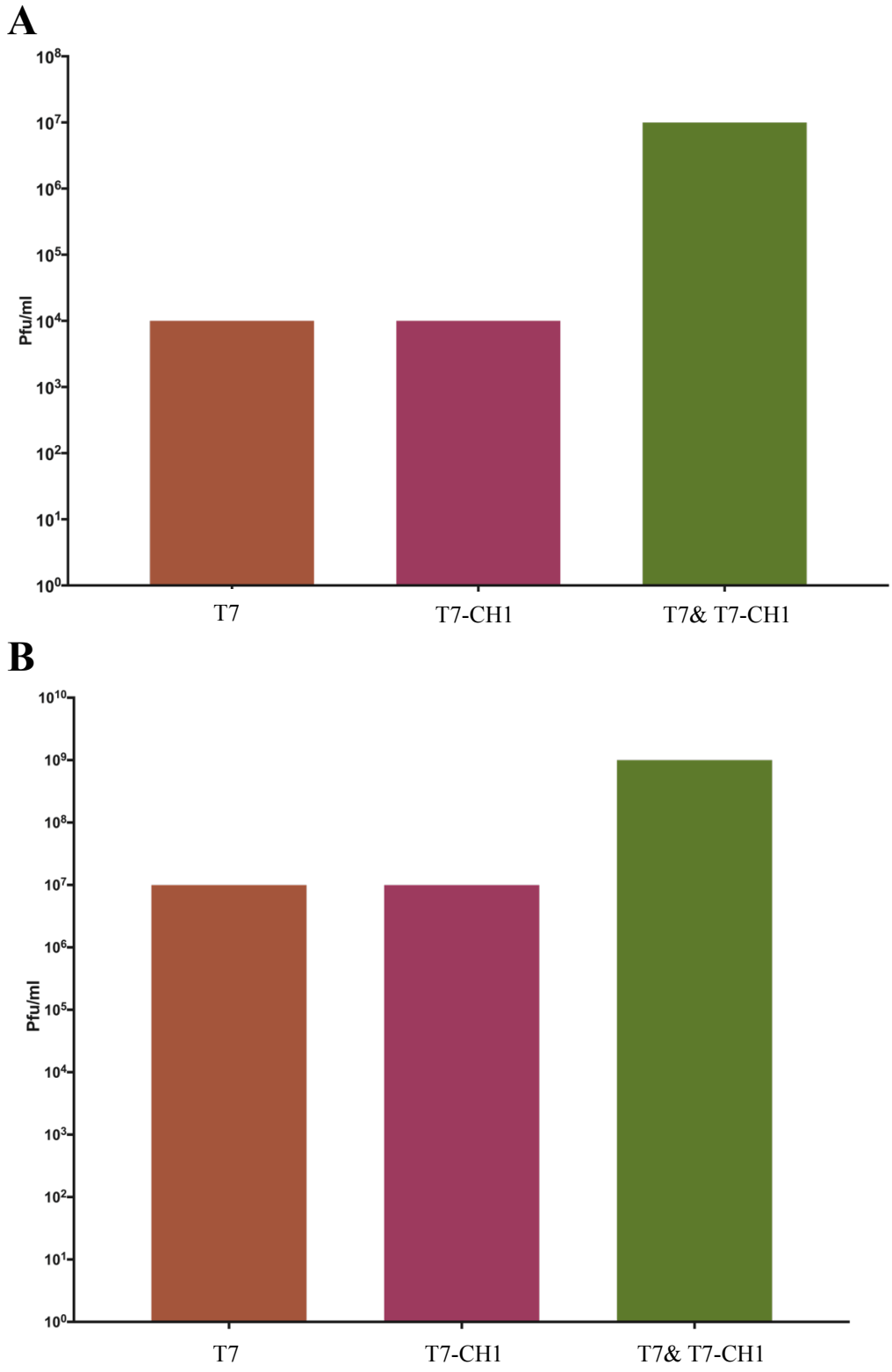


Figure 5-3. Phage enumeration was performed after the growth/lysis assays for MOIs of 0.001 and 3 presented in Figure 5.2.2 -1 on *E. coli* BW25113 cells. After completing the assay described in 5.2.2, plaque assays were performed on three samples from the plate reader wells: *E. coli ΔtrxA* cells containing T7 only, T7-CH1 only and T7 and T7-CH1 combined. Phages were enumerated using spot assay and the PFU/ml data are presented for one biological replicate per well. **(A)** Phage recovered from Figure 5.2.2 -1**,** A (MOI 0.001); the original amount of phage added was approximately 1 x 10^5 PFU/ml. **(B)** Phage recovered from Figure 5.2.2 -1, C (MOI of 3); the original amount of phage added was approximately 1 x 10^8 PFU/ml, n=1.

5.2.3 Removal of T7 replication using CRISPR-Cas type II

The results of Section 5.2.2 demonstrated that co-infection was required to produce lysis, it was reasoned that if the replication of one of these phages could be halted the replication of both phages would be stopped. This was tested using the previously developed CRISPR system (Section 3.3.2). Since CRISPR only shows 1000-fold reduction in T7 phage plaque formation when using the most efficient gRNA (Section 3.3.2), it was reasoned that CRISPR would only be effective at 'removing' the co-infection at MOIs lower than 1. In order to verify this the following conditions were set-up: *E. coliΔtrxA/*pAG_13 (gRNA2), *E. coliΔtrxA/*pAG_12 (non-targeting gRNA(scr)) and *E. coli ΔtrxA*. For each of the cell types one of the following was added: T7, T7-CH1 or both, at MOIs of 0.001 and 0.01 (Section 2.3.14). At all MOIs no lysis was observed when just T7 or phage T7-CH1 were added to *E. coli ΔtrxA* or *E. coli ΔtrxA*/pAG_13 cells, and lysis only occurred when T7 and T7-CH1 were added in combination to *E. coli ΔtrxA* cells. When the phages were combined and added to *E. coli* ΔtrxA/pAG_13 cells that contain a CRIPSR system to target T7 replication, no lysis was observed at any MOI (Figure 5-4). Thus, stopping the replication of one phage also stopped the replication of the phages combined.

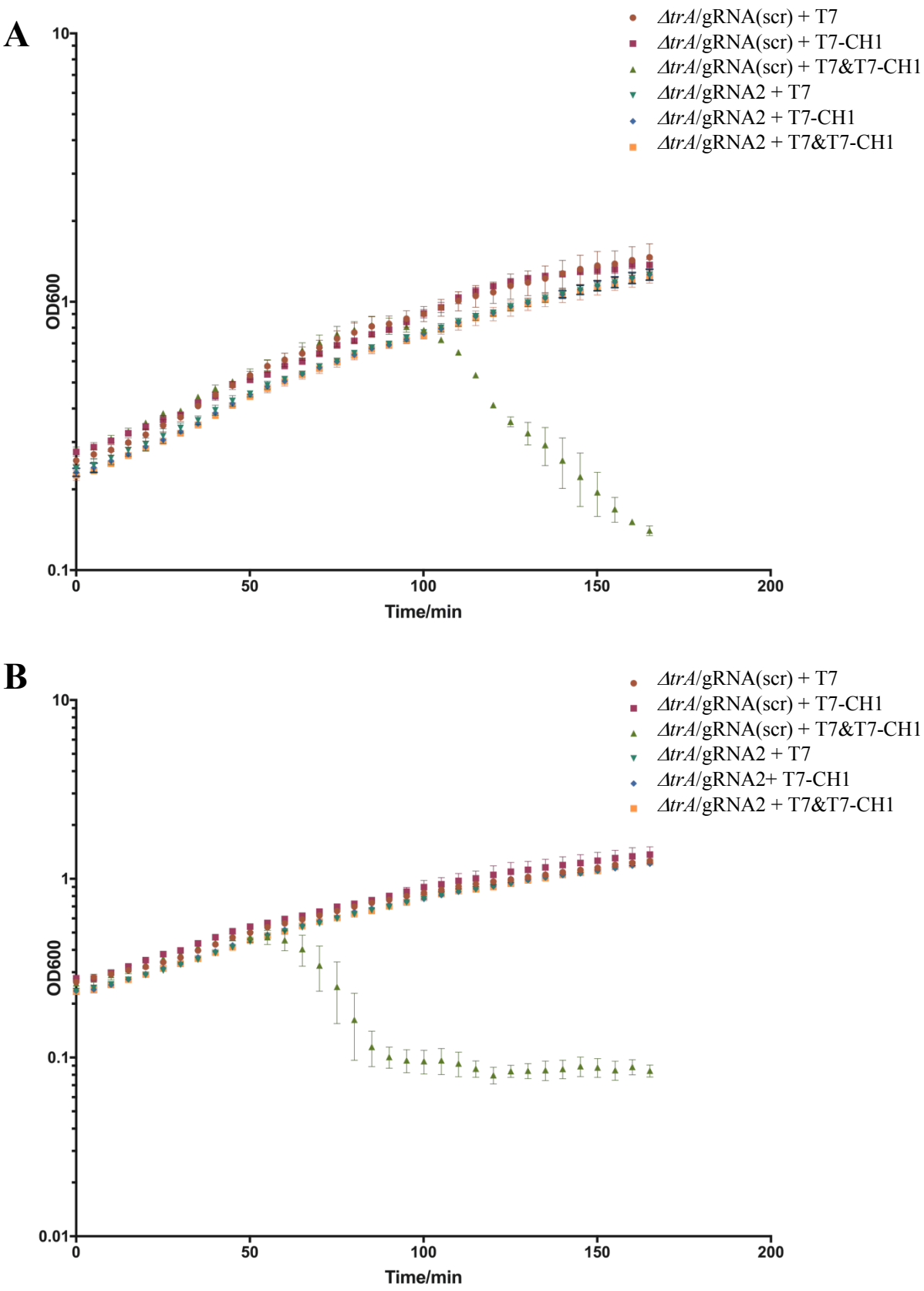


Figure 5-4. Bacterial growth/phage lysis assays. Each assay contains two cell types: *E. coli* Δ*trxA* cells expressing either the CRISPR type II system with a non-targeting scrambled gRNA (noted as *E. coli* Δ*trxA* + gRNA(scr)) or the CRISPR type II system with a targeting gRNA2 (noted as Δ*trxA* + gRNA2). In each of the cell types wild type T7, T7-CH1 or both wild type T7 and T7-CH1 were added. **(A)** Phages were added at an MOI of 0.001 (n=3 for each condition except for Δ*trxA* + gRNA(scr) with wild type T7 and T7-CH1 added, which was n=2 since the third biological replicate did not show lysis). **(B)** Phages were added at MOI of 0.01 (n=3 for each condition).

5.2.4 Delayed addition of a secondary phage effects on the potential co-infection

In order to investigate whether both phage need to be added at the same point for co-infection to occur, the addition of the second phage was delayed. The primary phage was added at an MOI of 0.5 to *E. coli ΔtrxA* cells followed by addition of a secondary phage at an MOI of 0.5 after 15, 30, 60 and 120 minutes marked as arrows in Figure 5.2.4-1 (Section 2.3.14). When the T7 was the primary phage the addition of T7-CH1 phage after 15, 30, 60 and 120 resulted in lysis after approximately 70, 90, 130 and 210 minutes respectively. When the reciprocal experiment was carried out, lysis occurred after 60, 90, 130 and 205 minutes respectively (Figure 5-5). This result further confirms that both phages are required to allow phage replication and lysis of the cells. In addition, this verifies that for co-infection to take place primary and secondary phages do not need to enter the cell at the same time.

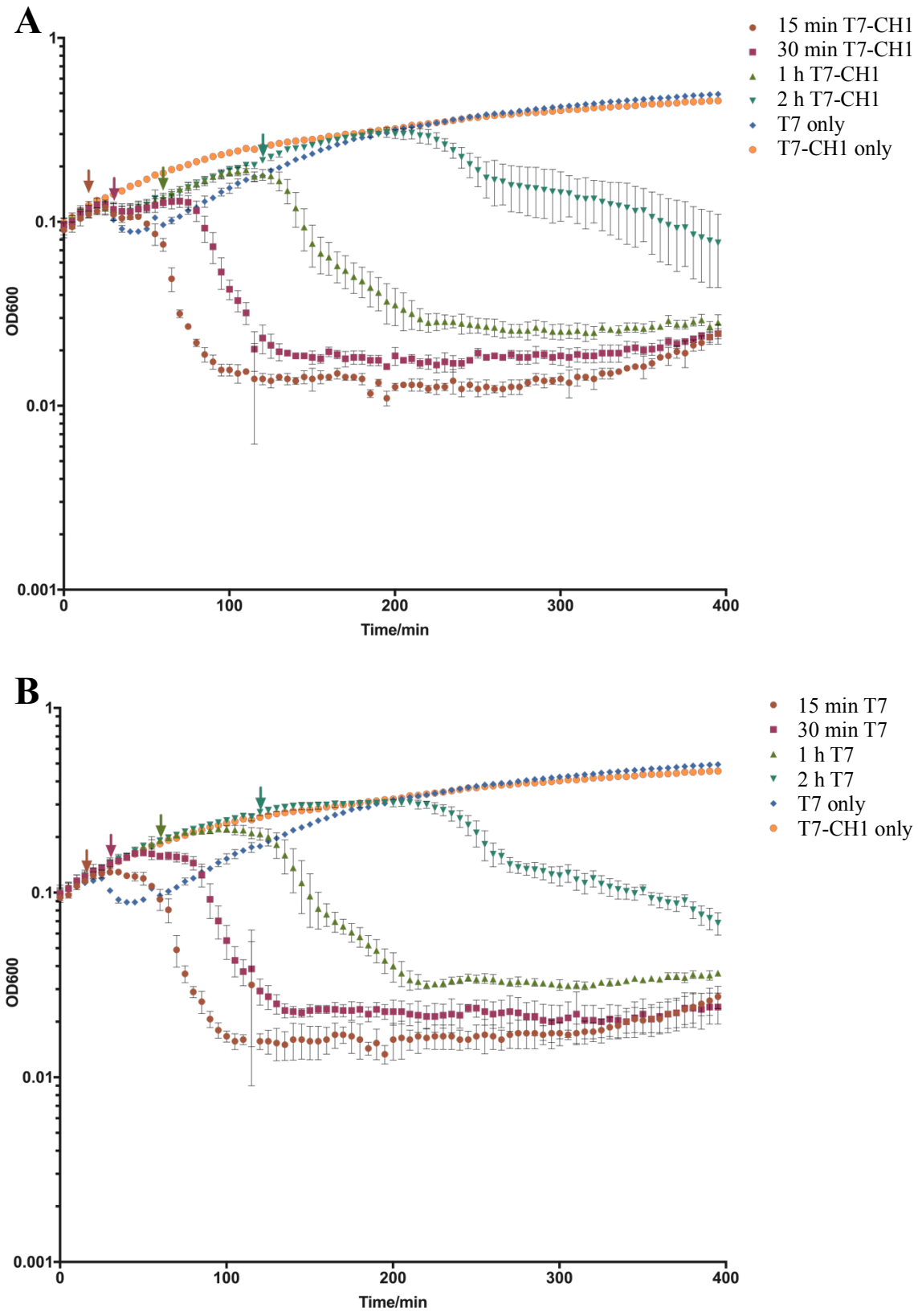


Figure 5-5. Bacterial growth/phage lysis assays in *E. coli* Δ*trxA* cells at an MOI of 1. **(A)** T7 and T7-CH1 are used as primary and secondary phages, respectively. At the start of the assay wild type T7 was added to four samples at an MOI of 0.5, followed by T7-CH1 at an MOI of 0.5 after either 15 min (15 min T7-CH1), 30 min (30 min T7-CH1), 1h (1 h T7-CH1), or 2h (2 h T7-CH1). Two additional samples only contained either wild type T7 or T7-CH1, added at the start of the assay at an MOI of 1. **(B)** T7-CH1 and wild type T7 are used as primary and secondary phages, respectively. At the start of the assay T7-CH1 was added to four samples at an MOI of 0.5, followed by wild type T7 at an MOI of 0.5 after either 15 min (15 min T7), 30 min (30 min T7), 1 h (1 hr T7), or 2 h (2 hr T7). Two additional samples only contained either wild type T7 or T7-CH1, added at the start of the assay at an MOI of 1. Arrows represent time points of phage addition. N=3.

5.2.5 Construction and characterisation of T7-*g5*::*trxA* mutant for further co-dependency verification

To investigate if co-infection only occurrs when targeting a tail fiber gene, a T7 mutant was constructed lacking *g5* which encodes DNA polymerase. To achieve this, vector pAG_60 (Appendix 1), containing *trxA* in between 100 bp long homologous recombination arms (before and after *g5*), was constructed using Gibson assembly (Section 2.1.7). This vector was then used to generate the T7-*g5::trxA* phage using homologous recombination (Section 2.3.11), followed by selection and purification using *E. coli ΔtrxA* with *g5* (pET24-gp5, see Appendix 1) *in trans* as well as *E. coli ΔtrxA* only. In the latter, after multiple rounds of purification, PCR was performed on the plaques (Figure 5-6) with the resulting PCR product examined using gel electrophoresis and sequenced. Both T7 as well as T7-*g5::trxA* phage corresponding bands were found, replicating the result previously seen for the tail fiber chimeras (section 5.2.1). The next step was to perform a ratio experiment (Section 5.2). As seen previously for the tail fiber mutant, the 1:1 ratio resulted in the highest amount of phage enumerated, approximately 5 x 10^5 PFU/ml (Figure 5-6). When one of the phages was reduced by 10-fold the resulting PFU was reduced approximately by 10-fold. Even greater PFU drops were seen when one of the phages was reduced 100 or 1000-fold, approximately 100 and 1000-fold respectively. To test the co-dependency in liquid medium, the experimental set-up used in section 5.2.2 was used, with assays performed at MOIs of 1 and 0.001 (Figure 5-7). As seen previously, lysis did not occur when either T7 or T7-*g5::trxA* phages were added to *E. coli* Δ*trxA* cells at either MOI, but did occur when both T7 and T7-*g5::trxA* phages were added to *E. coli* Δ*trxA* (Figure 5-7). These findings further support the replicative co-dependency between the two.

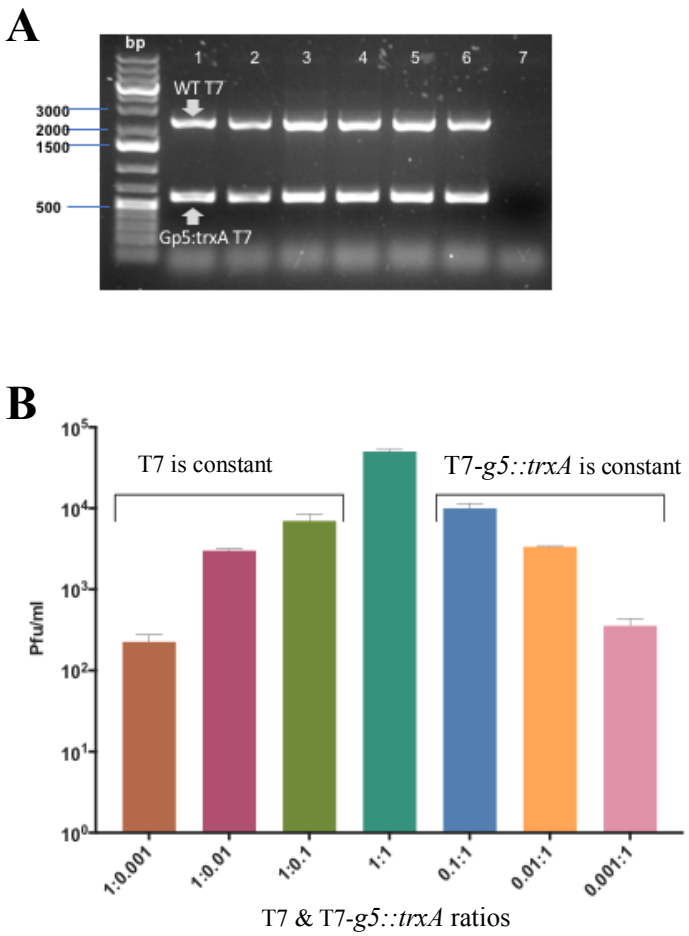
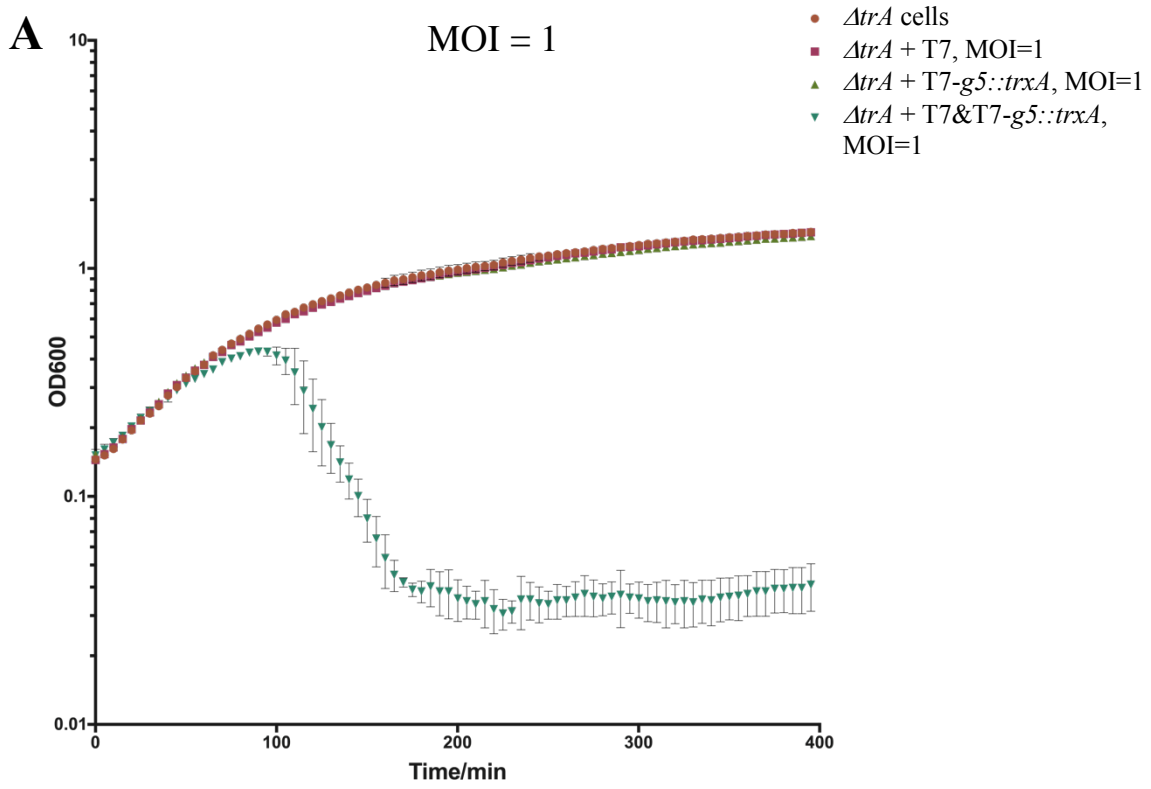


Figure 5-6. Evidence for T7-*g5::trx*A and wild type T7 co-dependence. **(A)** Chimeric phage verification in which the phage *g5* region from six *E. coli ΔtrxA* plaques was amplified using PCR with the products examined via gel electrophoresis. All six plaques (lanes 1 – 6) contained both T7 (~2 kb) and chimeric (~600 bp) *g5* regions. The far left lane, marked 'bp', contains the reference ladder GeneRuler 1 kb Plus DNA while the far right lane, lane 7, contains the negative control (no DNA template). **(B)** Phage enumeration of wild type T7 and chimeric T7-*g5::trxA* phage ratios. Phage stocks were diluted to approximately 10^8 PFU/ml and mixed in T7:chimeric ratios of 1:0.001, 1:0.01, 1:0.1, 1:1. Plaque assays were performed for each of the phage ratios, followed by phage enumeration. The PFU/ml data are presented as the mean of three independent experiments, n=3.



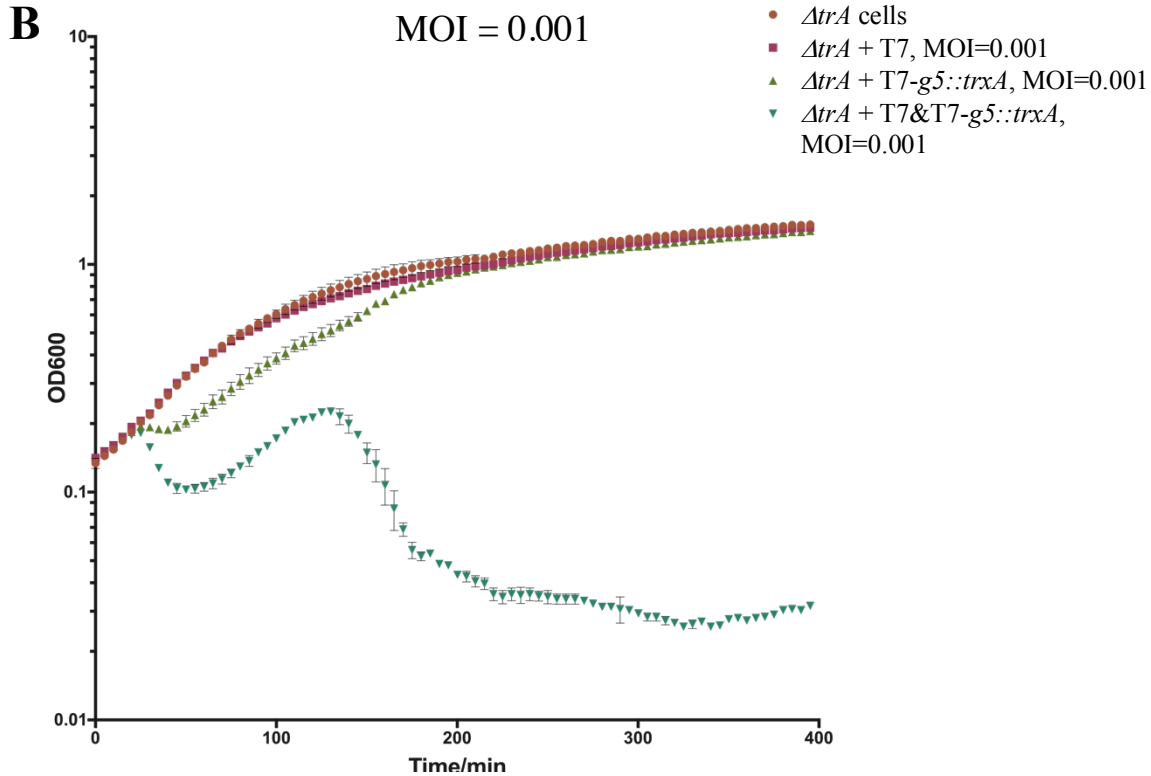


Figure 5-7. Bacterial growth/phage lysis assays. Each assay had the following conditions: *E. coli* Δ*trxA* cells only, *E. coli* Δ*trxA* cells with T7 added, *E. coli* Δ*trxA* cells with T7-*g5::trxA* added, and *E. coli* Δ*trxA* cells with T7 and T7-*g5::trxA* added. **(A)** Phages were added at an MOI of 1 (n=3 for each condition). **(B)** Phages were added at an MOI of 0.001 (n=3 for each condition).

5.3 Discussion

5.3.1 Stochasticity of infection at lower MOIs

In section 5.2.2 liquid medium experiments in *E. coli* Δ*trxA* cells, the samples containing both T7 and T7-CH phages yielded lysis in comparison to the samples containing each of the T7 and T7-CH phages separately at the same MOIs. However, at the lowest MOI of 0.001 tested one in three biological controls did not result in lysis (Figure 5-2). This could be potentially explained by very low total phage numbers present in the system making the 'lysis' outcome a low probability event, and somewhat stochastic, as both T7 and T7-CH need to enter the cell to generate the lysis. The probability of two distinct phage particles being able to enter a cell at the same time and at a given MOI was previously described using a Poisson distribution using the following formula: $P(n) = \frac{m^n \times e^{-m}}{n!} \times initial\ PFU$, where m is MOI of total phage added and n is number of distinct phage particles[380]. This formula is written under assumption that both particles enter the cell at the same time. Therefore, in this co-infection system at total MOI of 0.001 (m = 0.001) and the two particles were T7 and T7-CH phages (n = 2). The probability of successful co-infections at MOI of 0.001 seems to be relatively low $actual\ P(2) = 0.05$ (Figure 5-8).

(1) $Actual\ P(n) = \frac{m^n \times e^{-m}}{n!} \times initial\ PFU$

(2) $P(2) = \frac{0.001^2 \times e^{-0.001}}{2!} = \frac{1 \times 10^{-6} \times 9.9 \times 10^{-1}}{2} \approx 5 \times 10^{-7}$

(3) $initial\ PFU = 1 \times 10^{-3}\text{MOI} \times 2 \times 10^{8}\text{CFU} = 1 \times 10^{5}\text{PFU}$

(4) $Actual\ P(2) = 1 \times 10^{5}\text{PFU} \times 5 \times 10^{-7} = \mathbf{0.05}$

m = total MOI
n = no of phage particles

Figure 5-8. Determination of probability of successful co-infections taking place at MOI 0.001. **(1)** equation describes the relationship at two particles being able to infect a given cell under an assumption that they infect at the same time; this is then multiplied by the initial number of phages initial added. **(2)** Working out the P(2). **(3)** Working out the initial amount of phage added at MOI 0.001 and at CFU $2x10^8$. **(4)** Working out the Actual P(2).

This could explain the difference in outputs for the three biological controls at MOI = 0.001. Given that the *E. coli* cells are doubling at approximately every 25 min and the very low probability of successful co-infection, the *E. coli* may outpace the number of phages generated. This stochasticity was not present in the samples with higher MOIs because greater amount of each phage was present which meant a higher probability that both of the phages will enter the cell and yield lysis overcoming the growth rate effect. In the case of MOI = 1, the $actual\ P(2) = 0.18$. Therefore, the probability of successful co-infections increases with the higher MOI applied.

5.3.2 Partial T7 lysis of *E. coli ΔtrxA* at high MOIs

The addition of T7 to *E. coli ΔtrxA* cells at an MOI of 3 resulted in some reduction of growth followed by a quick recovery of growth. This could potentially be due to the fact T7 enters the cells, but replication does not occur (or at a very low level) since the *trxA* is absent from the cells, but lysis of the cells, due to phage entry occurs[381]. Once T7 enters the cell its genes are transcribed and later translated. The product of one of the early genes *g0.7* inactivates transcription of multiple host genes as well inhibits host proteins, including RNAse III and E, and *E. coli* polymerase[234]. This shuts down the *E. coli* metabolism preventing the bacterium from functioning, reproducing and eventually resulting in its death. The bacterial death after phage infection in such way could serve as phage sink. T7 would be able to enter its host, however, since close to no replication was taking place the concentration of holin and lysin in the host cytoplasm may not be sufficient to release the progeny. Therefore, the cell would undergo death prior to releasing T7. One of the potential ways to validate that T7 entry alone (without phage amplification) causes cell death could be to generate T7 mutant phage lacking in *g0.7.* Addition of such phage to *E. coli ΔtrxA* cells should less hinder bacterial growth since it would prevent host protein inhibition and show a better growth when compared to growth of *E. coli ΔtrxA* cells with T7 added.

Another possibility is that T7 progeny is released due to eventual accumulation of lysin and holin genes in *E. coli*, however, since T7 replication is disabled, the quantity of the progeny is significantly reduced and eventually outcompeted by *E. coli* growth, hence the bacterial growth after the increase at time 40 mins, Figure 5-2. This is supported by the spot assay (Figure 5-3) that shows no T7 phage number increase confirming no replication taking place. In the previous work, it has been shown that no replication takes place when T7 is used to infect *E. coli ΔtrxA*

cells, verifying that *trxA* is essential for T7 replication, supporting the results obtained in this chapter[328]. In another, *in vitro*, study, it has been shown that in absence of *trxA* T7 DNA polymerase sustained 1% of its original activity. This may not directly translate to *in vivo* results seen this chapter. However, it is a good indicator that DNA polymerase activity is inhibited and replication does not take place[382].

5.3.3 A potential cause of superinfection

In the previous studies it has been suggested that superinfection exclusion happens due to primary phage expressing proteins that inhibit second phage entry into the cell[373]. This was studied with amber mutants, where the addition of a secondary mutant past 7 minutes of primary mutant addition, productive infections were not affected by the secondary mutant addion[373]. However, the data collected in this chapter, surpasses the 7 minutes time point for the superinfection. The addition of a secondary phage 2 hours after the primary phage showed phage infection recovery. It is possible that probability of P (2) =0.05 described in section 5.3.1 is significantly larger given that co-infection does not depend on two particles entering at the same as evidenced by the delay experiments. Therefore, co-infection probability increases with time.

In both T7 and chimeric phage co-dependency cases, chimeric phage had the advantage of having *trxA* in its genome, whereas the wild type had either the correct tail fiber gene or DNA polymerase. In both instances the wild type phage is dependent on the chimeric phage to provide the means of replication, via its *trxA,* which is required for DNA polymerase processivity[383]. For the co-dependency between the T7 and T7-*g5::trxA*, both phages depended on each other for replication as the modified phage had its *g5* replaced with *trxA,* and T7 had the *g5*. When comparing ratio experiments between wild type phages with either *g17* and *g5* mutants, the latter reaches an approximately 10-fold higher PFU/ml when the same amount of phage was added. This could potentially be due to the fact that in *g5* mutant phage *trxA* is placed earlier in the genome (in *g5* rather than *g17* region) that could potentially account for an earlier availability of the TrxA in the system resulting in more progeny being produced.

5.3.4 Proposed mechanism for the observed co-infection phenomenon

To explain how this potential co-infection takes place, chimeric phage selection using the *E. coli ΔtrxA* cells only (described in Chapter 4.2.4) will be used as an example. In this example, all of the phage is plated on *E. coli ΔtrxA* cells after homologous recombination. The resulting

phage progeny is a mixture of wild type and successfully recombined chimeric T7. In the Section 3.3.3 assays, it was determined that wild type T7 is not able to replicate in *E. coli* Δ*trxA* cells, as *trxA* is necessary for the formation of the DNA polymerase replication complex[384,383]. Therefore, wild type T7 is dependent on the *trxA* found on the chimeric genome for successful replication. In the case of the chimeric phage, we know that its has a non-functional tail fiber, requiring a wild type tail fiber from T7 to be infective. If both of the phages enter the cell the following potential progeny is generated: A) Chimeric genome and chimeric tail fibers; B) Chimeric genome and T7 tail fibers; C) T7 genome and T7 tail fibers; D) T7 genome and chimeric tail fibers (Figure 5.3.5-1). Due to experiments carried out in Chapter 4.2.4.3 we know that chimeric tail fibers are non-functional therefore progeny A and D are non-infective. This means that the burst size for samples containing phages with chimeric and WT genomes should be smaller than for those containing only phages with WT genomes. This is because phage types A – D are formed in mixed samples, resulting in approximately 50% of phages being non-infective, while only phage types B and C are formed in WT samples, resulting in 100% of infective phages. This could potentially explain why smaller plaques were observed (Figure 4.2.4-3) and why a significant reduction of phage particles was seen in the ratio experiments (Section 5.2.1). Here 10^7 PFU of total phage was added; however only approximately10^3 and 10^5 PFU were recovered, for ratios with the tail fiber and *g5* chimeric phages, respectively.

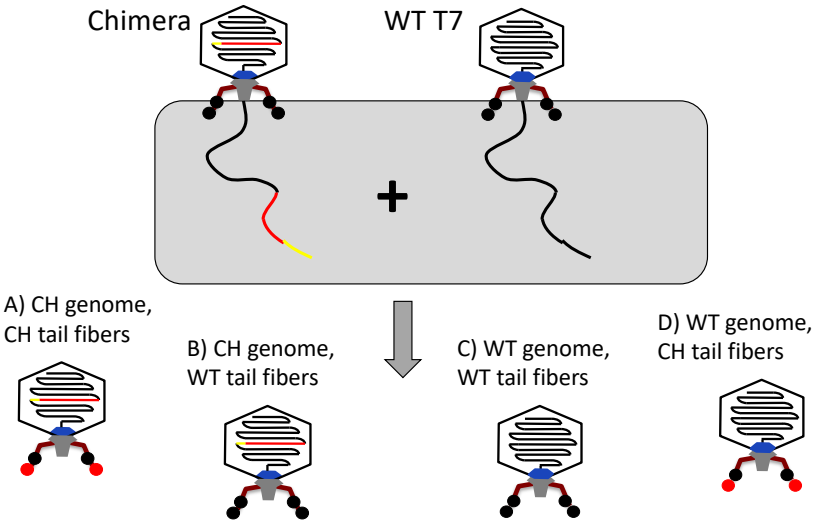


Figure 5-9. Representation of the proposed co-infection mechanism between T7 and chimeric phages. Once both phages enter the cell four types of phage progeny are generated: **(A)** chimeric genome and chimeric tail fibers; **(B)** Chimeric genome and T7 tail fibers; **(C)** T7 genome and T7 tail fibers; **(D)** T7 genome and chimeric tail fibers. **(A)** produces non-infective particles due to non-functional chimeric tail fibers. C produces phage that is unable to replicate due to the lack of *trxA* in the T7 genome.

Chapter 6 Expanding phage bacterial killing via cargo delivery

6.1 Introduction

In the previous three chapters phage engineering methods and their applications for phage therapeutics were addressed. Phages could also be used as a delivery vehicle for therapeutically actively cargo i.e. phage lysins. Lysins are enzymes naturally produced by phages at the end of their replication to cause host lysis and release of phage progeny[195]. In this study, the lysins will be referred to as lysins since they will no longer exclusively from within a cell. Phage lysins are promising alternatives to antibiotics[385,386]. They have multiple advantages over the current infection treatment e.g. elimination of bacterial colonisation of mucous membranes that often provide a start for a systematic infection[387]. The current standard method of producing a lysin is achieved by expressing it in a heterologous bacterial strain that the lysin will not lyse. If a phage was to be engineered with a lysin targeting different bacterium to that of phage host, each such phage infection would result in the phage as well as lysin production. This could expand the range of bacteria that are killed, as will have a direct target (host targeted by phage infection) and secondary targets (bacteria that are killed by lysin production). Previous work focussed on mutagenesis of phage T7, however due to DNA packaging system used by T7, only a small insertions, up to 103% of its original genome size (~1198 bp), can be incorporated into the genome[223,388–390]. Therefore, for this work a T4-like phage (Slur 96) that was recently isolated by P Sazinas which has an average nucleotide identity of >95% with T4 (NC_000866) which has the potential to allow large insertions in the genome[223,388–390] . Given the identity of this phage to T4, it will be referred to as T4 (see full T4-like sequence in the Appendix 4).

6.1.1 Lysins chosen for this study

Three lysins were chosen for study, all three have previously been purified and shown to have potential as therapeutic agents[286,391–393]. The first lysin chosen was lysH5, or H5, from the *Staphylococcus aureus* phage ΦH5[391,286]. The host for the phage ΦH5 and the target for the H5 lysin is the Gram-positive bacterium *Staphylococcus aureus*[391,286]. It has been shown that *S. aureus* can produce enterotoxins causing staphylococcal food poisoning, one of the leading causes for gastroenteritis[394]. *S. aureus* is also responsible for subclinical intramammary infections in dairy cows[395]. In a previous lysed multiple *S. aureus* bovine and clinical strains as well as *S. epidermidis* clinical strains when applied externally[396]. Up to this day very few phage lysins have been characterised, for example those of phages P68, phage K and phiWMY[272,283,397]. These lysins are made up from enzymatic (D-alanylglycyl endopeptidase,

L-muramoyl-L-alanine amidase, N-acetyl-glucosaminidase) and cell wall recognition domains[398,399] . In the previous study sequence analysis and phylogenetic position of the phage H5 lysin was used to identify its domains[286]. CHAP (cysteine, histidine-dependent amidohydrolases/ peptidases) was found as its enzymatic domain and Ami_2 (N-acetylmuramyl-L-alanine amidase) and SH3b (bacterial cell recognition) as H5 lysin host cell wall recognition domains[286].

The second lysin chosen for this study was PBC1 from phage PBC1[392,400]. The host and the target bacterium for the phage and the lysin PBC1 is *B. Cereus*[392,400]. PBC1 lysin has been shown to have a very narrow host range, showing lysis only 1 of 22 *B. cereus* strains i.e. ATCC 21768, when applied externally[400]. PBC1 was found to have a N-terminal type 3 amidase domain (PF01520), as its enzymatically active domain and a C-terminal Amidase02_C domain (PF12123), as its cell wall binding domain[400].

The third lysin chosen was CD27L isolated from CD27 phage infecting *C. difficile*, and demonstrated that it can cause cell lysis when applied externally[393]. *C. difficile* produces spores that are resistant to some disinfectants, contributing to its persistence in the environment and causes a frequently hospital-acquired infection[401,402]. CD27L has been shown to lyse the following strains: *Bacillus amyloliquefaciens*, *Bacillus cereus*, *Bacillus subtilis*, *Clostridium bifermentans*, *Clostridium sordellii* and *Listeria ivanovii* when applied externally[393]. CD27L is homologous to N-acetylmuramoyl-L-alanine amidases in the N-terminal part of the protein, while its C-terminal domain represents a novel fold that has been found in multiple lysins targeting *Clostridia bacteria*[403,404].

6.1.2 Acridines based T4 mutagenesis

It has been shown that acridines inhibit phage development at lower concentrations to those that inhibit bacterial growth[102,405]. The mechanism by which acridines inhibit phages is unknown[102,405]. It has been shown that when treated with acridine phage T4 DNA packaging has been interrupted[406]. It has also been shown that mutations in a T4 gene 52.2 , called *ac*, confers acridine resistance by a proposed mechanism where the cells permeability to acridines is reduced for the cells infected with phage[405,407]. Despite the mechanism not be being fully understood *ac* gene confers acridine resistance, the link between *ac gene* disruption and acridine has been used for T4 mutagenesis[408–410]. The disruption of *ac* will be used for selecting mutants in this chapter.

6.1.3 Aims

The main aim for this chapter is to determine if T4like phage can be engineered to kill a broader range of bacteria. The main aim was subdivided into the following sub-aims:

1) Engineer T4 phage with the three chosen lysins;
2) Characterise T4 mutants on solid and liquid medium;
3) Determine if the T4 mutant result in changed degree of virulence when compared to T4.

6.2 Methods

6.2.1 Engineering of phage T4

To engineer and select for phage T4 mutants an *ac* gene based method was adapted from a previous study[405]. The principle for gene insertion was to interrupt the *ac* gene region with a lysin gene of interest (X, C< B) as shown in Figure 6-1. To achieve this a plasmid containing homologous recombination arms (as described in section 6.2.2) was used (Section 2.3.11) followed by selection of phage that undergone a successful homologous recombination. The selection was carried out by performing modified plaque assay (Section 2.3.5) on a homologous recombination phage mixture. The modification consisted of addition of 15 ul of acriflavine (ACR) at a concentration of 1 mg/ml to each of the samples after mixing the phage, the cells and the soft agar, prior to plating on agar[405]. These samples were then plated on agar plates with ampicillin and acriflavine at 1 ng/ml. Ampicillin is required to maintain the homologous recombination vector (Figure 6-2) that facilitates a lysin insertion into phage T4, whereas acriflavine interferes with the T4 replication and hence allows for selection of engineered phage (phage that had lysin inserted and *ac* gene disrupted). The entire engineering protocol is outlined in Figure 6-2.

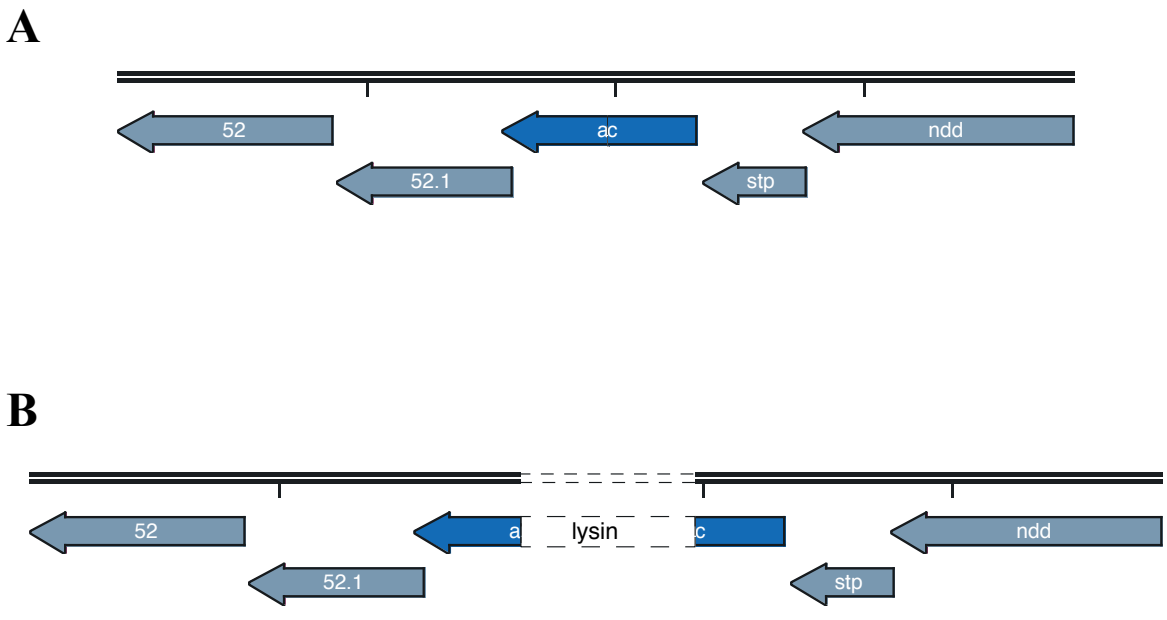


Figure 6-1. Schematic representation of lysin gene insertion into the T4 genome. **(A)** T4 genome with ac intact. **(B)** T4 genome with ac gene interrupted with a lysin insertion.

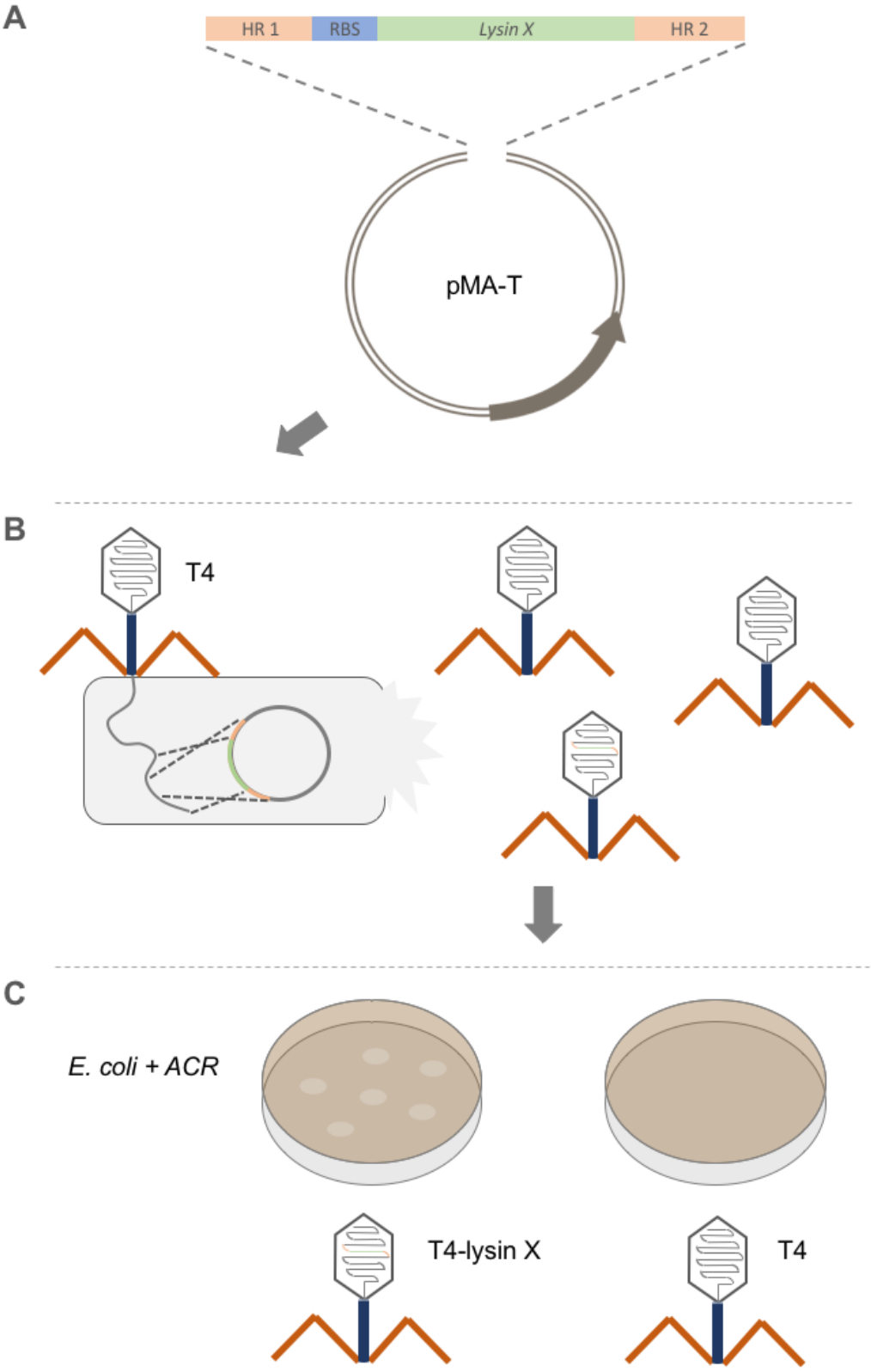


Figure 6-2. The outline of T4 engineering procedure used in this study. **(A)** design of homologous recombination plasmid (see section 6.2.2) containing homologous recombination arms, RBS B0034, a lysin gene of interest, all inserted into pMAT vector that confers resistance to ampicillin. **(B)** homologous recombination where T4 was used to infect *E. coli* MG1655/ pAG_50. **(C)** Selection for successfully recombined T4 mutant phage on 1 ng/ml acriflavine (ACR) containing agar plates.

6.2.2 Design of vector for homologous recombination

To insert a lysin of interest into the *ac* gene of phage T4, a vector was designed (Figure 6-2) containing an RBS (B0034) followed by a lysin gene of interest in between homologous recombination arms[411]. The first homologous recombination arm consisted of total of 200 bp, 122 bp before the *ac* gene and the first 78 bp of the *ac* gene. The second homologous recombination arm also consisted of total of 200 bp, the remaining 78 bp of *ac* as well as 122 bp that comes after *ac*. The design was such that the original *ac* promoter would be used for the transcription of the inserted lysin gene. The designed vector was then ordered as a synthetic gene in pMAT-1 vector from IDT (Appendix 1).

6.3 Results

6.3.1 Lysin insertion vector construction

Each of the lysin genes (*lysH5*, *lysPBC1* and *lysCD27L*) were inserted in between homologous recombination arms together with an RBS (B0034) (Sections 6.2.1 and 6.2.2) using Gibson cloning method generating homologous recombination vector. Each of the constructs were PCR verified and sequenced Table 6-1.

Table 6-1. Summary of lysin gene containing homologous recombination vectors generated in this study.

Name	Vector	HR arm 1	Lysin	HR arm 2
pAG_49	pMAT-1	200bp*	N/A	200bp**
pAG_50	pAG_49	200bp	H5	200bp
pAG_51	pAG_49	200bp	PBC1	200bp
pAG_52	pAG_49	200bp	CD27L	200bp
pAG_53	pAG_49	200bp	H5 and PBC1	200bp

* The first homologous recombination consists of total of 200 bp, 122 bp before the *ac* gene and first 78 bp of the *ac* gene.
** The second homologous recombination arm consists of total of 200 bp, the remaining 78 bp of *ac* as well as 122 bp that comes after *ac*.

6.3.2 Generation of phage T4 mutants

The vectors constructed in the previous section were used to yield T4 mutants (Table 6-2) using homologous recombination (Section 2.3.11) followed by mutant selection described in section 6.2.1.

Table 6-2. Summary of T4 mutants made for this study.

Vector Name	Lysin	Resultant Phage
pAG_50	H5	T4-*ac::lysH5*
pAG_51	PBC1	T4-*ac::lysPBC1*
pAG_52	CD27L	T4-*ac::lysCD27L*
pAG_53	H5 and PBC1	T4-*ac::lysH5lysPBC1*

6.3.3 Determination of lysin gene stability in T4

In order to determine if the lysin genes were able to remain in T4 genome over multiple generations serial transfer experiment 'T4 mutant stability assay' (Section 2.3.10) was performed over the period of 24 hours for each of the mutants. To verify the presence of the gene encoding for the lysin was maintained in the phage population PCR, on the T4-*ac::lysH5* population sample was performed after 4, 8, 16 and 24 hours (Figure 6-3 only 16 and 24-hour samples are shown). This confirms that the lysin gene is stable to remain in population for 96 (approximately 4 generations per hour) generations.

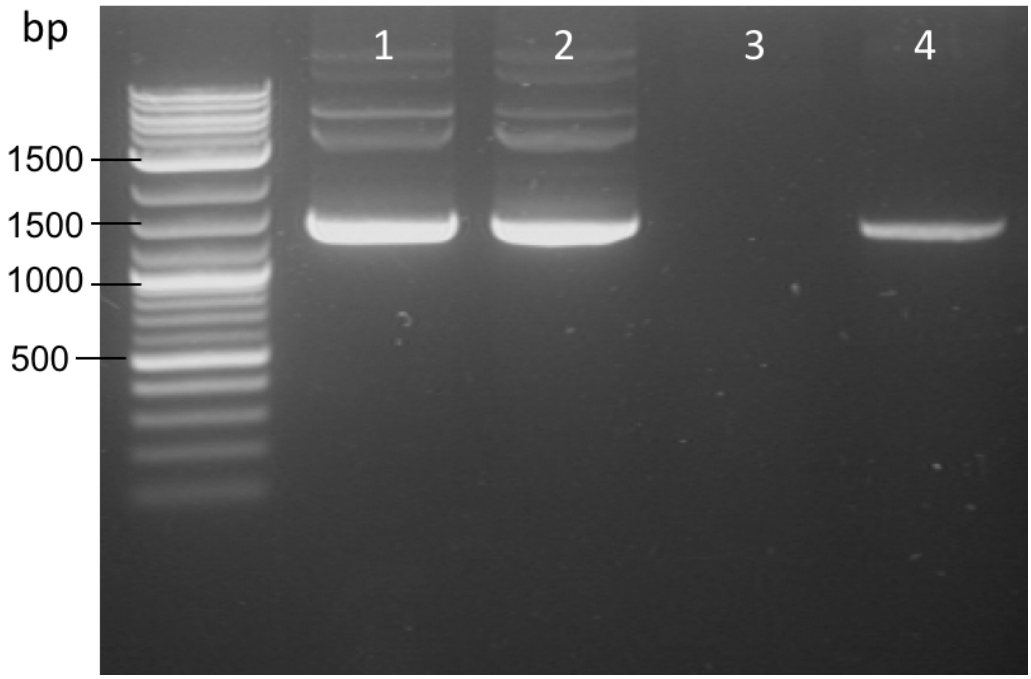


Figure 6-3. H5 lysin verification after phage mutant stability assay. Lane 1, shows H5 (~1.5 kb) after 16 hours of the assay. Lane 2, shows H5 (~1.5 kb) after 24 hours of the assay. The reference ladder is1 kb Plus DNA ladder. Lane 3 contains the negative control (no DNA template). Lane 4 contains the positive control (pAG_50).

6.3.4 Spot assays to detect lysin activity

In order to characterise each of the T4 mutants (H5, PBC1 and CD27L) the lysate from *E. coli MG1655* infected with T4 lysin mutants (Section 2.3.9) were spotted (Section 2.3.8) on a range of strains that the respective lysins maybe active against (Table 6-3). As a negative control, wild type T4 lysate was also spotted resulting in no lysis. All of the *C. difficile* strains presented in Table 6-3 were sequenced by Martha Clokie laboratory (University of Leicester). *B. subtilis* and *B. cereus* strains were sequence verified by Emma Denham (University of Bath) and Nick Waterfield (university of Warwick) laboratories. The remaining strains in Table 6-3 are clinical isolates that were not genome verified.

Phage T4-*ac::lysH5* was tested on *S. aureus* clinical isolates some of which, after streaking out on *Chromoagar Orientation,* turned out to be *Enterococcus* or *Klebsiella Enterobacter/Citrobacte* instead of *S. aureus* (Table 6-3). Phage T4-*ac::lysH5* lysed 15 out of 17 *S. aureus* and 4 out of 4 *Enterococcus* strains tested, and 2 out of 2 *Klebsiella, Enterobacter, Citrobacte* strains tested (Table 6-3). An example of lysis is showed in Figure 6-4. PBCD1 was tested on two strains *B. subtilis* and *B. cereus*, and showed clear lysis for *B. subtilis*, however, it generated some lysis for *B. cereus* (Table 6-3). CD27L was tested on 20 bacterial strains unfortunately generating no lysis.

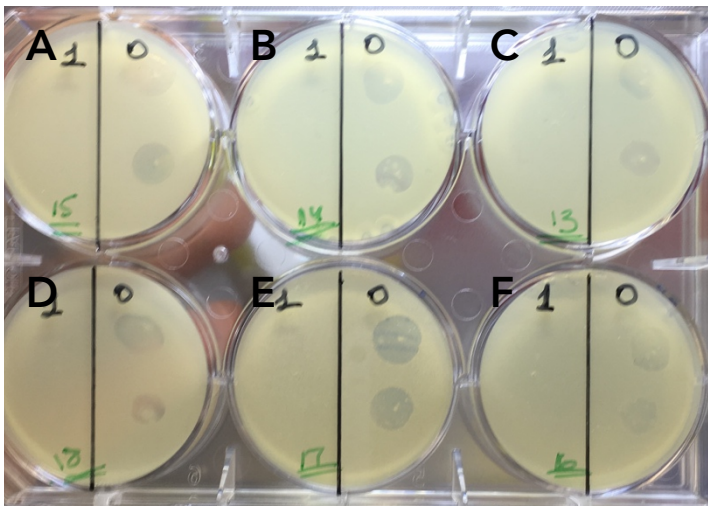


Figure 6-4. T4-*ac::lysH5* and T4 lysate spot assay against **(A)** *Klebsiella, Enterobacter, Citrobacte*, **(B)** *S. aureus*, **(C)** *S. aureus*, **(D)** *S. aureus*, **(E & F)** *Enterococcus.* Samples marked with 1 = T4 lysate, 0 = T4-ac::lysH5 For both phages, the lysates were generated by infecting *E. coli* MG1655.

Table 6-3. Summary of the T4-lysin output on a range of strains. All of the *C. difficile* strains presented here were sequenced by Martha Clokie laboratory (University of Leicester). *B. subtilis* and *B. cereus* strains were sequence verified by Emma Denham (University of Bath) and Nick Waterfield (university of Warwick) laboratories. The remaining strains are clinical isolates that were not genome verified.

Strain name /Clinical isolate no	Phage tested	Outcome (lysis vs no lysis)
Enterococcus sp1/3047211	*T4-ac::lysH5*	lysis
Enterococcus sp2/31608-21	*T4-ac::lysH5*	lysis
S. aureus sp1/33077-11	*T4-ac::lysH5*	lysis
S. aureus sp2/3047221	*T4-ac::lysH5*	lysis
S. aureus sp3/33077-31	*T4-ac::lysH5*	lysis
S. aureus sp4/32421-12	*T4-ac::lysH5*	lysis
S. aureus sp5/30472-31	*T4-ac::lysH5*	lysis
S. aureus sp6/31554-1	*T4-ac::lysH5*	lysis
S. aureus sp7/33503-12	*T4-ac::lysH5*	lysis
Klebsiella, Enterobacter, Citrobacter sp1*/31608-31	*T4-ac::lysH5*	lysis
S. aureus sp8/32827-31	*T4-ac::lysH5*	lysis
S. aureus sp9/32827-11	*T4-ac::lysH5*	lysis
Klebsiella, Enterobacter, Citrobacter sp2*/3393011	*T4-ac::lysH5*	lysis
S. aureus sp10/3047211	*T4-ac::lysH5*	lysis
S. aureus sp11/31608-11	*T4-ac::lysH5*	lysis
Enterococcus sp3/32957-11	*T4-ac::lysH5*	lysis
Enterococcus sp4/32957-12	*T4-ac::lysH5*	lysis
S. aureus sp12/31530-12	*T4-ac::lysH5*	lysis
S. aureus Newman	*T4-ac::lysH5*	no lysis
S. aureus sp13	*T4-ac::lysH5*	no lysis
S. aureus sp14	*T4-ac::lysH5*	lysis
S. aureus sp15	*T4-ac::lysH5*	lysis
S. aureus sp16	*T4-ac::lysH5*	lysis
S. aureus sp17	*T4-ac::lysH5*	lysis
Bacilus subtilis ATCC 6051	*T4-ac::lysPBC1*	lysis
Bacilus cereus G9241	*T4-ac::lysPBC1*	lysis
C. difficile K20	*T4-ac::lysCD27L*	no lysis
C. difficile H5C	*T4-ac::lysCD27L*	no lysis
C. difficile H431024	*T4-ac::lysCD27L*	no lysis
C. difficile M68	*T4-ac::lysCD27L*	no lysis
C. difficile AML	*T4-ac::lysCD27L*	no lysis
C. difficile A14	*T4-ac::lysCD27L*	no lysis
C. difficile R018	*T4-ac::lysCD27L*	no lysis
C. difficile K10	*T4-ac::lysCD27L*	no lysis
C. difficile CD89	*T4-ac::lysCD27L*	no lysis
C. difficile AJV	*T4-ac::lysCD27L*	no lysis
C. difficile R20291	*T4-ac::lysCD27L*	no lysis
C. difficile K6	*T4-ac::lysCD27L*	no lysis
C. difficile B19	*T4-ac::lysCD27L*	no lysis
C. difficile T6	*T4-ac::lysCD27L*	no lysis
C. difficile ATM	*T4-ac::lysCD27L*	no lysis
C. difficile 1342	*T4-ac::lysCD27L*	no lysis
C. difficile TL178	*T4-ac::lysCD27L*	no lysis
C. difficile AIP	*T4-ac::lysCD27L*	no lysis
C. difficile CD84	*T4-ac::lysCD27L*	no lysis
C. difficile AIJ	*T4-ac::lysCD27L*	no lysis

*****After streaking out the clinical isolates, *Chromoagar Orientation* generated metallic blue colour which is representative of one of the strains: *Klebsiella, Enterobacter, Citrobacte***.**

6.3.5 Co-culturing *E. coli* with *B. subtilis* and *S. aureus*

The results of spot assays suggested the production of an active lysin during T4 replication that could kill a Gram-positive host. To verify if the lysins inserted into the T4 genome are still able to lyse their target bacterium in the liquid culture, both *E. coli* and *S. aureus* or *B. subtilis* would have to be co-cultured together and phage added. Prior to doing this, it was necessary to test that *E. coli* could be cultured with *S. aureus* or *B. subtilis*. *C. difficile* was not included in these experiments due to anaerobic growth conditions required and the lack of any lysis based on spot assays (Table 6-3).

Both *E. coli* and *B. subtilis*, and *E. coli* and *S. aureus*, were co-cultured and allowed to grow until an OD_{600nm} 0.4 was reached (Section 2.2.5). In the case of *E. coli* and *B. subtilis*, *E. coli* MG1655/pSB6K1-RFP was used to distinguish between *E. coli* and *B. subtilis* colonies. *E. coli* colonies have a red morphology. In the case of *E. coli* and *S. aureus,* both strains were plated on *Chromoagar Orientation* selective medium (Section 2.1.2), that allows differentiation between *E. col*i, as it forms pink colonies, and *S. aureus*, as it forms white colonies.

When *E. coli* MG1655 was cultured separately and in co-culture with *B. subtilis,* it yielded approximately 6 x 10^7 cfu/ml in both conditions (Figure 6-5). Reciprocally *B. subtilis* when grown separately and in co-culture with *E. coli* MG1655 yielded 2.5 x10^7 cfu/ml (Figure 6-5). Therefore, co-culturing both strains does not interfere with their growth. A similar result was found when *E. coli* and *S. aureus* were co-cultured or cultured separately, with 6 x10^7 and 1 x10^8 cfu/ml found in all conditions for both *E. coli* and *S. aureus* respectively (Figure 6-5). Thus, again confirming the strains can be co-cultured together and do no effect each other's growth.

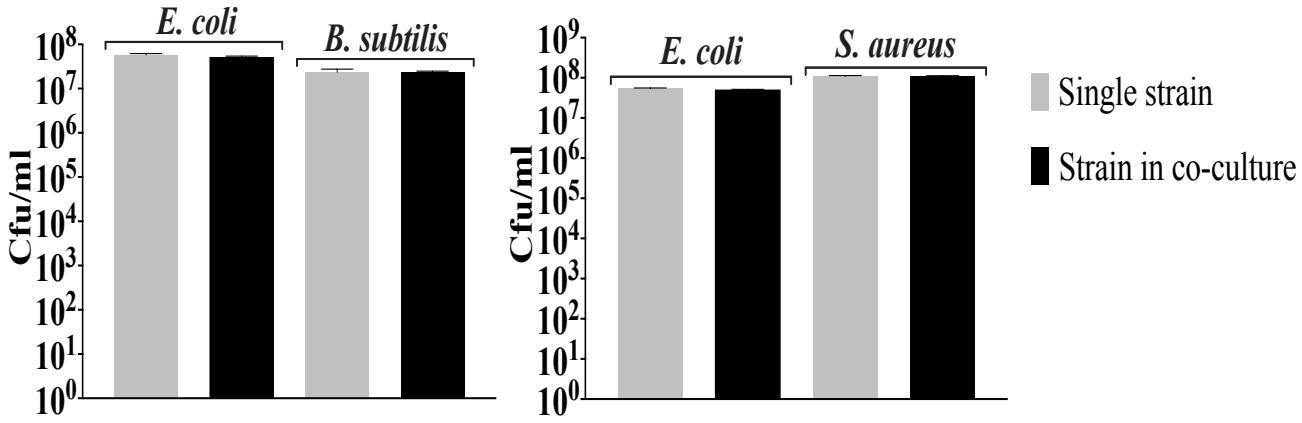


Figure 6-5. The colony enumeration assays for the co-culturing experiments. **(A)** *E. coli* MG1655 and *B. subtilis* strains were plated after growing them separately at OD_{600} 0.4. The colony counts for either *E. coli* or *B. subtilis*, were labelled as '*E. coli*' or '*B. subtilis*' respectively and represented in grey column 'single strain'. The strains were also plated after growing them both in co-culture at OD_{600} 0.4. The colony counts for either *E. coli* or *B. subtilis* were grown in co-culture were labelled as '*E. coli*' or '*B. subtilis*' respectively and represented in black column 'strain in co-culture'. **(B)** *E. coli* MG1655 and *S. aureus* strains were plated after growing them separately at OD_{600} 0.4. The colony counts for either *E. coli* or *S. aureus*, were labelled as '*E. coli'* or '*S. aureus*' respectively and represented in grey column 'single strain'. The strains were also plated after growing them both in co-culture at OD_{600} 0.4. The colony counts for either *E. coli* or *S. aureus* were grown in co-culture were labelled as '*E. coli*' or '*S. aureus*' respectively and represented in black column 'strain in co-culture'.

6.3.6 Assessing efficacy of T4 mutants in liquid medium

In order to characterise lysins in liquid medium the following experiment was set-up using co-culturing experiment described with two additional treatments. Either T4 or T4 mutants were added to the co-cultures when they reached an of OD_{600nm} of 0.4 (Section 2.2.5). *T4-ac::lysPBC1* was added to *E. coli* MG1655/*B. subtilis* co-culture and T4-*ac::lysH5* was added to the *E. coli/ S. aureus* co-culture, phage T4 was used as a control in both systems. Samples from co-cultures were collected after 30 minutes and 1 hour followed by colony enumeration (Section 2.3.5).

For *E. coli* and *B. subtilis* co-culture, there was no change in the number of colonies for *B. subtilis* when untreated, or the addition of T4 or T4-*ac::lysPBC1* (Figure 6-6 an example of the assay output colonies Figure 6-7). For *B. subtilis* the control and two treatments approximately 10^7 and 10^8 cfu/ml were enumerated for 30 min and 1-hour time points respectively. After 30 min the number of *E. coli* was reduced 10-fold for both T4 and *T4-ac::lysPBC1* treatments when compared with the no phage control (Figure 6-6). After 1-hour, *E. coli* was reduced >100-fold for both T4 and T4-*ac::lysPBC1* treatments when compared with the no phage added control. The same pattern was observed for *E. coli* and *S. aureus* co-cultures, with reduction in *E. coli* numbers when phage T4 or T4-*ac::lysH5* were added, without a reduction in *S. aureus.*

CFU assays presented in Figures 6-5 and 6-6 indicate higher *E. coli* numbers when compared to *B. subtilis*, and greater *S. aureus* numbers when compared to *E. coli*. To evaluate if co-culturing of *E. coli* and either *B. subtilis* or *S. aureus* would maintain lysin production and its associated lysis, the growth rate needs to be determined for each of the strains in question. If the growth rate of either of lysin target strains is much greater than that of *E. coli* and lysin production, the engineered phage will not be able to completely eliminate the target bacteria.

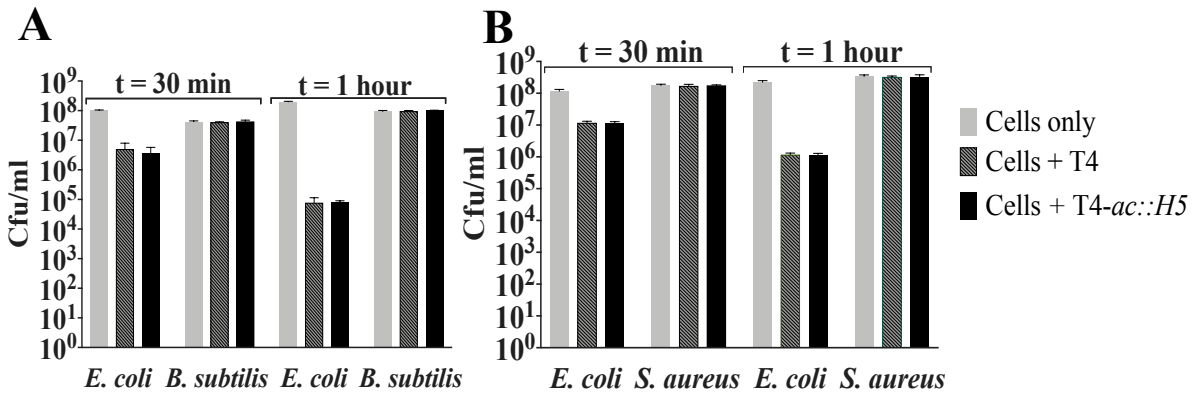


Figure 6-6. Colony enumeration assays after the co-culturing experiments. **(A)** *B. subtilis* and *E. coli* were co-cultured and once reached OD 0.4 were subjected to two treatments (addition of either T4 or T4-*ac::PBC1* phage at an MOI of 1) and a no treatment control (cells only). Bacterial numbers were determined after 30 min and 1 hour. **(B)** *S. aureus* and *E. coli* were co-cultured and once they had reached OD 0.4 were subjected to two treatments (addition of either phage T4 or T4-*ac::H5* at an MOI of 1) and a no treatment control (cells only). The samples were enumerated using cfu assays after two time points Bacterial numbers were determined after 30 min and 1 hour (n=3 for each condition).

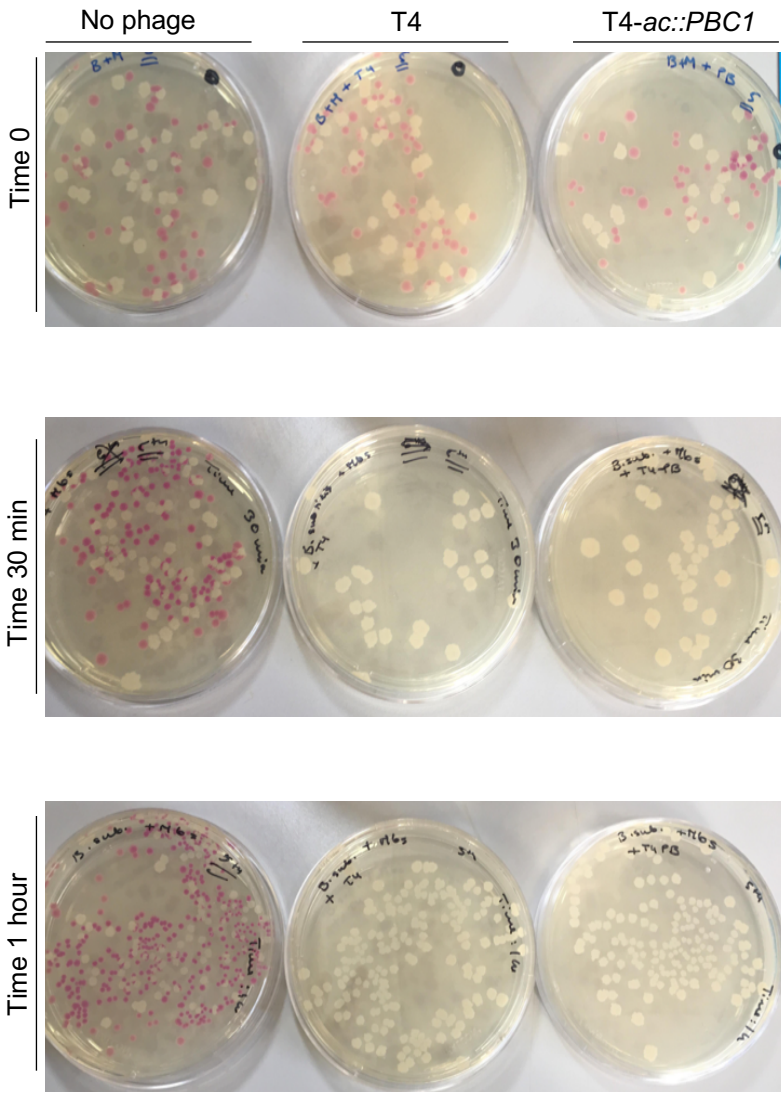


Figure 6-7. Summary of post *E. coli* MG1655/*B. subtilis* co-culture CFU assays. Each of the rows represent different time points, from top to bottom: time 0 (each of the strains were grown up to OD 0.4, mixed in equal ratios and 100ul of the mixture was plated), time 30 min (the mixture from time 0 was grown for 30 min), time 1 hour (the mixture from time 0 was grown for 1 hour). Each of the columns represent different treatment conditions, from left to right: no phage added, T4 added at MOI=0.1, T4-*ac::PBC1* added at MOI=0.1.

6.3.7 T4 mutant virulence assays

To determine if engineering T4 with additional payload conveys any burden for its infectivity, virulence assays were performed (Section 2.3.15) for T4 its mutants as well as T7 phages Table 6-4. All of the T4 mutants' virulence indices generated were the same as the T4 phage.

Table 6-4. Summary of virulence indices.

Phage	Virulence index
T4	~0.6
T4-*ac::lysH5*	~0.6
T4-*ac::lysPBC1*	~0.6
T4-*ac::lysCD27L*	~0.6
T7	~0.9

6.4 Discussion

6.4.1 Recombinant lysin expression

In multiple studies lysins have been over expressed in a heterologous bacterial system followed by purification of the lysin. In one instance, lysins of phages targeting *Listeria monocytogenes,* Ply118 and Ply511, were introduced into a vector in *Lactococcus lactis*[412]. Upstream of the lysin gene an purification tag sequence was inserted to allow the use of metal chelate affinity chromatography for the later lysin purification steps[413]. In another study, LysB4, an lysin from the *Bacillus cereus*-infecting bacteriophage B4 was expressed in *E. coli* with an N-terminal His-tag followed by purification using affinity chromatography[414]. His tag insertion followed by affinity chromatography in multiple other cases[415,416]. The ability to express lysins using phage expression system described in this chapter, would remove the need for a purification step, as large numbers of phage would be released during an infection cycle. However, this may result in increased lysin instability due to uncontrolled and possible suboptimal conditions during phage lysis. Each of the lysins used in this chapter were previously verified to be stable at 37°C pH 7, however what the conditions are following cells lysis are is unknown. It has been hypothesised that respiring cell produces a proton gradient just outside its cytoplasmic membrane[417]. Therefore, the pH may vary during lysis affecting lysin activity. Furthermore, it has been observed that lysin activity and autolysis are inhibited by growth at low pH in multiple bacterial species including pneumococci, staphylococci, streptococci, and *B. subtilis*[418].

6.4.2 The host range of T4 mutants' lysate on solid medium

It has been showed that a phage lysins have a broader lytic spectrum than the phages they are found in[419–421]. This was true for phage *ΦH5* and its encoded LysH5 lysin[396]. In the previous work, in addition to lysing the host bacterial strain, LysH5 managed to lyse multiple *S. aureus* strains despite their origin, bovine or human[396]. Furthermore, LysH5 showed activity against clinical strains *of S. epidermidis* isolated from humans. However, no lytic activity was observed against *Bacillus, Clostridium, Streptococcus, Enterococcus* and *Listeria*[396]. The results generated in this chapter confirms LysH5 activity against multiple *S. aureus* strains. However, here *Enterococcus* as well as potentially one of the *Klebsiella, Enterobacter, Citrobacter* strains were also lysed by LysH5. This implies that LysH5 is able to lyse both Gram-positive and Gram-negative bacteria.

Both Gram-negative and Gram-positive bacteria targeting lysins contain enzymatically active domain (EAD) as part of their lysin structure[201]. Since the Gram-positive bacteria do not have an outer membrane, a lysin can access the bacterial peptidoglycan when applied externally resulting in lysis[201]. In a previous study, lysin LysAB2 showed lytic activity against both Gram-negative and Gram-positive bacteria, *A. baumannii* and *S. aureus*[422]. It has been shown that some lysins, mainly the ones from Gram-positive bacteria are capable of disrupting bacterial cells by a mechanism independent to their enzymatic mechanism[423]. It has been showed that helix-forming amphipathic peptides containing basic amino acids cooperate with negatively charged membrane components such as lipopolysaccharides[424].
In the case of LysPBC1 lysin, the phage that encodes LysPC1 has a very narrow host range, infecting only 1 of 22 *B. cereus* strains (*B. cereus* ATCC 21768)[400392]. However, LysPBC1 showed lysis when applied against 22 out of 22 *B. cereus strains*, as well as 3 out of 4 *B. subtilis* strains, *B. thuringiensis*, *B. mycoides*, *B. weihenstephanensis*[400392]. The results generated in this chapter confirms LysPBC1 activity against *B. subtilis* ATCC 6051.

For the third lysin LysCD27L, the phage ΦCD27 has a reported narrow host range, infecting only four of the 30 *C. difficile* strains tested[404]. The host range of the lysin LysCD27L is significantly broader lysing all 30 strains tested, in addition to *Bacillus amyloliquefaciens*, *Clostridium bifermentans*, *Clostridium sordellii*, *Listeria ivanovii*, *Bacilus cereus* and *Bacilus subtilis*[393,404]. The results generated in this chapter were not able to confirm the lytic activity of LysCD27, as it showed no lysis when tested against 20 different *C. difficile* strains. The lack of lysin activity could be due to the lack of stability due to suboptimal conditions occurring during the lysis as discussed in 6.4.1.

Another explanation for reduced lysis of the three lysins in question could be due to insufficient gene expression levels. As previously mentioned, *ac* gene is a T4 gene found late in the genome (position 52.2), therefore lysin expression would only be initiated at a later stage of T4 infection. The strength of *ac* promoter has also not been evaluated in the literature, therefore prediction of successful transcription events is not possible. Late and potential weak transcription would delay the translation and may have a significant impact for the generation of the final numbers of lysin molecules.

6.4.3 The T4 mutants' bacterial killing in liquid medium

The classical way of assaying for lysin activity in liquid medium is turbidity reduction assay[270]. In this assay the decrease in light scattering due to live cells is measured[270]. Since the system, described in this chapter, requires a co-culture between *E. coli* and a lysin target the turbidity assay could not be used. As it would be impossible to determine which strain was being cleared solely based on OD reading. Therefore, co-culture followed by addition of phage T4-*ac::lysin*, incubation and subsequent enumeration of colony forming units was carried out. No lysis was detected in liquid conditions when T4 mutants' lysate was tested against different hosts. This could be due to the fact that in the previous studies, the lysins of interest were first purified and then used for the lysis assays, allowing higher concentrations of the active lysin. In this chapter, all of the lysins were produced using T4 infection a means of generating the lysin, which could potentially not reach the levels required for lysis generation. When comparing to the solid medium experiments, the lysis in liquid medium could potentially be affected by the change in kinetics and finding the host in the co-culture may require higher concentration of the lysin. In addition, it is important to note the possibility that lysins may be one-use enzymes[267]. As previously shown, a listeria phage lysin had a binding affinity to its substrate similar to that of IgG molecule to its substrate, meaning that several molecules may be needed to attack a local region to sufficiently weaken the cell wall to result in lysis[267].

Chapter 7 Engineering phage to be tissue specific

7.1 Introduction

Phages have multiple advantages over antibiotics and hold a lot of promise for an alternative therapy[133]. One of the major factors for effective phage therapy treatment of infection is sufficient exposure to the phage at the site of bacterial infection (a tissue or an organ)[150,181]. A limitation of phage therapy is insufficient localisation of phage to the tissue of interest (targeted tissue)[181]. Engineering phages to improve their tissue-specificity is a promising avenue that is set out to explore in details in this chapter. Furthermore, some clinically significant pathogens are able to invade and localize inside host cells which protects them from the host immune system as well as antibacterials[425–428]. The ineffective delivery of pathogen targeting drugs leads to only partial clearance of the pathogen resulting in persistent infections[429–431]. If tissue-specific phages were engineered this would potentially enhance their endocytosis in mammalian cells allowing greater killing of the intracellular bacteria.

7.1.1 Homing peptides and their potential to aid tissue specific phage targeting

Previous studies have shown that vascular endothelium cells are heterogenous and express tissue specific-markers[432]. Specifically, the vascular endothelium of healthy or diseased organs possess organ specific molecular markers or receptors, and ligands binding to them are referred to as ‘‘homing peptides (HP)’’[433]. *In vivo* phage display technology has been used for tissue-specific marker recognition, to identify extensive molecular differences in vasculature[434]. Multiple peptides targeting normal and tumour blood vessels, or tumour lymphatic vessels have been isolated[435–440]. In addition, homing peptides have been identified that target the vasculature of various organs, including the lung, heart, prostate, skeletal and cardiac muscle, and adipose tissue[435,436,441–443]. Homing peptides have been exploited for targeting drug molecules, liposomes, and inorganic nanoparticle to tissues[433]. Their use has not only increased the targeting specificity and efficacy of drug delivery but has also reduced drug associated side effects, presumably by reducing unspecific targeting of healthy tissues and organs[444,445]. Arming phages with homing peptides, a technique previously applied to other nano-particles, would potentially allow greater phage concentrations and hence greater efficacy at the site of infection[446].

7.1.2 Lungs as a target for tissue-specific phages

The human lung filter through approximately 10,000 litres of air making it susceptible to airborne pathogens on the daily basis[182]. This air can contain approximately 100,000 bacteria per litre, some of which will get into the lung and beyond[447]. A wide range of pulmonary infectious agents have been identified over the years including *Mycobacterium tuberculosis*, *Streptococcus pneumoniae Haemophilus influenzae* and many others[448–450]. Many pathogens overcame antibiotics that resulted in various degrees of resistance, compromising treatment of the pulmonary infections[183]. Having considered all of the above, the lung tissue has been chosen as a target tissue for the homing peptide engineered phage targeting.

7.1.3 T7 Capsid engineering

The T7 capsid has previously been modified for phage display revealing protein – protein interactions and protein – targeting by chemical compounds[451–454]. The T7 capsid consists of 10A (major) and 10B (minor) proteins[224,455]. They share the same start codon, but the 10B protein is generated by a shift in translational reading frame ahead of the normal termination codon, which adds 53 amino acids to the C-terminal end of the 10A protein[224]. This frameshift occurs rarely, which explains why 10B is transcribed at a lower frequency and makes up approximately ∼5–10% capsid protein translated[455,456]. Both the major (10A) and minor (10B) proteins are combined to make up the T7 phage capsid particle[455,457].

The 10B minor capsid protein has been mainly used for modifications as it is produced at lower frequency and hence is less likely to affect the capsid integrity if a ligand of interest is attached[458]. Previously the T7 capsid has been modified to incorporate a fluorescent protein enabling a direct visualisation of the interaction between displayed proteins and their binding particles[459]. In all studies so far, the modifications to the capsid protein have been done using *in trans* complementation, where modified capsid protein of interest is expressed in *E. coli* followed by T7 phage infection allowing the modified protein integration into the phage capsid[459]. This method allows an *in trans* complementation of the desired altered capsid protein resulting in some protein being complemented and some retaining the wild type features. If the phage genome was to be modified all of the capsid proteins would be retain the desired modification.

7.1.4 Aims

The main aim of the chapter is to determine whether T7 engineered with homing peptides could allow increased adsorption/absorption to the mammalian cells. This aim was divided into the following sub-aims:

1) To select homing peptides and design their insertions in T7 minor and major capsid proteins;
2) To engineer/insert homing peptides in the minor and major capsid proteins of phage T7;
3) To verify if HP insertion has any effect on capsid structure;
4) To compare T4 vs T7 adsorption to lung endothelial tissue cells;
5) To compare wild type and homing peptides – engineered T7 phages' adsorption/absorption to lung endothelial tissue cells.

7.2 Methods

7.2.1 Determination of homing peptides and the mammalian cell line to be used

The target organ for this study was chosen to be the lungs. Lung epithelial tissue targeting HPs have been previously extensively identified (Table 7-1)[446,460–464]. Therefore, adenocarcinomic human alveolar basal epithelial cells, specifically A549 cells, were chosen to mimic lung epithelial tissue for the *in vitro* experiments detailed in this chapter[465–467]. It has to be noted that A549 is a cancerous lung epithelial cell line, and hence the expression profile and cellular membrane surface differ from that of a healthy cell line. However, A549 cell line has not been reported to upregulate or downregulate the receptor membrane dipeptidase that is the receptor for GFE-1/2 homing peptides, suggesting that its levels are the same to that of a healthy tissue[468]. The receptor for MTDH domain is a putative one therefore its regulation in cancerous vs healthy epithelial cell line was not identified.

The small reported DNA sequence sizes of homing peptides do not surpass the 1 kb genomic insertion limit for the T7 phage, therefore their insertions should be tolerated by T7 phage[446,460–464].

Table 7-1. Summary of homing peptides chosen to be inserted in T7 capsid proteins.

HP No.	Name	HP Sequence	Receptor
1	GFE-1	CGFECVRQCPERC	membrane dipeptidase[460,462]
2	GFE-2	CGFELETC	membrane dipeptidase[460,462]
3	MTDH domain	GLNGLSSADPSSDWNAPAEEWGN WVDEDRASLLKSQEPISNDQKVSD DDKEKGEGALPTGKSK	lung endothelium putative receptor[421]

7.2.2 Marker based selection for T7 capsid engineering

Previously it was demonstrated that *trxA* marker-based selection is the most efficient method of engineering bacteriophage T7 (Section 3.3.4). Therefore, to insert a homing peptide into the capsid proteins, homologous recombination was carried out followed by marker-based selection. For this homologous recombination a vector was designed that contains ~100 bp

homologous recombination arms flanking the homing peptide, an RBS site and the *trxA* gene (Figure 7-1).

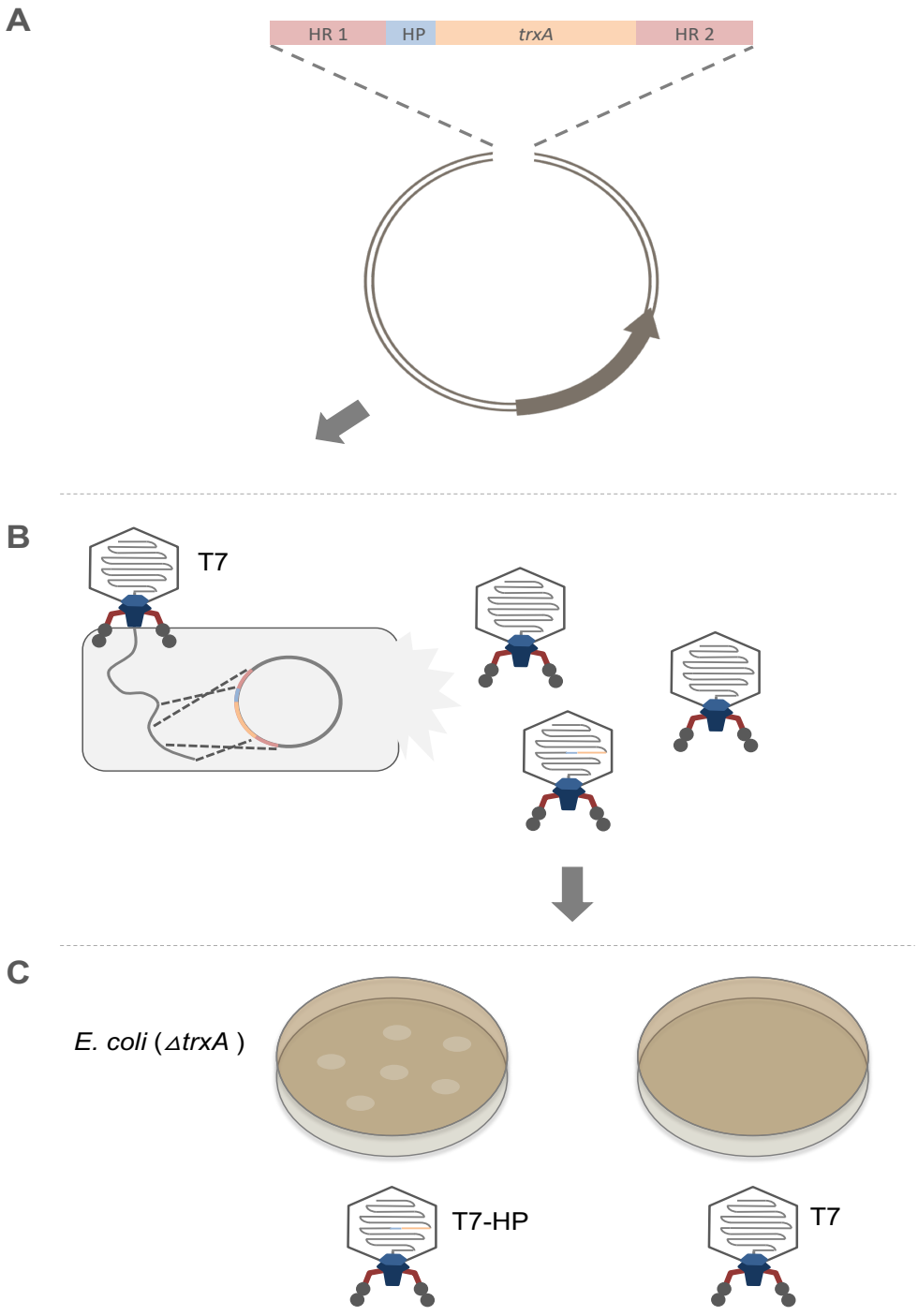


Figure 7-1. Engineering and selection of T7 containing homing peptides (T7-HP). **(A)** The homologous recombination (HR) vector was constructed by inserting a sequence encoding the homing peptide DNA and *trxA*, flanked by homologous recombination arms, into the backbone pSMART. **(B)** Transformation of *E. coli* BW15113 with the HR vector, followed by the addition of wild type T7 phages. **(C)** Phage selection on *E. coli* Δ*trxA*.

7.2.3 Design of capsid 10A and 10B fusion proteins for homing peptide display

The C-terminus of 10B has been used for insertions in multiple studies[451–454,458]. However, the C-terminus of 10A has not been used in a similar manner, potentially due to its higher translation frequency and the inherent risk of compromising the capsid integrity. For all homing peptides (listed in section 7.2.1) it was decided to exploit the C-terminus end of not only 10B but also 10A capsid proteins. To this end, the translational frame shift required for 10B is represented in Figure 7.2.3-1. 10Bi represents the part of the 10B gene translated using frame 0, whereas 10Bii is translated using frame -1 (Figure 7-2). Having this in mind, it was possible to append the homing peptides to both 10B (Figure 7-2) and 10A capsid proteins (Figure 7-2) using homologous recombination regions 1 and 2 (HR1/2) Table 7-2. By inserting a HP at the end of 10B C terminus, wild type functional 10A is made at the same time as 10B now containing the HP of interest. By inserting a HP at the end of 10A C terminus, 53 aa of wild type 10B are removed generating 10B that may or may not be functional. At the same time 10A is made with a HP of interest. Each insertion for a homing peptide was followed by *trxA* insertion as described in the methods section and section 7.2.2.

Table 7-2. Homologous recombination regions 1 and 2 (HR1/2) for 10A and 10B homing peptide insertions.

	HR1	**HR2**
10A	cgtgctaacttccaagcggaccagattatc gctaagtacgcaatgggccacggtggtcttc gcccagaagctgctggtgcagtggttttcaaagtggag	agcataaccccttggggcctctaaacgggtc ttgaggggttttttgctgaaaggaggaactatat gcgctcatacgatatgaacgttgagactgccg
10B	acagaagaaaccttaacgccagcacaggag gccgcacgcacccgcgctgctaacaaagcccg aaaggaagctgagttggctgctgccaccgctgagcaa	ctagcataaccccttggggcctctaaacgggtc ttgaggggttttttgctgaaaggaggaactatatg cgctcatacgatatgaacgttgagactgccg

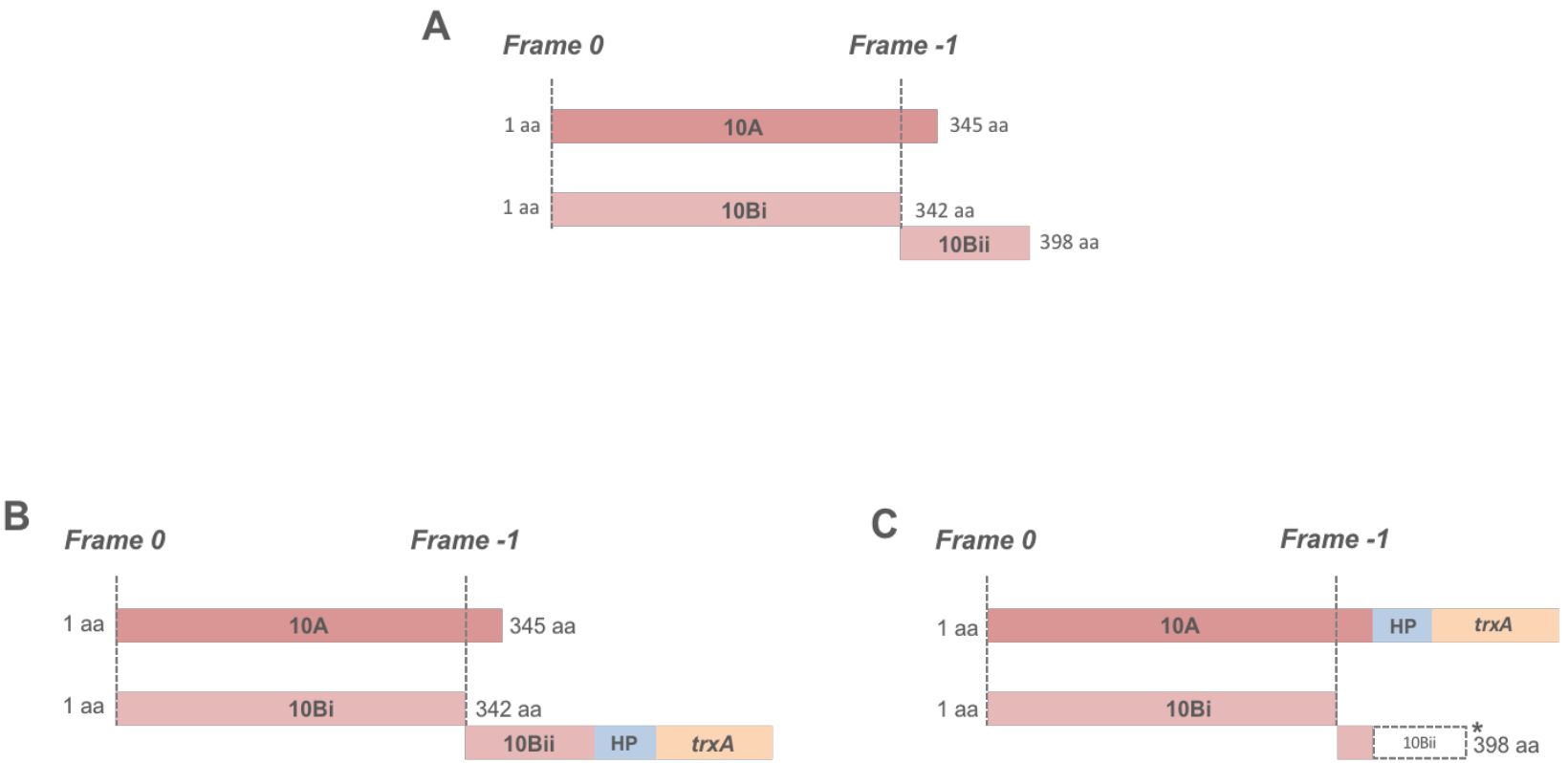


Figure 7-2. HP insertion strategies. **(A)** Representation of 10A and 10B translational shift. When 10A is translated (Frame 0), only a hypothetical portion of 10B is translated (10Bi), however 10B itself is not produced. When 10A is being translated and before its stop codon, it undergoes a translational shift (Frame -1), the full 10B is translated 10% of the time. **(B)** Representation of HP insertion after 10B. This means that only 10B protein is going to have HP attached to it and 10A will not be altered. **(C)** Representation of HP insertion after 10A. This means that 10A will have HP attached to it. 10B will not be formed at all as the insertion of HP disrupts sequence after 10A, removing part of 10B. This means that no 10B will be formed. (*) represents the part of 10B disrupted by HP insertion.

7.3 Results

7.3.1 T7 engineered with homing peptides

The vectors designed in the methods section 7.2.3 were transformed into *E. coli* Δ*trxA* cells and then used for homologous recombination (Section 2.3.11). This was followed by marker-based selection (section 7.2.2) and phage plaque purification (Chapter 2.3.13). This resulted in the generation of the mutant T7 phages that were PCR verified and sequenced (see Appendix 4) Table 7-3.

Table 7-3. Summary of T7 mutants generated in this study.

Plasmid used*	10A/B	HP	RBS/trxA	Name
pAG_54	A	GFE-1	∨	**T7-HP1**
pAG_55	A	GFE-2	∨	**T7-HP2**
pAG_56	A	MTDH	∨	**T7-HP3**
pAG_57	B	GFE-1	∨	**T7-HP4**
pAG_58	B	GFE-2	∨	**T7-HP5**
pAG_59	B	MTDH	∨	**T7-HP6**

*-vectors used here were synthesised by IDT, for the vector map see Appendix 1, for the insert sequences see Appendix 4.

7.3.2 10A modifications do not affect the capsid structure

To the best of my knowledge nobody has ever attempted to modify the major 10A capsid protein. Interestingly, we were able to generate T7 phages with mutated 10A. These phages were not only able to infect and complement *E. coli* ΔtrxA (Figure 7-4) but were also not displaying any structural defects when viewed under TEM (Figure 7-3) (Section 2.3.16). Indeed, phage *MTDH::10A*-T7 (Figure 7-3) looked similar in shape and size to the wild type T7 phage (Figure 7-3).

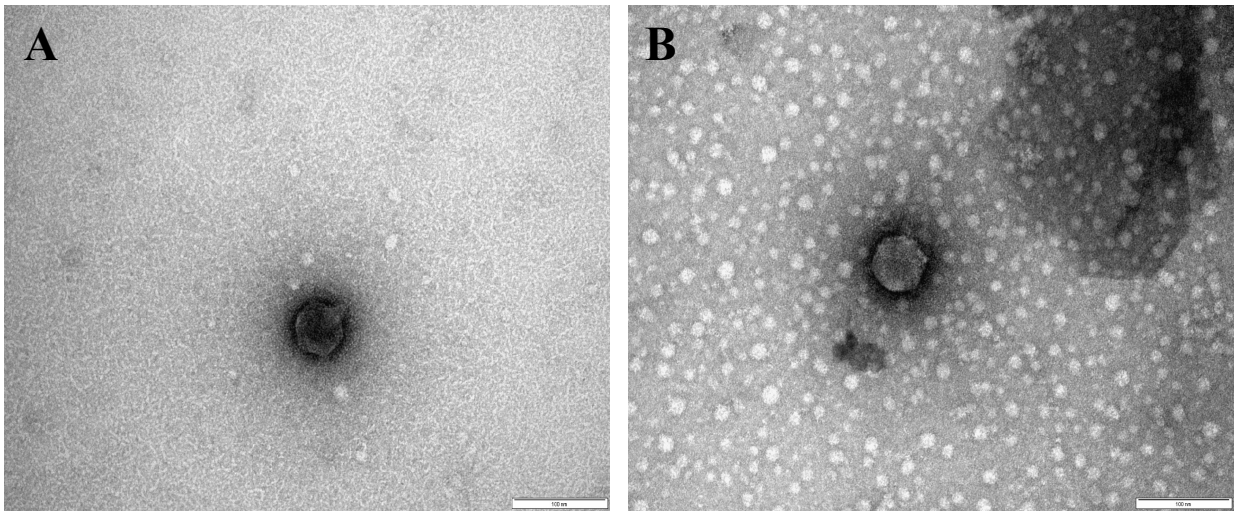


Figure 7-3. TEM images for mutant T7 and T7 phages. **(A)** T7-*10A::MTDH*. **(B)** phage T7.

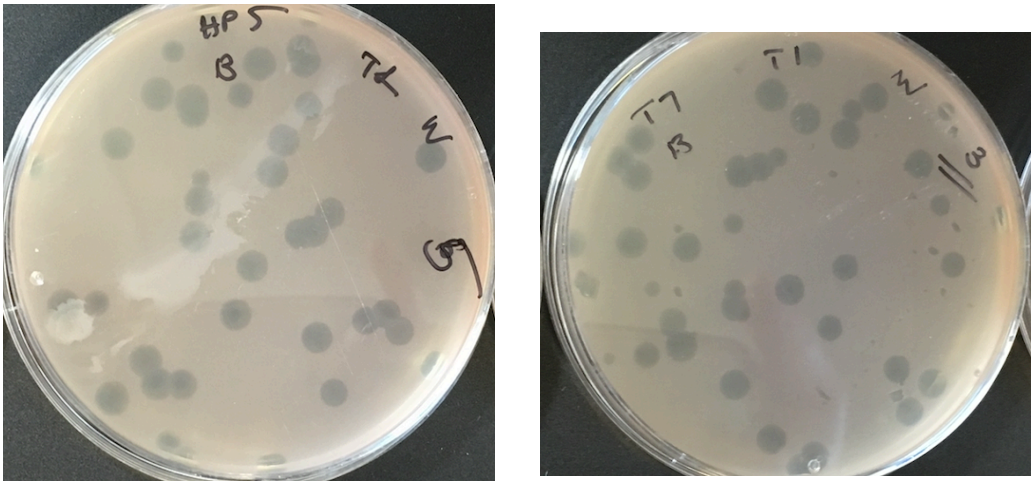


Figure 7-4. Plaque comparison between **(A)** T7-*10A::MTDH*. **(B)** phage T7.

7.3.3 Comparison of T4 and T7 adsorption to A549 cells

Prior to testing mutant T7 phages, the adsorption of wild type phage T4 and T7 to A549 cells was tested. This was done by adsorption/absorption assays (Section 2.4.3), where the wash step allows the identification of loosely bound phages whereas the lysis step allows quantification of phages either bound strongly to the cell surface or that have entered the cell. T4 and T7 were used at six different concentrations (10^3 to 10^8 PFU/ml) and two phage incubation time points (1 and 4 hours) to evaluate if higher phage amount application allows greater phage entry/adsorption to the epithelial cells. For this experiment total phage recovery was evaluated by combining the total amount of loosely bound phages (after the wash step) with the total amount of cell associated phages (lysis). In general, adding higher amounts of phage to the samples resulted in higher phage recovery after the wash and lysis steps. The highest number of total phage recovered after adding 10^8 PFU/ml, was approximately 10^4 PFU/ml for T4 and T7 phages after 1 hour of incubation (Figure 7-5). The incubation time was found to have different effects on T4 and T7 recovery. When increasing the incubation time from 1 to 4 hours, after addition of 10^8 PFU/ml of each phage, approximately 10^6 PFU/ml and 10^4 PFU/ml were recovered for T4 and T7, respectively (Figure 7-5). This assay showed that both T4, and more importantly, T7 phage can be adsorbed to and potentially be internalized into mammalian cells.

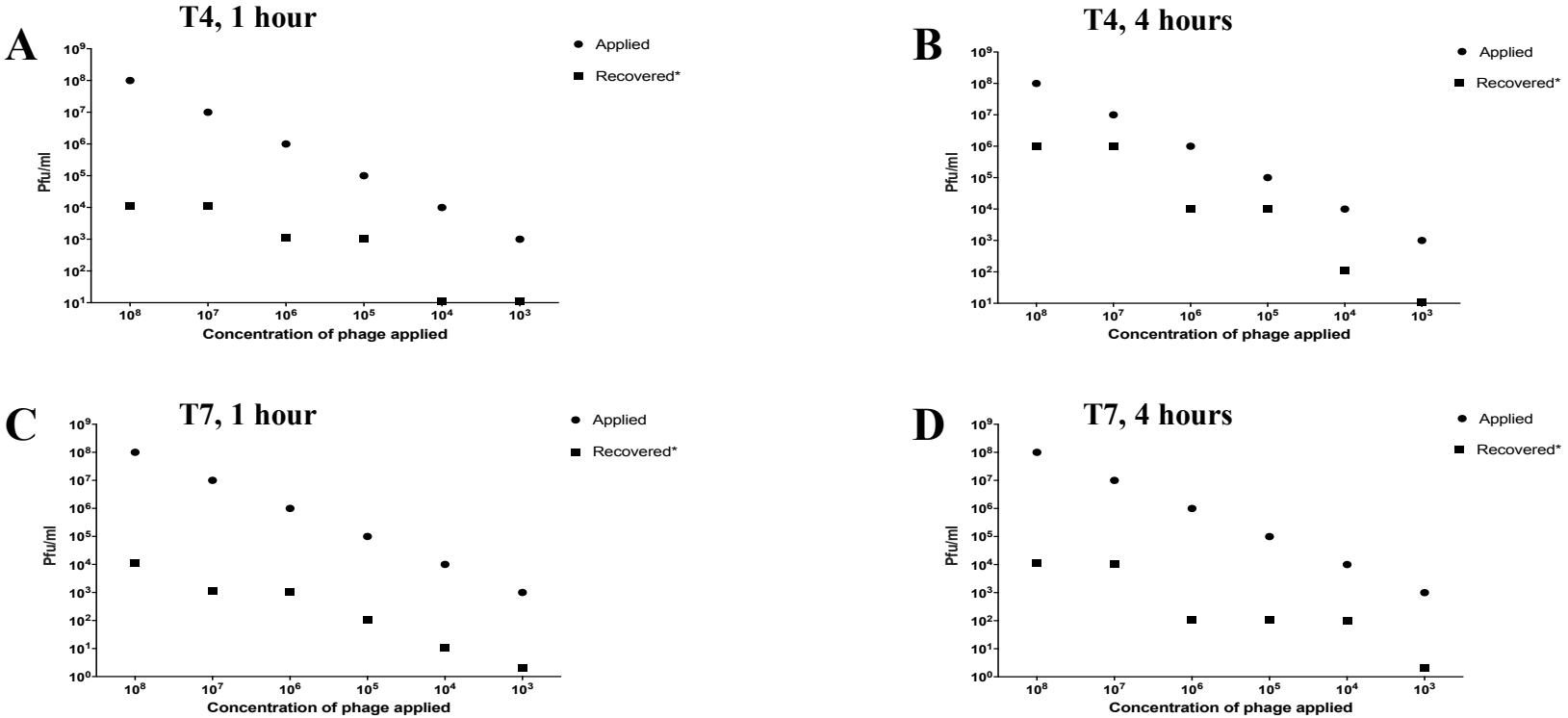


Figure 7-5. Phage recovery from A459 cells. 10^3 to 10^8 PFU/ml of T7 or T4 phage was incubated with A549 cells (80-90% confluence) for 1 or 4 hours. The cells were then washed and lysed, with the phages recovered and counted after both steps. PFU/ml data represented here is from one biological sample, no replicates (N=1 for all conditions). The PFU for each sample was determined on the *E. coli* strain BW25113 using the spot assay method. Phage species and incubation times: **(A)** T4, 1 hour, **(B)** T4, 4 hours, **(C)** T7, 1 hour, and **(D)** T47 4 hours. Recovered (*) refers to the total amount of phage recovered when phage numbers after the lysis and wash steps combined.

7.3.4 Mutant T7 and T7 phage adsorption to A549 cells

Another set of adsorption/absorption assays were carried out on mutant and wild type T7 phages to determine if the addition of homing peptides provides greater binding to the mammalian cell surface. After 1 h of infection, all T7 phages with homing peptides incorporated showed increased numbers of loosely bound phages (wash step) as compared to the wild type T7 phage (Figure 7-6). Although, only three of them had significantly higher values to that of T7: HP1, HP3 and HP4. This held true for cell-associated phages (cell-adherent and intracellular phages) after the cell lysis step with the exception of the *GFE-2::10B*-T7 (HP5), which had approximately a 10-fold lower recovery post-lysis than the wild type (Figure 7-6). A prolonged infection (4h) resulted in greater phage recovery numbers for both wash and lysis steps. In the case of the lysis HP2, HP3, HP4 and HP5 showed significantly greater values to that of T7; whereas strongly cell-associated phages, after the lysis step, showed significantly greater amounts of phage recovered for all HP-engineered phages when compared to T7 (7-6).

In order to determine if there is a significant difference between the amount of phage recovered for different phage types (T7 and T7 mutants) a two-way ANOVA was performed using Statistical Package for the Social Sciences (SPSS) software. This statistical method tested three hypotheses, listed in Table 7-4, to determine if there is both a significant difference between the phages recovered given the type of phage (T7 vs T7 mutants) and incubation time, and whether there is a significant interaction between phage type and incubation time. The analysis was performed separately for the wash and lysis steps (Appendix 5).

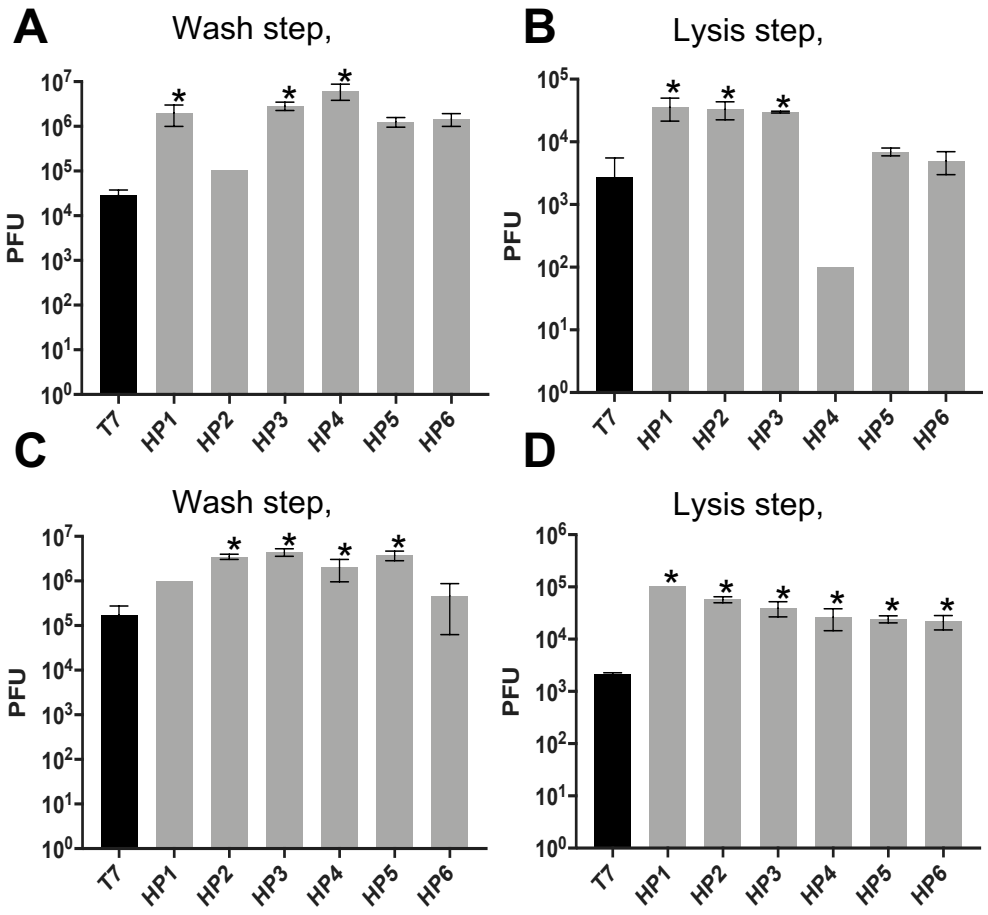


Figure 7-6. Phage enumeration after recovery from A549 cells after wash and lysis steps. 2 x 10^8 PFU/ml of each mutant T7 (HP1-HP6) and T7 phages were incubated with A549 cells (80-90% confluence) for 1 (A, B) or 4 hours (C, D). The cells were then washed and lysed, with phages recovered and counted after wash (A, C) and lysis (B, D) steps. PFU/ml data are from three technical replicates (N=3 for all conditions). The PFU for each sample was determined on the *E. coli* BW25113 strain using the plaque assay. * - $p < 0.05$, when compared to T7 (Appendix 5). N=1, error bars come from technical replicates and in some instances too small to be represented on the chart.

Table 7-4. Summary of the tested hypotheses.

Set No	H_0	H_1
1	The mean amounts of phages recovered after the wash/lysis step across each phage type are equal.	The mean amounts of phages recovered after the wash/lysis step across each phage type are not all equal.
2	The mean amounts of phages recovered after the wash/lysis step across each incubation time are equal.	The mean amounts of phages recovered after the wash/lysis step across each incubation time are not all equal.
3	There is no significant correlation between the phage type and incubation time for the wash/lysis step.	There is significant correlation between the phage type and incubation time for the wash/lysis step.

The results of the two-way ANOVA for the wash step allowed rejection of the null hypothesis ($H_0$1) (p-value = 0.091) demonstrating there is a statistically significant difference in the recovery for phage types across samples. In the case of the second hypothesis set the p-value was found to be 0.501, implying that there is not enough evidence to reject $H_0$2 and support $H_1$2. Therefore, the mean amount of phage recovered could be the same for the two incubation times. For the third hypothesis, there was statistically significant difference between the phage type and incubation time (p-value <0.001). Given that the first and the third hypothesis sets were statistically significant further analyses were carried out. For the first hypothesis set a post-hoc test (Least Significant Difference) was used to determine which phage types had statistically significant different mean amounts of recovered phage. The test revealed that there was a statistically significant increase in phage recovery for phages that contained homing peptides (HP1-5) with p-values <0.05. For the third hypothesis set t-tests were used to determine which phage types and incubation times had significant interactions T7 mutant phages with (HP2, HP5 and HP6) yielded p-values <0.05 meaning that the incubation time has a significant effect upon phage recovery.

Applying the statistical techniques used for the wash step, it was found that there was a statistically significant increase in phage recovery for phages that contained homing peptides (HP1 – 6) with p-values <0.05. When the mean amount of phage recovered between incubation time points for each phage type was analysed T7 mutant phages (HP2, HP5 and HP6) yielded p-values <0.05 meaning that the incubation time has a significant effect upon phage recovery.

When p-values for a post-hoc test were compared for each of the lysis and wash steps (Appendix 5) T7-HP3 mutant yielded the most efficient binding.

7.4 Discussion

7.4.1 Stability of T7 capsid

10B protein was used more frequently for phage display capsid modifications as it is expressed at lower frequency than 10A, and hence its modifications were less likely to affect the integrity of the capsid[458]. Perhaps a more frequent use of 10B for phage display is a safer option that does not affect the capsid protein integrity when fusing larger proteins to the capsid. It has been shown that stable phage capsids can be formed only from 10A protein, suggesting that 10B is not necessary for stable capsid formation[469,455]. Furthermore, in the previous study, T7 10A protein was tagged to display peptides homing directly to the placenta. Even though, the capsid was not modified at the genome level (complementation of 10A capsid with appended peptide *in trans* was used) the result suggested that major capsid protein tagging did not disrupt the overall structure of the capsid[470]. In this study, it has been confirmed with both TEM as well as plaque assays that 10A protein tagging with homing peptides does not alter T7 structure and infectivity respectively. GFE-1 and GFE-2 are relatively small peptides 13 and 8 amino acids in length respectively, both identified using phage display method[464]. In contrast, the third homing peptide (MTDH) is 63 amino acids long and been identified as a protein that migrates and binds to lungs during breast cancer metastasis[463]. The size of this homing peptide could potentially be riskier when tagged after 10A protein. However, this was not the case as the T7-HP3 not only forms a visibly intact capsid as supported with TEM data but also shows the greatest increase of phage recovered for both lysis and wash steps. Which could be due to the fact that the peptide was attached to 10A capsid protein, resulting in more abundant peptide on the capsid protein.

7.4.2 Adsorption/absorption assay distinguishes between loosely and tightly associated phage

In this work an adsorption/absorption assay was used to quantify the phage recovered from wash and lysis steps. The wash step allows the identification of loosely bound phages. Given the recent study where phage entry into different mammalian tissues was demonstrated, the assay used in this study did not allow

the discrimination between phages bound strongly to the cell surface or phages that have entered the cell[471]. In this recent study an *in vitro* method was used to demonstrate the rapid and directional transcytosis of different phages across the lung, gut, kidney, liver and brain[471]. It was also shown that phage transcytosis across cell layers had a significant preferential directionality for apical-to-basolateral transport[471].

In a previous study, with a similar experimental set-up to this chapter, T7 and T4 phage adsorption/absorption onto mammalian cells was evaluated on A549 cells[471]. It was found that after applying approximately 10^8 PFU/ml of T4 and T7 to the cell surface, approximately 10^3 to10^4 PFU/ml was recovered for both phages after two hours of incubation. In the results presented in this chapter, a similar output of approximately 10^4 PFU/ml was recovered for both phages after one hour of incubation, and for the T7 phage after 4 hours of incubation. However, incubating the T4 phage for 4 hours resulted in a phage recovery level of approximately 100-fold higher than the amount recovered by this previous study[471].

7.4.3 HP engineered phages towards the greater concentration at a site of infection

Homing peptides have previously been used to enhance treatment specificity and drug concentrations at target site[472–474]. In all instances, a drug tagged with a homing peptide resulted in greater amounts at the target site when compared to the non-tagged drug[472–474]. In one of the studies a nanoparticle drug, abraxane, was modified with tumour specific homing peptides[472]. It was shown that LyP-1 homing peptide tagged abraxane, improved the efficacy of the drug[472]. The results generated in this chapter showed statistically significant increase in phage recovery for phages that contained a homing peptide for HP3-5 for the strongly-associated phages and HP1-3 for the loosely bound phages, as showed by wash and lysis phage enumeration results respectively.

In general, all of the T7 mutants showed greater amounts of phage recovery, suggesting that homing peptides facilitate greater attachment of T7 phage to the lung tissue epithelium cells. The amount of phage recovered for phage T7 was

approximately 10^4 Pfu/ml whereas for some T7 mutants it was approximately 10^7 Pfu/ml. Given, the surface are of the flask (75 cm^2), the confluence that was approximately 85%, as well as approximate size of A549 cell (100μm^2), the approximate number of cells per flask was 6 x 10^7. This suggest that approximately one phage per mammalian cell was strongly bound or internalised. It has been shown previously that to achieve effective killing at a site of infection the phage concentration should reach the concentration of 10^7 or higher[150]. This suggests that T7 mutants, generated in this study, would be preferable for therapeutic applications over the T7 phage. Furthermore, phages generated in this study would provide a more effective treatment when targeting intracellular bacteria such as *Listeria monocytogenes*, *Salmonella typhi*, *Mycobacterium spp*[475].

7.4.4 Potential use of homing peptides for tissue targeting *in vivo*

Results generated in this chapter suggest that when using the phage T7 tagged with homing peptides, higher phage concentrations can be achieved at a target site *in vitro.* To confirm if this translates to *in vivo* systems further experimentation is required. However, previous studies have shown comparability between *in vitro* and *in vivo* outcomes[471,473,476]. In one of the studies, F3 peptide, homing to tumour endothelial cells, when bound to cisplatin loaded nanoparticles showed greater drug concentration and binding specificity to the target tissue (when compared to non-tagged drug) for both *in vitro* and *in vivo* studies[476]. In another study, cage-like protein nanoparticle genetically engineered with LyP-1 peptide showed greater binding to vascular macrophages *in vitro* and *in vivo* studies. Even though these studies did not use phage, they strongly suggest that, *in vivo* studies with homing peptide tagged phages could confirm the very promising results obtained *in vitro*.

When considering a target mammalian tissue for phage targeting and localisation a potential constrains need to be considered e.g. the presence of lysosomes. Due to its acidic pH, lysosomes may affect bacteriophage activity[477]. In the previous study, in order to determine if phage activity may be altered by the lysosomes, the lysosomal acidic environment was neutralised[478]. Phage activity prior and post neutralisation was compared, showing approximately 2-fold greater total phage recovery indicating that the lysosome does inactivate the internalised phages to an

extent[478]. Therefore, lysosomal activity effects when administering phages need to be considered.

Chapter 8 Conclusions and Future Directions

In this thesis, the limitations for the phage therapy were assessed and potential engineering solutions to address them were suggested. In this chapter, the conclusions and future work of this thesis are discussed in the context of the original aims of the study highlighted in Chapter 1.24.

Firstly (Aim 1), comparison of the efficiencies of marker-based and marker-less methods in the selection efficiency for engineered bacteriophages was sought. It was determined that *trxA* marker-based method was the most efficient method for engineered bacteriophage selection. Out of two marker-less methods compared, CRISPR type II was the more efficient one at targeting wild type sequence and hence allowing more effective engineered phage selection. For both, marker-based and marker-less, methods T7 essential structural gene was targeted for the first time. These results will aid future bacteriophage engineering and more optimal method design.

In order to improve CRISPR type I and type II efficiencies a pre-induction step before the addition of homologously recombined phage could be included in the future work. IPTG induction would be used to prevent T7 promoter inhibition whereas moving into BL21-AI strain and consequent induction with arabinose would provide even further control of the Cas protein expression. In the case of CRISPR type II, Cas9, trcRNA as well as the spacer region control could be exchanged with *E. coli* inducible promoters instead of *S. pyogens* native ones. This would then allow induction of the system in question. In reference to section 3.7.3 the genomic contexts of target DNA as well as the GC percentage contribute to sgRNA performance therefore both should be considered when designing CRISPR gRNAs. In addition, base pair positions might be considered when designing gRNAs as it has been showed that some base pairs might be preferred over others for the gRNA design, for example: 1) cytosine is preferred at the -3 position, which is the DNA cleavage site by the CRISPR/Cas9 complex, 2) guanines are strongly preferred at the -1 and the -2 positions proximal to the PAM sequence, which are associated with the sequence preference in Cas9 loading[479–481]. Moreover, screening for escape mutants should be carried out to find out if these mutants emerge when essential genes of phage are targeted, in the same manner as it has been shown for non-essential genes[329,327].

Secondly (Aim 2), the possibility of T7 host range of alteration when tail fibre domains from a genetically distant host are appended to it, was disseminated. The finding from Aim 1 on *trxA* being the most effective selection for engineered phage, facilitated the means of obtaining the Aim 2. When *trxA* method without *in trans* complementation was used, a persistent wild type T7 contaminant was detected throughout the purification process and after; therefore, *trxA* method with *in trans* complementation was used to generate seven different chimeric tail fibre phages. Fusion tail fiber characterisation *in vitro* indicated that foldon, as a linker between the fusion tail fiber components, was not sufficient to ensure tail fiber stability and folding. In addition, varying the fusion tail fiber length did not influence the tail fiber stability. The data obtained indicates that further factors need to be considered to aid future fusion tail fiber designs, if fusions between such distant phages, as T7 and BPP1, are feasible at all.
Furthermore, an interesting find was stumbled upon, where T7 showed lysis against *B. bronchiseptica* strain on the solid medium, implying previously unexpectedly broader host range of phage T7.

To potentially improve the stability of fusion tail fibers, future studies should consider a wider range linkers to be included between two of the domains fused as well as inclusion of the β-prism domain reinforce overall trimeric assembly[482]. In addition, shorter regions of C-terminal domain of Mtd could be used for the fusions, in particular, only the variable regions to enhance the fusion tail fiber stability. Moreover, to better understand *B. bronchiseptica* infection by T7, further examination is needed. This could be done by carrying out 16S to verify the strain as well as transcriptomics to understand if LPS expression profile is maintained throughout the experimentation and perhaps if its changes lead to susceptibility to T7.

The finding in Chapter 4 (to answer aim 2), where a persistent wild type T7 contaminant was present throughout the purification process when no *in trans* method was used to select for chimeric phages, led to further experiments in Chapter 5. Here the aim was to determine if co-infection is possible between T7

and recombinant phage mutants. To achieve this T7 and chimeric phage interactions were assessed in solid and liquid media. In Chapter 5, for the first time I have demonstrated co-infection in phage T7. This is an important finding as it implies novel possible ways of phage co-evolution and cross-effects that haven't been considered before. This result indicates that if two T7 phages can co-evolve in the laboratory a similar behaviour must be present in the nature. Therefore, lytic phage evolution may not occur in isolation but with other phages. The data also shows that combining the marker-based method with CRISPR/Cas results in removal of the co-infection, providing a solution for future phage engineering.

Future work should seek to unravel the potential mechanism suggested in Chapter 5.3.4. This could be achieved by incorporating fluorescent markers in both T7 and chimeric phages to understand their better interplay.

The Aim 3, that was assessed in Chapter 6, sought to determine if T4 can be engineered to kill a broader range of bacteria. To achieve this, the study yielded four T4 mutants, for three of which their infectivity on solid medium was tested against multiple strains. This work confirms that T4 can be engineered with different lysis yielding functional phage that is capable of cargo delivery. LysH5 when expressed as part of T4-*ac::lysH5* lysate after infection of *E. coli MG1655,* showed broad host range, including lysis of Gram-negative bacterium that hasn't been shown before. LysPBC1 showed lysis on *B. subtilis* and partial lysis on *B. cereus.* LysCD27L generated no lysis in this study. None of the three lysins tested showed lysis in liquid culture. Lysin gene stability in T4 genome as well as virulence assays showed that the lysin is stable in T4 genome for over 72 generations and does not alter phage infectivity when compared to wild type T4. The activity of two lysins was verified on solid medium validating that lysin production in T4 serves as a reliable expression system that can generate stable lysin with its catalytic intact. T4 engineered with a lysin is a promising therapeutic that would allow two-fold, Gram-negative and Gram-positive targeting, by T4 and its cargo respectively.

Future work should include showing lysin expression by RT-PCR or proteomics. In addition, lysin quantification in the T4 lysate should be attempted by modifying

lysin sequence to include by His-tag followed affinity chromatography. Tunability of lysin expression should be also achieved by placing the lysin gene under different promoters in addition to that of T4 *ac* gene. Further liquid characterisation assays including MIC assays should be performed.

Chapter 7, Aim 4, examined the potential for applying phage engineering to determine whether T7 engineered with homing peptides could allow increased adsorption/absorption to the mammalian cells. In this chapter it was determined that both wild type T4 and T7 phages can be adsorbed to and potentially absorbed by mammalian cells. T7 engineered with three different homing peptides provided greater association with lung epithelium tissue when compared to the phage T7. This is the first study to show homing peptide benefits for phage delivery. This finding proposes that the potential use of these engineered phages in therapeutic setting would yield higher pathogen kill at a site of infection.

Future work should assess phage T7 engineered with homing peptides using Imunno-gold TEM should to quantify the amount of homing peptides in the T7 capsid. Furthermore, T7-HP phages should be characterised on different tissues to verify their selectivity to lung epithelium tissue only.

References

1. World Health Organization. WHO | Global action plan on antimicrobial resistance. *WHO* (2015).
2. Spellberg, B. & Shlaes, D. Prioritized current unmet needs for antibacterial therapies. *Clinical Pharmacology and Therapeutics* **96,** 151–153 (2014).
3. Bush, K. *et al.* Tackling antibiotic resistance. *Nature Reviews Microbiology* **9,** 894–896 (2011).
4. Czaplewski, L. *et al.* Alternatives to antibiotics-a pipeline portfolio review. *The Lancet Infectious Diseases* **16,** 239–251 (2016).
5. Sulakvelidze, A., Alavidze, Z. & Morris, J. G. Bacteriophage Therapy. *Antimicrob. Agents Chemother.* **45,** 649–659 (2001).
6. Kęsik-Szeloch, A. *et al.* Characterising the biology of novel lytic bacteriophages infecting multidrug resistant Klebsiella pneumoniae. *Virol. J.* **10,** (2013).
7. Oliveira, H. *et al.* A lytic Providencia rettgeri virus of potential therapeutic value is a deepbranching member of the T5virus genus. *Appl. Environ. Microbiol.* **83,** (2017).
8. Dcosta, V. M. *et al.* Antibiotic resistance is ancient. *Nature* **477,** 457–461 (2011).
9. Woolhouse, M., Ward, M., Van Bunnik, B. & Farrar, J. Antimicrobial resistance in humans, livestock and the wider environment. *Philos. Trans. R. Soc. B Biol. Sci.* **370,** (2015).
10. Woolhouse, M. E. J. & Ward, M. J. Sources of antimicrobial resistance. *Science* **341,** 1460–1461 (2013).
11. D'Costa, V. M., Griffiths, E. & Wright, G. D. Expanding the soil antibiotic resistome: exploring environmental diversity. *Current Opinion in Microbiology* **10,** 481–489 (2007).
12. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bull. World Health Organ.* **79,** 780–790 (2001).
13. Kardos, N. & Demain, A. L. Penicillin: The medicine with the greatest impact on therapeutic outcomes. *Applied Microbiology and Biotechnology* **92,** 677–687 (2011).

14. Schatz, A., Bugie, E. & Waksman, S. A. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. 1944. *Clin. Orthop. Relat. Res.* 3–6 (2005). doi:10.1097/01.blo.0000175887.98112.fe
15. Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. Drugs for bad bugs: Confronting the challenges of antibacterial discovery. *Nature Reviews Drug Discovery* **6,** 29–40 (2007).
16. Davies, J. Where have all the antibiotics gone? in *Canadian Journal of Infectious Diseases and Medical Microbiology* **17,** 287–290 (2006).
17. Sukkar, E. Why are there so few antibiotics in the research and development pipeline? *Pharm. J.* **291,** 520 (2013).
18. Lewis, K. Recover the lost art of drug discovery. *Nature* **485,** 439–440 (2012).
19. Palliyil, S., Downham, C., Broadbent, I., Charlton, K. & Porter, A. J. High-sensitivity monoclonal antibodies specific for homoserine lactones protect mice from lethal Pseudomonas aeruginosa infections. *Appl. Environ. Microbiol.* **80,** 462–469 (2014).
20. Smith, H. W. & Huggins, M. B. Effectiveness of Phages in Treating Experimental Escherichia coli Diarrhoea in Calves, Piglets and Lambs. *Microbiology* **129,** 2659–2675 (1983).
21. Hua, L. *et al.* Assessment of an anti-alpha toxin mAb for prevention and treatment of Staphylococcus aureus induced pneumonia. *Antimicrob. Agents Chemother.* **58,** 1108–1117 (2014).
22. Villena, J. & Kitazawa, H. Modulation of intestinal TLR4-inflammatory signaling pathways by probiotic microorganisms: Lessons learned from Lactobacillus jensenii TL2937. *Frontiers in Immunology* **4,** (2014).
23. Goldenberg, J. *et al.* Probiotics for the prevention of Clostridium difficile-associated diarrhea in adults and children (Review). *Cochrane Libr.* (2013). doi:10.1002/14651858.CD006095.pub3.Copyright
24. Bak, H. & Rathkjen, P. H. Reduced use of antimicrobials after vaccination of pigs against porcine proliferative enteropathy in a Danish SPF herd. *Acta Vet. Scand.* **51,** (2009).
25. Matthews, L. *et al.* Predicting the public health benefit of vaccinating cattle against Escherichia coli O157. *Proc. Natl. Acad. Sci.* **110,** 16265–

16270 (2013).

26. Moriel, D. G. *et al.* Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic Escherichia coli. *Proc. Natl. Acad. Sci.* **107,** 9072–9077 (2010).
27. Schuch, R. *et al.* Combination therapy with lysin CF-301 and antibiotic is superior to antibiotic alone for treating methicillin-resistant staphylococcus aureus-induced murine bacteremia. *J. Infect. Dis.* **209,** 1469–1478 (2014).
28. Yang, H., Yu, J. & Wei, H. Engineered bacteriophage lysins as novel anti-infectives. *Frontiers in Microbiology* **5,** (2014).
29. Jassim, S. A. A. & Limoges, R. G. Impact of external forces on cyanophage-host interactions in aquatic ecosystems. *World Journal of Microbiology and Biotechnology* **29,** 1751–1762 (2013).
30. Bergh, Ø., Børsheim, K. Y., Bratbak, G. & Heldal, M. High abundance of viruses found in aquatic environments. *Nature* **340,** 467–468 (1989).
31. Jassim, S. A. A., Abdulamir, A. S. & Bakar, F. A. Novel phage-based bio-processing of pathogenic Escherichia coli and its biofilms. *World J. Microbiol. Biotechnol.* **28,** 47–60 (2012).
32. Hibma, A. M., Jassim, S. A. A. & Griffiths, M. W. Infection and removal of L-forms of Listeria monocytogenes with bred bacteriophage. *Int. J. Food Microbiol.* **34,** 197–207 (1997).
33. Bright, N. G., Carroll, R. J. & Applegate, B. M. A model system for pathogen detection using a two-component bacteriophage/bioluminescent signal amplification assay. in *Proceedings of SPIE* **5271,** 13 (2004).
34. Favrin, S. J., Jassim, S. A. & Griffiths, M. W. Application of a novel immunomagnetic separation-bacteriophage assay for the detection of Salmonella enteritidis and Escherichia coli O157:H7 in food. *Int. J. Food Microbiol.* **85,** 63–71 (2003).
35. Shan, J., Jia, Y., Clokie, M. R. J. & Mann, N. H. Infection by the 'photosynthetic' phage S-PM2 induces increased synthesis of phycoerythrin in Synechococcus sp. WH7803. *FEMS Microbiol. Lett.* **283,** 154–161 (2008).
36. Bragg, J. G. & Chisholm, S. W. Modeling the fitness consequences of a cyanophage-encoded photosynthesis gene. *PLoS One* **3,** (2008).

37. Šulčius, S. *et al.* Characterization of a lytic cyanophage that infects the bloom-forming cyanobacterium Aphanizomenon flos-aquae. *FEMS Microbiol. Ecol.* **91,** (2015).
38. Kot, W., Hammer, K., Neve, H. & Vogensen, F. K. Identification of the receptor-binding protein in lytic Leuconostoc pseudomesenteroides bacteriophages. *Appl. Environ. Microbiol.* **79,** 3311–3314 (2013).
39. Orlova, E. V. Bacteriophages and Their Structural Organisation. in *Bacteriophages* (2012). doi:10.5772/34642
40. Catalão, M. J., Gil, F., Moniz-Pereira, J., São-José, C. & Pimentel, M. Diversity in bacterial lysis systems: Bacteriophages Show the way. *FEMS Microbiol. Rev.* **37,** 554–571 (2013).
41. Abedon, S. T. Lysis from without. *Bacteriophage* **1,** 46–49 (2011).
42. Delbruck, M. THE GROWTH OF BACTERIOPHAGE AND LYSIS OF THE HOST. *J. Gen. Physiol.* **23,** 643–660 (1940).
43. Young, R. Bacteriophage lysis: mechanism and regulation. *Microbiol. Rev.* **56,** 430–481 (1992).
44. Arisaka, F., Kanamaru, S., Leiman, P. & Rossmann, M. G. The tail lysozyme complex of bacteriophage T4. *International Journal of Biochemistry and Cell Biology* **35,** 16–21 (2003).
45. Tarahovsky, Y. S., Ivanitsky, G. R. & Khusainov, A. A. Lysis of Escherichia coli cells induced by bacteriophage T4. *FEMS Microbiol. Lett.* **122,** 195–199 (1994).
46. Cota-Robles, E. H. Electron microscopy of 'lysis from within' of Escherichia coli by coliphage T2. *J. Ultrasructure Res.* **11,** 112–122 (1964).
47. Mardanov, A. V. & Ravin, N. V. Conversion of Linear DNA with Hairpin Telomeres into a Circular Molecule in the Course of Phage N15 Lytic Replication. *J. Mol. Biol.* **391,** 261–268 (2009).
48. Kim, M. & Ryu, S. Antirepression System Associated with the Life Cycle Switch in the Temperate Podoviridae Phage SPC32H. *J. Virol.* **87,** 11775–11786 (2013).
49. Ravin, V. K. & Shulga, M. G. Evidence for extrachromosomal location of prophage N15. *Virology* **40,** 800–807 (1970).
50. Allardet-Servent, A., Michaux-Charachon, S., Jumas-Bilak, E., Karayan,

L. & Ramuz, M. Presence of one linear and one circular chromosome in the Agrobacterium tumefaciens C58 genome. *J. Bacteriol.* **175,** 7869–7874 (1993).

51. Erez, Z. *et al.* Communication between viruses guides lysis-lysogeny decisions. *Nature* **541,** 488–493 (2017).
52. Menouni, R., Hutinet, G., Petit, M. A. & Ansaldi, M. Bacterial genome remodeling through bacteriophage recombination. *FEMS Microbiology Letters* **362,** (2015).
53. Burns, N., James, C. E. & Harrison, E. Polylysogeny magnifies competitiveness of a bacterial pathogen in vivo. *Evol. Appl.* **8,** 346–351 (2015).
54. Allison, H. E. *et al.* Immunity profiles of wild-type and recombinant Shiga-like toxin-encoding bacteriophages and characterization of novel double lysogens. *Infect. Immun.* **71,** 3409–3418 (2003).
55. Chung, I. Y., Bae, H. W., Jang, H. J., Kim, B. o. & Cho, Y. H. Superinfection exclusion reveals heteroimmunity between Pseudomonas aeruginosa temperate phages. *J. Microbiol.* **52,** 515–520 (2014).
56. Refardt, D. Within-host competition determines reproductive success of temperate bacteriophages. *ISME J.* **5,** 1451–1460 (2011).
57. Schuch, R. & Fischetti, V. A. The secret life of the anthrax agent Bacillus anthracis: Bacteriophage-mediated ecological adaptations. *PLoS One* **4,** (2009).
58. Los, M., Wegrzyn, G. & Neubauer, P. A role for bacteriophage T4 rI gene function in the control of phage development during pseudolysogeny and in slowly growing host cells. *Res. Microbiol.* **154,** 547–552 (2003).
59. Clokie, M. R. J., Millard, A. D., Letarov, A. V. & Heaphy, S. Phages in nature. *Bacteriophage* **1,** 31–45 (2011).
60. Los, M. & Wegrzyn, G. Pseudolysogeny. in *Bacteriophages: Part A* 396 (2012).
61. Ogunseitan, O. A., Sayler, G. S. & Miller, R. V. Dynamic interactions of Pseudomonas aeruginosa and bacteriophages in lake water. *Microb. Ecol.* **19,** 171–185 (1990).
62. Kokjohn, T., Sayler, G. & Miller, R. Attachment and replication of Pseudomonas aeruginosa bacteriophages under conditions simulating

aquatic environments. *J. Gen. Microbiol.* 661–666 (1991). doi:10.1099/00221287-137-3-661

63. ROMIG, W. R. & BRODETSKY, A. M. Isolation and preliminary characterization of bacteriophages for Bacillus subtilis. *J. Bacteriol.* **82,** 135–141 (1961).
64. Ripp, S. & Miller, R. V. The role of pseudolysogeny in bacteriophage-host interactions in a natural freshwater environment. *Microbiology* **143,** 2065–2070 (1997).
65. Ripp, S. & Miller, R. V. Dynamics of the pseudolysogenic response in slowly growing cells of Pseudomonas aeruginosa. *Microbiology* **144,** 2225–2232 (1998).
66. Wilson, W. H., Carr, N. G. & Mann, N. H. The effect of phosphate status on the kinetics of cyanophage infection in the oceanic cyanobacterium Synechococcus sp. WH7803. *J. Phycol.* **32,** 506–516 (1996).
67. Chanishvili, N. Phage Therapy-History from Twort and d'Herelle Through Soviet Experience to Current Approaches. *Adv. Virus Res.* **83,** 3–40 (2012).
68. Wittebole, X., De Roock, S. & Opal, S. M. A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. *Virulence* **5,** 226–235 (2013).
69. Ramesh, V., Fralick, J. A. & Rolfe, R. D. Prevention of Clostridium difficile -induced ileocecitis with Bacteriophage. *Anaerobe* **5,** 69–78 (1999).
70. Biswas, B. *et al.* Bacteriophage Therapy Rescues Mice Bacteremic from a Clinical Isolate of Vancomycin-Resistant Enterococcus Bacteriophage Therapy Rescues Mice Bacteremic from a Clinical Isolate of Vancomycin-Resistant Enterococcus faecium. *Infect. Immun.* **70,** 204–210 (2002).
71. Wang, J. *et al.* Use of bacteriophage in the treatment of experimental animal bacteremia from imipenem-resistant Pseudomonas aeruginosa. *Int. J. Mol. Med.* **17,** 309–317 (2006).
72. Watanabe, R. *et al.* Efficacy of bacteriophage therapy against gut-derived sepsis caused by Pseudomonas aeruginosa in mice. *Antimicrob. Agents Chemother.* **51,** 446–452 (2007).
73. Soothill, J. S. Treatment of experimental infections of mice with

bacteriophages. *J. Med. Microbiol.* **37,** 258–261 (1992).

74. Reardon, S. Phage therapy gets revitalized. *Nature* **510,** 15–16 (2014).
75. Górski, A. *et al.* Phage Therapy: Beyond Antibacterial Action. *Front. Med.* **5,** (2018).
76. Międzybrodzki, R. *et al.* Clinical aspects of phage therapy. *Adv. Virus Res.* **83,** 73–121 (2012).
77. Weber-Dabrowska, B., Mulczyk, M. & Górski, a. Bacteriophage therapy of bacterial infections: an update of our institute's experience. *Arch. Immunol. Ther. Exp. (Warsz).* **48,** 547–551 (2000).
78. Weber-Dabrowska, B. *et al.* Bacteriophage procurement for therapeutic purposes. *Frontiers in Microbiology* **7,** (2016).
79. Abedon, S. T., Kuhl, S. J., Blasdel, B. G. & Kutter, E. M. Phage treatment of human infections. *Bacteriophage* **1,** 66–85 (2011).
80. Slopek, S., Kucharewicz-Krukowska, A., Weber-Dabrowska, B. & Dabrowski, M. Results of bacteriophage treatment of suppurative bacterial infections. IV. Evaluation of the results obtained in 370 cases. *Arch. Immunol. Ther. Exp. (Warsz).* **33,** 219–40 (1985).
81. Slopek, S. *et al.* Results of bacteriophage treatment of suppurative bacterial infections. II. Detailed evaluation of the results. *Archivum immunologiae et therapiae experimentalis* **31,** 293–327 (1987).
82. Slopek, S., Kucharewicz-Krukowska, A., Weber-Dabrowska, B. & Dabrowski, M. Results of bacteriophage treatment of suppurative bacterial infections. V. Evaluation of the results obtained in children. *Arch. Immunol. Ther. Exp. (Warsz).* **33,** 241–59 (1985).
83. Slopek, S., Kucharewicz-Krukowska, A., Weber-Dabrowska, B. & Dabrowski, M. Results of bacteriophage treatment of suppurative bacterial infections. VI. Analysis of treatment of suppurative staphylococcal infections. *Arch. Immunol. Ther. Exp. (Warsz).* **33,** 261–73 (1985).
84. Verbeken, G. *et al.* European regulatory conundrum of phage therapy. *Future Microbiol.* **2,** 485–491 (2007).
85. Verbeken, G. *et al.* Call for a dedicated European legal framework for bacteriophage therapy. *Arch. Immunol. Ther. Exp. (Warsz).* **62,** 117–29 (2014).
86. Huys, I. *et al.* Paving a regulatory pathway for phage therapy. *EMBO Rep.*

14, 951–954 (2013).

87. Henein, A. What are the limitations on the wider therapeutic use of phage? *Bacteriophage* **3,** e24872 (2013).
88. Pelfrene, E., Willebrand, E., Cavaleiro Sanches, A., Sebris, Z. & Cavaleri, M. Bacteriophage therapy: A regulatory perspective. *J. Antimicrob. Chemother.* **71,** 2071–2074 (2016).
89. Fauconnier, A. Regulating phage therapy. *EMBO Rep.* **18,** 198–200 (2017).
90. Program, P. C., Resources, R., Act, C., Pharmacology, A. & Protocols, C. Investigational New Drug (IND) Application. *fda.gov website* 1–5 (2018).
91. Kutateladze, M. & Adamia, R. Phage therapy experience at the Eliava Institute. *Med. Mal. Infect.* **38,** 426–430 (2008).
92. Fish, R. *et al.* Bacteriophage treatment of intransigent diabetic toe ulcers: a case series. *J. Wound Care* **25,** S27–S33 (2016).
93. Jennes, S. *et al.* Use of bacteriophages in the treatment of colistin-only-sensitive Pseudomonas aeruginosa septicaemia in a patient with acute kidney injury-a case report. *Critical Care* **21,** (2017).
94. Duplessis, C. *et al.* Refractory Pseudomonas Bacteremia in a 2-Year-Old Sterilized by Bacteriophage Therapy. *J. Pediatric Infect. Dis. Soc.* (2017). doi:10.1093/jpids/pix056
95. Schooley, R. T. *et al.* Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant Acinetobacter baumannii infection. *Antimicrob. Agents Chemother.* **61,** (2017).
96. Monk, A. B., Rees, C. D., Barrow, P., Hagens, S. & Harper, D. R. Bacteriophage applications: Where are we now? *Letters in Applied Microbiology* **51,** 363–369 (2010).
97. Nannapaneni, R. & Soni, K. A. Use of Bacteriophages to Remove Biofilms of Listeria monocytogenes and other Foodborne Bacterial Pathogens in the Food Environment. in *Biofilms in the Food Environment: Second Edition* 131–144 (2015). doi:10.1002/9781118864036.ch5
98. Endersen, L. *et al.* Phage Therapy in the Food Industry. *Annu. Rev. Food Sci. Technol.* **5,** 327–349 (2014).

99. Schofield, D. A. *et al.* Bacillus anthracis diagnostic detection and rapid antibiotic susceptibility determination using 'bioluminescent' reporter phage. *J. Microbiol. Methods* **95,** 156–161 (2013).
100. Barlow, S. *et al.* The use and mode of action of bacteriophages in food production 1 Scientific Opinion of the Panel on Biological Hazards Adopted on 22 April 2009. *EFSA J.* **10,** 1–9 (2009).
101. Hyman, P. & Abedon, S. T. Bacteriophage host range and bacterial resistance. *Advances in applied microbiology* **70,** 217–248 (2010).
102. Adams, M. Bacteriophages. *Bacteriophages* 620 (1959). doi:58-12722
103. Deveau, H., Garneau, J. E. & Moineau, S. CRISPR/Cas System and Its Role in Phage-Bacteria Interactions. *Annu. Rev. Microbiol.* **64,** 475–493 (2010).
104. Fineran, P. C. *et al.* The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl. Acad. Sci.* **106,** 894–899 (2009).
105. Jacobs-Sera, D. *et al.* On the nature of mycobacteriophage diversity and host preference. *Virology* **434,** 187–201 (2012).
106. Bielke, L., Higgins, S., Donoghue, A., Donoghue, D. & Hargis, B. M. Salmonella host range of bacteriophages that infect multiple genera. *Poult. Sci.* **86,** 2536–2540 (2007).
107. Jensen, E. C. *et al.* Prevalence of broad-host-range lytic bacteriophages of Sphaerotilus natans, Escherichia coli, and Pseudomonas aeruginosa. *Appl. Environ. Microbiol.* **64,** 575–580 (1998).
108. Samson, J. E., Magadán, A. H., Sabri, M. & Moineau, S. Revenge of the phages: Defeating bacterial defences. *Nature Reviews Microbiology* **11,** 675–687 (2013).
109. Rothenberg, E. *et al.* Single-virus tracking reveals a spatial receptor-dependent search mechanism. *Biophys. J.* **100,** 2875–2882 (2011).
110. Parisien, A., Allain, B., Zhang, J., Mandeville, R. & Lan, C. Q. Novel alternatives to antibiotics: Bacteriophages, bacterial cell wall hydrolases, and antimicrobial peptides. *Journal of Applied Microbiology* **104,** 1–13 (2008).
111. Merabishvili, M. *et al.* Quality-controlled small-scale production of a well-defined bacteriophage cocktail for use in human clinical trials. *PLoS*

One **4,** (2009).

112. Žliobaitė, I., Fortelius, M. & Stenseth, N. C. Reconciling taxon senescence with the Red Queen's hypothesis. *Nature* **552,** 92–95 (2017).
113. Ormälä, A.-M. & Jalasvuori, M. Phage therapy: Should bacterial resistance to phages be a concern, even in the long run? *Bacteriophage* **3,** e24219 (2013).
114. Qimron, U., Marintcheva, B., Tabor, S. & Richardson, C. C. Genomewide screens for Escherichia coli genes affecting growth of T7 bacteriophage. **2006,** (2006).
115. Haggard-Ljungquist, E., Halling, C. & Calendar, R. DNA sequences of the tail fiber genes of bacteriophage P2: Evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J. Bacteriol.* **174,** 1462–1477 (1992).
116. Riede, I., Eschbach, M. L. & Henning, U. Presence of DNA, encoding parts of bacteriophage tail fiber genes, in the chromosome of Escherichia coli K-12. *J. Bacteriol.* **163,** 832–836 (1985).
117. Montag, D., Hashemolhosseini, S. & Henning, U. Receptor-recognizing proteins of T-even type bacteriophages. The receptor-recognizing area of proteins 37 of phages T4 TuIa and TuIb. *J. Mol. Biol.* **216,** 327–334 (1990).
118. Golomidova, A. K. *et al.* Branched lateral tail fiber organization in T5-like bacteriophages DT57C and DT571/2 is revealed by genetic and functional analysis. *Viruses* **8,** (2016).
119. Heller, K. J. & Bryniok, D. O antigen-dependent mutant of bacteriophage T5. *J. Virol.* **49,** 20–5 (1984).
120. Mondigler, M., Holz, T. & Heller, K. J. Identification of the receptor-binding regions of pb5 proteins of bacteriophages T5 and BF23. *Virology* **219,** 19–28 (1996).
121. Zivanovic, Y. *et al.* Insights into Bacteriophage T5 Structure from Analysis of Its Morphogenesis Genes and Protein Components. *J. Virol.* **88,** 1162–1174 (2014).
122. Heller, K. & Braun, V. Accelerated adsorption of bacteriophage T5 to Escherichia coli F, resulting from reversible tail fiber-lipopolysaccharide binding. *J. Bacteriol.* **139,** 32–38 (1979).

123. Heller, K. & Braun, V. Polymannose O-antigens of Escherichia coli, the binding sites for the reversible adsorption of bacteriophage T5+ via the L-shaped tail fibers. *J. Virol.* **41,** 222–7 (1982).
124. Liu, M. *et al.* Reverse transcriptase-mediated tropism switching in Bordetella bacteriophage. *Science* **295,** 2091–2094 (2002).
125. Dai, W. *et al.* Three-dimensional structure of tropism-switching Bordetella bacteriophage. *Proc. Natl. Acad. Sci.* **107,** 4347–4352 (2010).
126. Uhl, M. A. & Miller, J. F. Integration of multiple domains in a two-component sensor protein: the Bordetella pertussis BvgAS phosphorelay. *EMBO J.* **15,** 1028–36 (1996).
127. Akerley, B. J., Cotter, P. A. & Miller, J. F. Ectopic expression of the flagellar regulon alters development of the bordetella-host interaction. *Cell* **80,** 611–620 (1995).
128. Mattoo, S., Foreman-Wykert, A. K., Cotter, P. A. & Miller, J. F. Mechanisms of Bordetella Pathogenesis. *Front. Biosci.* **6,** 168–186 (2001).
129. Liu, M. *et al.* Genomic and Genetic Analysis of Bordetella Bacteriophages Encoding Reverse Transcriptase-Mediated Tropism-Switching Cassettes. *J. Bacteriol.* **186,** 1503–1517 (2004).
130. Liu, M. *et al.* Tropism switching in Bordetella bacteriophage defines a family of diversity-generating retroelements. *Nature* **431,** 476–481 (2004).
131. Carlton, R. M. Phage therapy: past history and future prospects. *Arch. Immunol. Ther. Exp. (Warsz).* **47,** 267–274 (1999).
132. Stratton, C. W. Dead bugs don't mutate: Susceptibility issues in the emergence of bacterial resistance. *Emerging Infectious Diseases* **9,** 10–16 (2003).
133. Loc-Carrillo, C. & Abedon, S. Pros and cons of phage therapy. *Bacteriophage* **1,** 111–114 (2011).
134. Capparelli, R. *et al.* Bacteriophage therapy of Salmonella enterica: a fresh appraisal of bacteriophage therapy. *J. Infect. Dis.* **201,** 52–61 (2010).
135. Skurnik, M. therapy: F. and fiction & Strauch, E. Phage therapy: Facts and fiction. *International Journal of Medical Microbiology* **296,** 5–14 (2006).
136. Kutter, E. *et al.* Phage Therapy in Clinical Practice: Treatment of Human Infections. *Curr. Pharm. Biotechnol.* **11,** 69–86 (2010).
137. Skurnik, M., Pajunen, M. & Kiljunen, S. Biotechnological challenges of

phage therapy. *Biotechnology Letters* **29,** 995–1003 (2007).

138. Clokie, M. R. J. & Kropinski, A. M. *Bacteriophages : methods and protocols Volume 1: Isolation, Characterization, and Interactions*. *Methods in molecular biology* (2009). doi:10.1007/978-1-60327-164-6
139. Krylov, V. N. Phage Therapy in Terms of Bacteriophage Genetics: Hopes, Prospects, Safety, Limitations. *Russian Journal of Genetics* **37,** 715–730 (2001).
140. Bentley, R. & Bennett, J. W. What is an antibiotic? Revisited. *Advances in Applied Microbiology* **52,** 303–331 (2003).
141. Abedon, S. T. *Bacteriophages and biofilms: Ecology, phage therapy, plaques*. *Bacteriophages and Biofilms: Ecology, Phage Therapy, Plaques* (2011).
142. Gupta, R. & Prasad, Y. Efficacy of polyvalent bacteriophage P-27/HP to control multidrug resistant staphylococcus aureus associated with human infections. *Curr. Microbiol.* **62,** 255–260 (2011).
143. Mann, N. H. The potential of phages to prevent MRSA infections. *Res. Microbiol.* **159,** 400–405 (2008).
144. Mackowiak, M., Leifels, M., Hamza, I. A., Jurzik, L. & Wingender, J. Distribution of Escherichia coli, coliphages and enteric viruses in water, epilithic biofilms and sediments of an urban river in Germany. *Sci. Total Environ.* **626,** 650–659 (2018).
145. Dhillon, T. S., Chan, Y. S., Sun, S. M. & Chau, W. S. Distribution of coliphages in Hong Kong sewage. *Appl. Microbiol.* **20,** 187–91 (1970).
146. Ashelford, K. E., Day, M. J. & Fry, J. C. Elevated abundance of bacteriophage infecting bacteria in soil. *Appl. Environ. Microbiol.* **69,** 285–289 (2003).
147. Parasion, S., Kwiatek, M., Gryko, R., Mizak, L. & Malm, A. Bacteriophages as an alternative strategy for fighting biofilm development. *Polish Journal of Microbiology* **63,** 137–145 (2014).
148. Gill, J. J. & Hyman, P. Phage choice, isolation, and preparation for phage therapy. *Curr. Pharm. Biotechnol.* **11,** 2–14 (2010).
149. Qadir, M. I., Mobeen, T. & Masood, A. Phage therapy: Progress in pharmacokinetics. *Brazilian Journal of Pharmaceutical Sciences* **54,** (2018).

150. Abedon, S. T. & Thomas-Abedon, C. Phage therapy pharmacology. *Curr. Pharm. Biotechnol.* **11,** 28–47 (2010).
151. Dennehy, J. J. What Can Phages Tell Us about Host-Pathogen Coevolution? *Int. J. Evol. Biol.* **2012,** 1–12 (2012).
152. Nilsson, A. S. Phage therapy-constraints and possibilities. *Upsala Journal of Medical Sciences* **119,** 192–198 (2014).
153. Levin, B. R. & Bull, J. J. Population and evolutionary dynamics of phage therapy. *Nat. Rev. Microbiol.* **2,** 166–173 (2004).
154. Gasiunas, G., Sinkunas, T. & Siksnys, V. Molecular mechanisms of CRISPR-mediated microbial immunity. *Cellular and Molecular Life Sciences* **71,** 449–465 (2014).
155. Szczepankowska, A. Role of CRISPR/cas System in the Development of Bacteriophage Resistance. *Adv. Virus Res.* **82,** 289–338 (2012).
156. Dy, R. L., Przybilski, R., Semeijn, K., Salmond, G. P. C. & Fineran, P. C. A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism. *Nucleic Acids Res.* **42,** 4590–4605 (2014).
157. Ofir, G. *et al.* DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.* **3,** 90–98 (2018).
158. LURIA, S. E. & HUMAN, M. L. A nonhereditary, host-induced variation of bacterial viruses. *J. Bacteriol.* **64,** 557–569 (1952).
159. Szermer-Olearnik, B. & Boratyński, J. Removal of endotoxins from bacteriophage preparations by extraction with organic solvents. *PLoS One* **10,** (2015).
160. Dofferhoff, A. S. M. *et al.* Effects of different types and combinations of antimicrobial agents on endotoxin release from gram-negative bacteria: An in-vitro and in-vivo study. *Scand. J. Infect. Dis.* **23,** 745–754 (1991).
161. Rietschel, E. T. *et al.* Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J.* **8,** 217–225 (1994).
162. Gorbet, M. B. & Sefton, M. V. Endotoxin: The uninvited guest. *Biomaterials* **26,** 6811–6817 (2005).
163. de Oliveira Magalhães, P. *et al.* Methods of endotoxin removal from biological preparations: A review. *Journal of Pharmacy and Pharmaceutical Sciences* **10,** 388–404 (2007).

164. Alisky, J., Iczkowski, K., Rapoport, A. & Troitsky, N. Bacteriophages show promise as antimicrobial agents. *J. Infect.* **36,** 5–15 (1998).
165. Górski, A., Borysowski, J., Międzybrodzki, R. & Weber-Dąbrowska, B. Bacteriophages in Medicine. in *Bacteriophage: Genetics and Molecular Biology* 125–158 (2007).
166. Kutateladze, M. & Adamia, R. Bacteriophages as potential new therapeutics to replace or supplement antibiotics. *Trends Biotechnol.* **28,** 591–595 (2010).
167. Goodridge, L. D. Designing phage therapeutics. *Curr. Pharm. Biotechnol.* **11,** 15–27 (2010).
168. Fischer, S., Kittler, S., Klein, G. & Glünder, G. Impact of a Single Phage and a Phage Cocktail Application in Broilers on Reduction of Campylobacter jejuni and Development of Resistance. *PLoS One* **8,** (2013).
169. Gu, J. *et al.* A method for generation phage cocktail with great therapeutic potential. *PLoS One* **7,** (2012).
170. Bogovazova, G. G., Voroshilova, N. N. & Bondarenko, V. M. The efficacy of Klebsiella pneumoniae bacteriophage in the therapy of experimental Klebsiella infection. *Zh. Mikrobiol. Epidemiol. Immunobiol.* 5–8 (1991).
171. Dabrowska, K., Switała-Jelen, K., Opolski, A., Weber-Dabrowska, B. & Gorski, A. A review: Bacteriophage penetration in vertebrates. *Journal of Applied Microbiology* **98,** 7–13 (2005).
172. Keller, R. Fate of Bacteriophage Particles Introduced into Mice by Various Routes. *Dep. Microbiol. Sch. Med. Univ. Missouri, Colunzbia* 1–4
173. Chibani-Chennoufi, S. *et al.* In vitro and in vivo bacteriolytic activities of Escherichia coli phages: Implications for phage therapy. *Antimicrob. Agents Chemother.* **48,** 2558–2569 (2004).
174. Brüssow, H. Phage therapy: The Escherichia coli experience. *Microbiology* **151,** 2133–2140 (2005).
175. Ochs, H. D., Davis, S. D. & Wedgwood, R. J. Immunologic responses to bacteriophage ϕX 174 in immunodeficiency diseases. *J. Clin. Invest.* **50,** 2559–2568 (1971).
176. Fogelman, I. *et al.* Evaluation of CD4+ T cell function In vivo in HIV-

infected patients as measured by bacteriophage phiX174 immunization. *J. Infect. Dis.* **182,** 435–41 (2000).

177. D browska, K. *et al.* Immunogenicity Studies of Proteins Forming the T4 Phage Head Surface. *J. Virol.* **88,** 12551–12557 (2014).
178. Górski, A. *et al.* Phage as a Modulator of Immune Responses. Practical Implications for Phage Therapy. *Adv. Virus Res.* **83,** 41–71 (2012).
179. Payne, R. J. H. & Jansen, V. A. A. Understanding bacteriophage therapy as a density-dependent kinetic process. *Journal of Theoretical Biology* **208,** 37–48 (2001).
180. Abedon, S. Phage Therapy Pharmacology. Calculating Phage Dosing. *Adv. Appl. Microbiol.* **77,** 1–40 (2011).
181. Górski, A. *et al.* Phages targeting infected tissues: Novel approach to phage therapy. *Future Microbiology* **10,** 199–204 (2015).
182. Shimer, D. a, Jenkins, P. L., Hui, S. P. & Adams, W. C. 132 Measurement of Breathing Rate and Volume in Routinely Performed Daily Activities. *Epidemiology* **6,** S30 (1995).
183. Song, J. H. & Chung, D. R. Respiratory infections due to drug-resistant bacteria. *Infectious Disease Clinics of North America* **24,** 639–653 (2010).
184. Wedgwood, R. J., Ochs, H. D. & Davis, S. D. The recognition and classification of immunodeficiency diseases with bacteriophage phiChi 174. *Birth Defects Orig. Artic. Ser.* **11,** 331–338 (1975).
185. Bearden, C. M. *et al.* Rituximab inhibits the in vivo primary and secondary antibody response to a neoantigen, bacteriophage phiX174. *Am. J. Transplant.* **5,** 50–57 (2005).
186. Sokoloff, A. V., Bock, I., Zhang, G., Sebestyén, M. G. & Wolff, J. A. The interactions of peptides with the innate immune system studied with use of T7 phage peptide display. *Mol. Ther.* **2,** 131–139 (2000).
187. Merril, C. R. *et al.* Long-circulating bacteriophage as antibacterial agents. *Proc. Natl. Acad. Sci.* **93,** 3188–3192 (1996).
188. Dabrowska, K. *et al.* Hoc protein regulates the biological effects of T4 phage in mammals. *Arch. Microbiol.* **187,** 489–498 (2007).
189. Miernikiewicz, P. *et al.* T4 Phage and Its Head Surface Proteins Do Not Stimulate Inflammatory Mediator Production. *PLoS One* **8,** (2013).
190. Molenaar, T. J. M. *et al.* Uptake and processing of modified bacteriophage

M13 in mice: Implications for phage display. *Virology* **293,** 182–191 (2002).

191. Russel, M., Linderoth, N. A. & Andrej, S. Filamentous phage assembly : variation on a protein export theme 1. **192,** 23–32 (1997).
192. Endemann, H. & Model, P. Location of filamentous phage minor coat proteins in phage and in infected cells. *J. Mol. Biol.* **250,** 496–506 (1995).
193. Smilowitz, H. Bacteriophage f1 infection: fate of the parental major coat protein. *J. Virol.* **13,** 94–9 (1974).
194. Ohkawa, I. & Webster, R. E. The orientation of the major coat protein of bacteriophage f1 in the cytoplasmic membrane of Escherichia coli. *J Biol Chem* **256,** 9951–9958 (1981).
195. Young, R., Wang, I. N. & Roof, W. D. Phages will out: Strategies of host cell lysis. *Trends in Microbiology* **8,** 120–128 (2000).
196. Iedaite ˙, G. Ž., Daugelavičius, R., Bamford, J. K. H. & Bamford, D. H. The Holin Protein of Bacteriophage PRD1 Forms a Pore for Small-Molecule and Endolysin Translocation. *J. Bacteriol.* **187,** 5397–5405 (2005).
197. Campbell, J. H. & Rolfe, B. G. Evidence for a dual control of the initiation of host-cell lysis caused by phage lambda. *MGG Mol. Gen. Genet.* **139,** 1–8 (1975).
198. Steiner, M., Lubitz, W. & Blasi, U. The missing link in phage lysis of gram-positive bacteria: Gene 14 of Bacillus subtilis phage φ29 encodes the functional homolog of lambda S protein. *J. Bacteriol.* **175,** 1038–1042 (1993).
199. Dunne, M. *et al.* Crystal structure of the CTP1L endolysin reveals how its activity is regulated by a secondary translation product. *J. Biol. Chem.* **291,** 4882–4893 (2016).
200. Borysowski, J. A. N. Bacteriophage Endolysins as a Novel Class of Antibacterial Agents. 366–377 (2006).
201. Fischetti, V. A. Bacteriophage endolysins: A novel anti-infective to control Gram-positive pathogens. *International Journal of Medical Microbiology* **300,** 357–362 (2010).
202. Miller, E. S. *et al.* Bacteriophage T4 Genome. *Microbiol. Mol. Biol. Rev.* **67,** 86–156 (2003).

203. Ochman, H. & Wilson, A. C. Evolution in bacteria: Evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* **26,** 74–86 (1987).
204. Olsen, G. J., Woese, C. R. & Overbeek, R. The winds of (evolutionary) change: Breathing new life into microbiology. *Journal of Bacteriology* **176,** 1–6 (1994).
205. Ackermann, H. W. *et al.* Aeromonas bacteriophages:Reexamination and classification. *Ann. l'Institut Pasteur Virol.* **136,** 175–199 (1985).
206. Tétart, F., Repoila, F., Monod, C. & Krisch, H. M. Bacteriophage T4 host range is expanded by duplications of a small domain of the tail fiber adhesin. *J. Mol. Biol.* **258,** 726–731 (1996).
207. Riede, I. Receptor specificity of the short tail fibres (gp12) of T-even type Escherichia coli phages. *MGG Mol. Gen. Genet.* **206,** 110–115 (1987).
208. Structure and function of bacteriophage T4Yap, M. L. & Rossmann, M. G. Structure and function of bacteriophage T4. *Future Microbiol.* **9,** 1319–1327 (2014).
209. Baschong, W. *et al.* Mass analysis of bacteriophage T4 proheads and mature heads by scanning transmission electron microscopy and hydrodynamic measurements. *J. Struct. Biol.* **106,** 93–101 (1991).
210. Baschong, W. *et al.* Head structure of bacteriophages T2 and T4. *J. Ultrastruct. Res. Mol. Struct. Res.* **99,** 189–202 (1988).
211. Onorato, L., Stirmer, B. & Showe, M. K. Isolation and characterization of bacteriophage T4 mutant preheads. *J. Virol.* **27,** 409–426 (1978).
212. Karam, J. D. & Drake, J. W. *Molecular biology of bacteriophage T4*. *Molecular biology of bacteriophage T4* (1994).
213. Wilkens, K., Tiemann, B., Bazan, F. & Ruger, W. ADP-ribosylation and early transcription regulation by bacteriophage T4. *Adv Exp Med Biol* **419,** 71–82 (1997).
214. Hinton, D. M. Transcription from a bacteriophage T4 middle promoter using T4 motA protein and phage-modified RNA polymerase. *J. Biol. Chem.* **266,** 18034–18044 (1991).
215. Ouhammouch, M., Adelman, K., Harvey, S. R., Orsini, G. & Brody, E. N. Bacteriophage T4 MotA and AsiA proteins suffice to direct Escherichia coli RNA polymerase to initiate transcription at T4 middle promoters.

Proc. Natl. Acad. Sci. U. S. A. **92,** 1451–5 (1995).

216. Ouhammouch, M., Orsini, G. & Brody, E. N. The asiA gene product of bacteriophage T4 is required for middle mode RNA synthesis. *J. Bacteriol.* **176,** 3956–3965 (1994).
217. Mosig, G. RECOMBINATION AND RECOMBINATION-DEPENDENT DNA REPLICATION IN BACTERIOPHAGE T4. *Annu. Rev. Genet.* **32,** 379–413 (1998).
218. Noble, E., Spiering, M. M. & Benkovic, S. J. Coordinated DNA replication by the bacteriophage T4 replisome. *Viruses* **7,** 3186–3200 (2015).
219. Venkatesan, M. & Nossal, N. G. Bacteriophage T4 gene 44/62 and gene 45 polymerase accessory proteins stimulate hydrolysis of duplex DNA by T4 DNA polymerase. *J. Biol. Chem.* **257,** 12435–12443 (1982).
220. Karam, J. D. & Konigsberg, W. H. DNA polymerase of the T4-related bacteriophages. *Prog. Nucleic Acid Res. Mol. Biol.* **64,** 65–96 (2000).
221. Venkatesan, M., Silver, L. L. & Nossal, N. G. Bacteriophage T4 gene 41 protein, required for the synthesis of RNA primers, is also a DNA helicase. *J. Biol. Chem.* **257,** 12426–12434 (1982).
222. Young, M. C., Schultz, D. E., Ring, D. & Von Hippel, P. H. Kinetic parameters of the translocation of bacteriophage t4 gene 41 protein helicase on single-stranded DNA. *J. Mol. Biol.* **235,** 1447–1458 (1994).
223. Molineux-The T7 group-2005.pdf.
224. Dunx, J. J. & Studier, F. W. Complete Nucleotide Sequence of Bacteriophage T7 D N A and the Locations of T7 Genetic Elements. (1983).
225. Serwer, P., Wright, E. T., Hakala, K. W. & Weintraub, S. T. Evidence for bacteriophage T7 tail extension during DNA injection. *BMC Res. Notes* **1,** (2008).
226. Englers, M. J. & Richardson, C. C. Bacteriophage T7 DNA replication. *J. Biol. Chem.* **258,** 11197–11205 (1983).
227. Lee, S.-J. & Richardson, C. C. Choreography of bacteriophage T7 DNA replication. *Curr. Opin. Chem. Biol.* **15,** 580–6 (2011).
228. Zavriev, S. K. & Shemyakin, M. F. RNA polymerase-dependent mechanism for the stepwise T7 phage DNA transport from the virion into

E. coli. *Nucleic Acids Res.* **10,** 1635–1652 (1982).

229. Moffatt, B. A. & Studier, F. W. Entry of bacteriophage T7 DNA into the cell and escape from host restriction. *J. Bacteriol.* **170,** 2095–2105 (1988).
230. Garcia, L. R. & Molineux, I. J. Rate of translocation of bacteriophage T7 DNA across the membranes of Escherichia coli. *J. Bacteriol.* **177,** 4066–4076 (1995).
231. Studier, F. W. Genetic analysis of non-essential bacteriophage T7 genes. *J. Mol. Biol.* **79,** 227–236 (1973).
232. Moak, M. & Molineux, I. J. Role of the Gp16 lytic transglycosylase motif in bacteriophage T7 virions at the initiation of infection. *Mol. Microbiol.* **37,** 345–355 (2000).
233. Walkinshaw, M. D. *et al.* Structure of Ocr from bacteriophage T7, a protein that mimics b-form DNA. *Mol. Cell* **9,** 187–194 (2002).
234. Gomez, B. & Nualart, L. Requirement of the bacteriophage T7 0.7 gene for phage growth in the presence of the Col 1b factor. *J. Gen. Virol.* **35,** 99–106 (1977).
235. Lyakhov, D. L. *et al.* Pausing and termination by bacteriophage T7 RNA polymerase. *J. Mol. Biol.* **280,** 201–213 (1998).
236. Tahlrov, T. H. *et al.* Structure of a T7 RNA polymerase elongation complex at 2.9 Å resolution. *Nature* **420,** 43–50 (2002).
237. Yin, Y. W. & Steitz, T. A. The Structural Mechanism of Translocation and Helicase Activity in T7 RNA Polymerase. *Cell* **116,** 393–404 (2004).
238. Nechaev, S. & Severinov, K. Inhibition of Escherichia coli RNA polymerase by bacteriophage T7 gene 2 protein. *J. Mol. Biol.* **289,** 815–826 (1999).
239. Doublié, S., Tabor, S., Long, A. M., Richardson, C. C. & Ellenberger, T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391,** 251–258 (1998).
240. Lee, J., Chastain, P. D., Griffith, J. D. & Richardson, C. C. Lagging strand synthesis in coordinated DNA synthesis by bacteriophage T7 replication proteins. *J. Mol. Biol.* **316,** 19–34 (2002).
241. Lee, J., Chastain, P. D., Kusakabe, T., Griffith, J. D. & Richardson, C. C. Coordinated leading and lagging strand DNA synthesis on a minicircular template. *Mol. Cell* **1,** 1001–1010 (1998).

242. Molineux, I. J. Host-parasite interactions: recent developments in the genetics of abortive phage infections. *New Biol.* **3,** 230–6 (1991).
243. Molineux, I. J. No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. *Molecular Microbiology* **40,** 1–8 (2001).
244. Steven, A. C. *et al.* Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. *J. Mol. Biol.* **200,** 351–365 (1988).
245. Wang, I.-N., Smith, D. L. & Young, R. Holins: The Protein Clocks of Bacteriophage Infections. *Annu. Rev. Microbiol.* **54,** 799–825 (2000).
246. White, J. H. & Richardson, C. C. Gene 18 protein of bacteriophage T7. Overproduction, purification, and characterization. *J. Biol. Chem.* **262,** 8845–8850 (1987).
247. Kemp, P., Garcia, L. R. & Molineux, I. J. Changes in bacteriophage T7 virion structure at the initiation of infection. *Virology* **340,** 307–317 (2005).
248. Agirrezabala, X. *et al.* Structure of the connector of bacteriophage T7 at 8 Å resolution: Structural homologies of a basic component of a DNA translocating machinery. *J. Mol. Biol.* **347,** 895–902 (2005).
249. Garcia-Doval, C. & Van Raaij, M. J. Crystallization of the C-terminal domain of the bacteriophage T7 fibre protein gp17. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **68,** 166–171 (2012).
250. Casjens, S. R. & Molineux, I. J. Short noncontractile tail machines: Adsorption and DNA delivery by podoviruses. *Adv. Exp. Med. Biol.* **726,** 143–179 (2012).
251. Richardson, C. C. Bacteriophage T7: Minimal requirements for the replication of a duplex DNA molecule. *Cell* **33,** 315–317 (1983).
252. Hamdan, S. M. & Richardson, C. C. Motors, Switches, and Contacts in the Replisome. *Annu. Rev. Biochem.* **78,** 205–243 (2009).
253. Qimron, U., Lee, S. J., Hamdan, S. M. & Richardson, C. C. Primer initiation and extension by T7 DNA primase. *EMBO J.* **25,** 2199–2208 (2006).
254. Kolodner, R. & Richardson, C. C. Replication of duplex DNA by bacteriophage T7 DNA polymerase and gene 4 protein is accompanied by hydrolysis of nucleoside 5'-triphosphates. *Proc Natl Acad Sci U S A* **74,**

1525–1529 (1977).

255. Vlamis-Gardikas, A. & Holmgren, A. Thioredoxin and glutaredoxin isoforms. *Methods Enzymol.* **347,** 286–296 (2002).

256. Collet, J.-F. & Messens, J. Structure, Function, and Mechanism of Thioredoxin Proteins. *Antioxid. Redox Signal.* **13,** 1205–1216 (2010).

257. Singha, N. C., Vlamis-Gardikas, A. & Holmgren, A. Real-time kinetics of the interaction between the two subunits, Escherichia coli thioredoxin and gene 5 protein of phage T7 DNA polymerase. *J. Biol. Chem.* **278,** 21421–21428 (2003).

258. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol. Syst. Biol.* **2,** (2006).

259. Tran, N. Q., Lee, S. J., Akabayov, B., Johnson, D. E. & Richardson, C. C. Thioredoxin, the processivity factor, sequesters an exposed cysteine in the thumb domain of bacteriophage T7 DNA polymerase. *J. Biol. Chem.* **287,** 39732–39741 (2012).

260. Huber, H. E., Tabor, S. & Richardson, C. C. Escherichia coli thioredoxin stabilizes complexes of bacteriophage T7 DNA polymerase and primed templates. *J. Biol. Chem.* **262,** 16224–16232 (1987).

261. Tabor, S., Huber, H. E. & Richardson, C. C. Escherichia coli thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem.* **262,** 16212–16223 (1987).

262. Patel, S. S., Wong, I. & Johnson, K. A. Pre-Steady-State Kinetic-Analysis Of Processive Dna-Replication Including Complete Characterization Of An Exonuclease-Deficient Mutant. *Biochemistry* **30,** 511–525 (1991).

263. Doyle, R. J. & Marquis, R. E. Elastic, flexible peptidoglycan and bacterial cell wall properties. *Trends Microbiol.* **2,** 57–60 (1994).

264. Schmelcher, M., Donovan, D. M. & Loessner, M. J. Bacteriophage endolysins as novel antimicrobials. doi:10.2217/fmb.12.97

265. López, R. & García, E. Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. *FEMS Microbiology Reviews* **28,** 553–580 (2004).

266. Loessner, M. J., Wendlinger, G. & Scherer, S. Heterogeneous endolysins in Listeria monocytogenes bacteriophages: a new class of enzymes and evidence for conserved holin genes within the siphoviral lysis cassettes.

Mol. Microbiol. **16,** 1231–1241 (1995).

267. Loessner, M. J., Kramer, K., Ebel, F. & Scherer, S. C-terminal domains of Listeria monocytogenes bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. *Mol. Microbiol.* **44,** 335–349 (2002).
268. Cheng, X., Zhang, X., Pflugrath, J. W. & Studier, F. W. The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proc. Natl. Acad. Sci.* **91,** 4034–4038 (1994).
269. Briers, Y. *et al.* Muralytic activity and modular structure of the endolysins of Pseudomonas aeruginosa bacteriophages φKZ and EL. *Mol. Microbiol.* **65,** 1334–1344 (2007).
270. Nelson, D. C. *et al.* Endolysins as Antimicrobials. *Adv. Virus Res.* **83,** 299–365 (2012).
271. Becker, S. C., Foster-Frey, J., Stodola, A. J., Anacker, D. & Donovan, D. M. Differentially conserved staphylococcal SH3b_5 cell wall binding domains confer increased staphylolytic and streptolytic activity to a streptococcal prophage endolysin domain. *Gene* **443,** 32–41 (2009).
272. Yokoi, K. J. *et al.* The two-component cell lysis genes holWMY and lysWMY of the Staphylococcus warneri M phage φWMY: Cloning, sequencing, expression, and mutational analysis in Escherichia coli. *Gene* **351,** 97–108 (2005).
273. Loessner, M. J. Bacteriophage endolysins - Current state of research and applications. *Current Opinion in Microbiology* **8,** 480–487 (2005).
274. Callewaert, L., Walmagh, M., Michiels, C. W. & Lavigne, R. Food applications of bacterial cell wall hydrolases. *Current Opinion in Biotechnology* **22,** 164–171 (2011).
275. Hermoso, J. A., García, J. L. & García, P. Taking aim on bacterial pathogens: from phage therapy to enzybiotics. *Current Opinion in Microbiology* **10,** 461–472 (2007).
276. Rodríguez-Rubio, L. *et al.* 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,** (2013).
277. Spratt, B. G. Resistance to antibiotics mediated by target alterations.

Science (80-.). **264,** 388–393 (1994).

278. Donovan, D. M. *et al.* Peptidoglycan hydrolase enzyme fusions for treating multi-drug resistant pathogens. in *Bio Tech International* **21,** 6–10 (2009).
279. Loeffler, J. M. & Fischetti, V. A. Synergistic lethal effect of a combination of phage lytic enzymes with different activities on penicillin-sensitive and -resistant Streptococcus pneumoniae strains. *Antimicrob. Agents Chemother.* **47,** 375–377 (2003).
280. Loeffler, J. M., Djurkovic, S. & Fischetti, V. A. Phage Lytic Enzyme Cpl-1 as a Novel Antimicrobial for Pneumococcal Bacteremia. *Infect. Immun.* **71,** 6199–6204 (2003).
281. Schuch, R., Nelson, D. & Fischetti, V. A. A bacteriolytic agent that detects and kills Bacillus anthracis. *Nature* **418,** 884–889 (2002).
282. Cheng, Q., Nelson, D., Zhu, S. & Fischetti, V. A. Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. *Antimicrob. Agents Chemother.* **49,** 111–117 (2005).
283. O'Flaherty, S., Coffey, A., Meaney, W., Fitzgerald, G. F. & Ross, R. P. The recombinant phage lysin LysK has a broad spectrum of lytic activity against clinically relevant staphylococci, including methicillin-resistant Staphylococcus aureus. *J. Bacteriol.* **187,** 7161–7164 (2005).
284. Horgan, M. *et al.* Phage lysin LysK can be truncated to its CHAP domain and retain lytic activity against live antibiotic-resistant staphylococci. *Appl. Environ. Microbiol.* **75,** 872–874 (2009).
285. Becker, S. C., Foster-Frey, J. & Donovan, D. M. The phage K lytic enzyme LysK and lysostaphin act synergistically to kill MRSA. *FEMS Microbiol. Lett.* **287,** 185–191 (2008).
286. Obeso, J. M., Martínez, B., Rodríguez, A. & García, P. Lytic activity of the recombinant staphylococcal bacteriophage ΦH5 endolysin active against Staphylococcus aureus in milk. *Int. J. Food Microbiol.* **128,** 212–218
287. Donovan, D. M., Lardeo, M. & Foster-Frey, J. Lysis of staphylococcal mastitis pathogens by bacteriophage phi11 endolysin. *FEMS Microbiol. Lett.* **265,** 133–139 (2006).

288. Schmelcher, M., Donovan, D. M. & Loessner, M. J. Bacteriophage endolysins as novel antimicrobials. *Future Microbiology* **7,** 1147–1171 (2012).

289. Nau, R. & Eiffert, H. Modulation of release of proinflammatory bacterial compounds by antibacterials: Potential impact on course of inflammation and outcome in sepsis and meningitis. *Clinical Microbiology Reviews* **15,** 95–110 (2002).

290. Entenza, J. M., Loeffler, J. M., Grandgirard, D., Fischetti, V. A. & Moreillon, P. Therapeutic effects of bacteriophage Cpl-1 lysin against Streptococcus pneumoniae endocarditis in rats. *Antimicrob. Agents Chemother.* **49,** 4789–4792 (2005).

291. Witzenrath, M. *et al.* Systemic use of the endolysin Cpl-1 rescues mice with fatal pneumococcal pneumonia. *Crit. Care Med.* **37,** 642–649 (2009).

292. Jado, I. *et al.* Phage lytic enzymes as therapy for antibiotic-resistant Streptococcus pneumoniae infection in a murine sepsis model. *J. Antimicrob. Chemother.* **52,** 967–973 (2003).

293. Rashel, M. *et al.* Efficient Elimination of Multidrug-Resistant *Staphylococcus aureus* by Cloned Lysin Derived from Bacteriophage φMR11. *J. Infect. Dis.* **196,** 1237–1247 (2007).

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

295. Breitbart, M. & Rohwer, F. Here a virus, there a virus, everywhere the same virus? *Trends in Microbiology* **13,** 278–284 (2005).

296. Chibani-Chennoufi, S., Bruttin, A., Dillmann, M. L. & Brüssow, H. Phage-host interaction: An ecological perspective. *Journal of Bacteriology* **186,** 3677–3686 (2004).

297. Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482,** 331–338 (2012).

298. Fineran, P. C. & Charpentier, E. Memory of viral infections by CRISPR-Cas adaptive immune systems: Acquisition of new information. *Virology* **434,** 202–209 (2012).

299. Sturino, J. M. & Klaenhammer, T. R. Engineered bacteriophage-defence systems in bioprocessing. *Nature Reviews Microbiology* **4,** 395–404 (2006).

300. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science (80-.).* **315,** 1709–1712 (2007).
301. Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* **327,** 167–170 (2010).
302. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J. Bacteriol.* **169,** 5429–5433 (1987).
303. Jansen, R., Embden, J. D. A. van, Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43,** 1565–1575 (2002).
304. Haft, D. H., Selengut, J., Mongodin, E. F. & Nelson, K. E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1,** 0474–0483 (2005).
305. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct* **1,** (2006).
306. Kunin, V., Sorek, R. & Hugenholtz, P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* **8,** (2007).
307. Horvath, P. *et al.* Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int. J. Food Microbiol.* **131,** 62–70 (2009).
308. Grissa, I., Vergnaud, G. & Pourcel, C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **8,** (2007).
309. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151,** 653–663 (2005).
310. Bolotin, A., Quinquis, B., Sorokin, A. & Dusko Ehrlich, S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of

extrachromosomal origin. *Microbiology* **151,** 2551–2561 (2005).

311. Horvath, P. *et al.* Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *J. Bacteriol.* **190,** 1401–1412 (2008).
312. Patterson, A. G., Chang, J. T., Taylor, C. & Fineran, P. C. Regulation of the type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Res.* **43,** 6038–6048 (2015).
313. Swarts, D. C., Mosterd, C., van Passel, M. W. J. & Brouns, S. J. J. CRISPR interference directs strand specific spacer acquisition. *PLoS One* **7,** (2012).
314. Patterson, A. G. *et al.* Quorum Sensing Controls Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems. *Mol. Cell* **64,** 1102–1108 (2016).
315. Levy, A. *et al.* CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* **520,** 505–510 (2015).
316. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).
317. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* **4,** e5553 (2009).
318. Charles, I. *et al.* Expression of the Bordetella pertussis P.69 pertactin adhesin in Escherichia coli: Fate of the carboxy-terminal domain. *Microbiology* **140,** 3301–3308 (1994).
319. Jaschke, P. R., Lieberman, E. K., Rodriguez, J., Sierra, A. & Endy, D. A fully decompressed synthetic bacteriophage øX174 genome assembled and archived in yeast. *Virology* **434,** 278–284 (2012).
320. Clara, S. *et al.* (12) United States Patent. **2,** (2012).
321. Ando, H., Lemire, S., Pires, D. P. & Lu, T. K. Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing. *Cell Syst.* **1,** 187–196 (2015).
322. Court, D. L., Sawitzke, J. A. & Thomason, L. C. Genetic Engineering Using Homologous Recombination. *Annu. Rev. Genet.* **36,** 361–388 (2002).
323. Marinelli, L. J. *et al.* BRED: A simple and powerful tool for constructing

mutant and recombinant bacteriophage genomes. *PLoS One* **3,** (2008).

324. Fehér, T., Karcagi, I., Blattner, F. R. & Pósfai, G. Bacteriophage recombineering in the lytic state using the lambda red recombinases. *Microb. Biotechnol.* **5,** 466–476 (2012).
325. Van Kessel, J. C. & Hatfull, G. F. Mycobacterial recombineering. *Methods Mol. Biol.* **435,** 203–215 (2008).
326. Marinelli, L. J., Hatfull, G. F. & Piuri, M. Recombineering: A powerful tool for modification of bacteriophage genomes. *Bacteriophage* **2,** 5–14 (2012).
327. Kiro, R., Shitrit, D. & Qimron, U. Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. *RNA Biol.* **11,** 42–44 (2014).
328. Qimron, U., Marintcheva, B., Tabor, S. & Richardson, C. C. Genomewide screens for Escherichia coli genes affecting growth of T7 bacteriophage. *Proc. Natl. Acad. Sci.* **103,** 19039–19044 (2006).
329. Martel, B. & Moineau, S. CRISPR-Cas: An efficient tool for genome engineering of virulent bacteriophages. *Nucleic Acids Res.* **42,** 9504–9513 (2014).
330. Brouns. Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Genome Res.* **964,** 960–964 (2008).
331. Luo, M. L., Mullis, A. S., Leenay, R. T. & Beisel, C. L. Repurposing endogenous type i CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Res.* **43,** 674–681 (2015).
332. Shah, S. A., Erdmann, S., Mojica, F. J. M. & Garrett, R. A. Protospacer recognition motifs: Mixed identities and functional diversity. *RNA Biol.* **10,** 891–899 (2013).
333. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31,** 233–239 (2013).
334. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (80-.).* **337,** 816–821 (2012).
335. Brouns, S. J. J. *et al.* Small Crispr Rnas Guide Antiviral Defense in Prokaryotes. *Cancer Epidemiol. Biomarkers Prev.* **2,** 531–535 (1993).
336. Westra, E. R. *et al.* Type I-E CRISPR-Cas Systems Discriminate Target

from Non-Target DNA through Base Pairing-Independent PAM Recognition. *PLoS Genet.* **9,** (2013).

337. Tran, N. Q., Rezende, L. F., Qimron, U., Richardson, C. C. & Tabor, S. Gene 1.7 of bacteriophage T7 confers sensitivity of phage growth to dideoxythymidine. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 9373–9378 (2008).
338. Gerdes, S. *et al.* Experimental determination and system level analysis of essential genes in Escherichia coli MG1655. *J. Bacteriol.* **185,** 5673–5684 (2003).
339. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *J. Bacteriol.* **190,** 1390–1400 (2008).
340. Duplessis, M., Russell, W. M., Romero, D. A. & Moineau, S. Global gene expression analysis of two Streptococcus thermophilus bacteriophages using DNA microarray. *Virology* **340,** 192–208 (2005).
341. Lévesque, C. *et al.* Genomic organization and molecular analysis of virulent bacteriophage 2972 infecting an exopolysaccharide-producing Streptococcus thermophilus strain. *Appl. Environ. Microbiol.* **71,** 4057–4068 (2005).
342. Hao, M., Cui, Y. & Qu, X. Analysis of CRISPR-cas system in Streptococcus thermophilus and its application. *Frontiers in Microbiology* **9,** (2018).
343. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152,** 1173–1183 (2013).
344. Liu, X. *et al.* Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Sci. Rep.* **6,** (2016).
345. ADAMS, M. H. & WADE, E. Classification of bacterial viruses; the relationship of two Serratia phages to coli-dysentery phages T3, T7, and D44. *J. Bacteriol.* **68,** 320–325 (1954).
346. Bertani, L. E. & Bertani, G. Genetics of P2 and Related Phages. *Adv. Genet.* **16,** 199–237 (1971).
347. PICKEN, R. N. & BEACHAM, I. R. Bacteriophage-resistant Mutants of Escherichia coli K12. Location of Receptors within the Lipopolysaccharide. *J. Gen. Microbiol.* **102,** 305–318 (1977).
348. González-García, V. A. *et al.* Characterization of the initial steps in the T7

DNA ejection process. *Bacteriophage* **5,** e1056904 (2015).

349. Garcia-Doval, C. & van Raaij, M. J. Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proc. Natl. Acad. Sci.* **109,** 9390–9395 (2012).
350. McMahon, S. A. *et al.* The C-type lectin fold as an evolutionary solution for massive sequence variation. *Nat. Struct. Mol. Biol.* **12,** 886–892 (2005).
351. Emsley, P., Charles, I. G., Fairweather, N. F. & Isaacs, N. W. Structure of Bordetella pertussis virulence factor P.69 pertactin. *Nature* **381,** 90–92 (1996).
352. Diavatopoulos, D. A., Hijnen, M. & Mooi, F. R. Adaptive evolution of the Bordetella autotransporter pertactin. **19,** 1931–1938 (2006).
353. Feinberg, H. *et al.* Structure of a C-type carbohydrate recognition domain from the macrophage mannose receptor. *J. Biol. Chem.* **275,** 21539–21548 (2000).
354. Batchelor, M. *et al.* Structural basis for recognition of the translocated intimin receptor (Tir) by intimin from enteropathogenic Escherichia coli. *Embo J.* **19,** 2452–64 (2000).
355. Luo, Y. *et al.* Crystal structure of enteropathogenic Escherichia coli intimin-receptor complex. *Nature* **405,** 1073–1077 (2000).
356. Garcia-Doval, C. & van Raaij, M. J. Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proc. Natl. Acad. Sci.* **109,** 9390–9395 (2012).
357. Ando, H. *et al.* Engineering Modular Viral Scaffolds. *Cell Syst.* **1,** 187–196 (2015).
358. Miyazaki, J. I., Ryo, Y., Fujisawa, H. & Minagawa, T. Mutation in bacteriophage T3 affecting host cell lysis. *Virology* **89,** 327–329 (1978).
359. Kato, T. & Kondo, S. Two types of X-ray-sensitive mutants of Escherichia coli B: Their phenotypic character compared with UV-sensitive mutants. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **4,** 253–263 (1967).
360. Kaplan, S., Stretton, A. O. W. & Brenner, S. Amber suppressors: Efficiency of chain propagation and suppressor specific amino acids. *J. Mol. Biol.* **14,** IN17-IN24 (1965).
361. Mitraki, A., Papanikolopoulou, K. & Van Raaij, M. J. Natural triple beta-

stranded fibrous folds. *Adv. Protein Chem.* **73,** 97–124 (2006).

362. Papanikolopoulou, K. *et al.* Adenovirus fibre shaft sequences fold into the native triple beta-spiral fold when N-terminally fused to the bacteriophage T4 fibritin foldon trimerisation motif. *J. Mol. Biol.* **342,** 219–227 (2004).
363. Nguyen, T. H. *et al.* Crystal structure of the fibre head domain of bovine adenovirus 4, a ruminant atadenovirus. *Virol. J.* **12,** (2015).
364. González-García, V. A. *et al.* Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *J. Biol. Chem.* **290,** 10038–10044 (2015).
365. Tran, N. Q., Lee, S., Akabayov, B., Johnson, D. E. & Richardson, C. C. Thioredoxin , the Processivity Factor , Sequesters an Exposed Cysteine in the Thumb Domain of Bacteriophage T7 DNA. **287,** 39732–39741 (2012).
366. Quertermous, T. Purification of bacteriophage clones. *Curr. Protoc. Mol. Biol. / Ed. by Frederick M.Ausubel ...[et al.]* **Chapter 6,** Unit6.5 (2001).
367. Hyman, P. & Abedon, S. T. Practical methods for determining phage growth parameters. *Methods Mol. Biol.* **501,** 175–202 (2009).
368. McAllister, W. T. & Barrett, C. L. Superinfection exclusion by bacteriophage T7. *J. Virol.* **24,** 709–11 (1977).
369. Geurtsen, J. *et al.* Identification of a novel lipopolysaccharide core biosynthesis gene cluster in Bordetella pertussis, and influence of core structure and lipid A glucosamine substitution on endotoxic activity. *Infect. Immun.* **77,** 2602–2611 (2009).
370. Ross, A., Ward, S. & Hyman, P. More is better: Selecting for broad host range bacteriophages. *Front. Microbiol.* **7,** (2016).
371. Ghasemian, A., Bavand, M. & Moradpour, Z. A broad-host range coliphage against a clinically isolated E. coli O157: Isolation and characterization. *J. Appl. Pharm. Sci.* **7,** 123–128 (2017).
372. Pierce, J. C. & Masker, W. E. A single-base change in gene 10 of bacteriophage T7 permits growth on Shigella sonnei. *J. Virol.* **62,** 4369–71 (1988).
373. Aksiyote Benbasat, J., Bauman Burck, K. & Miller, R. C. Superinfection exclusion and lack of conservative transfer of bacteriophage T7 DNA. *Virology* **87,** 164–171 (1978).
374. Bauman, K., And, B. & Miller, R. C. Marker rescue and partial replication

of bacteriophage T7 DNA (DNA synthesis/UV irradiation/cross-reactivation). *Genetics* **75,** 6144–6148 (1978).

375. Karska-wysocki, B., Racine, J. & Mamet-, M. D. Alkylation of T7 Bacteriophage Blocks Superinfection Exclusion. **44,** 708–710 (1982).
376. DELBRUCK, M. Interference between bacterial viruses; the mutual exclusion effect and the depressor effect. *J. Bacteriol.* **50,** 151–170 (1945).
377. Cheng, C., Lee, N., Mortazavi, A. & Pritchard, C. T7 Production in Escherichia coli B23 Is Not Inhibited By ΦX-174 at a Multiplicity of Infection of 10. *J. Exp. Microbiol. Immunol.* **18,** 150–155 (2014).
378. Hirsch-Kauffmann, M., Pfennig-Yeh, M. li, Ponta, H. & Herrlich, P. A virus-specified mechanism for the prevention of multiple infection-T7- and T3-mutual and superinfection exclusion. *MGG Mol. Gen. Genet.* **149,** 243–249 (1976).
379. Dunn, J. J., Bautz, F. A. & Bautz, E. K. F. Different template specificities of phage T3 and T7 rna polymerases. *Nat. New Biol.* **230,** 94–96 (1971).
380. Ellis, E. L. THE GROWTH OF BACTERIOPHAGE. *J. Gen. Physiol.* **22,** 365–384 (1939).
381. Lee, S.-J., Chowdhury, K., Tabor, S. & Richardson, C. C. Rescue of bacteriophage T7 DNA polymerase of low processivity by suppressor mutations affecting gene 3 endonuclease. *J. Virol.* **83,** 8418–27 (2009).
382. Hori, K., Mark, D. F. & Richardson, C. C. Deoxyribonucleic acid polymerase of bacteriophage T7. Purification and properties of the phage-encoded subunit, the gene 5 protein. *J. Biol. Chem.* **254,** 11591–7 (1979).
383. Tabor, S., Huber, H. E. & Richardson, C. C. Escherichia coli thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J Biol Chem* **262,** 16212–16223 (1987).
384. Mark, D. F. & Richardson, C. C. Escherichia coli thioredoxin: a subunit of bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **73,** 780–4 (1976).
385. Schmelcher, M. & Loessner, M. J. Bacteriophage endolysins: Applications for food safety. *Current Opinion in Biotechnology* **37,** 76–87 (2016).
386. Hosseini, E. S., Moniri, R., Goli, Y. D. & Kashani, H. H. Purification of Antibacterial CHAPKProtein Using a Self-Cleaving Fusion Tag and Its Activity Against Methicillin-Resistant Staphylococcus aureus. *Probiotics*

Antimicrob. Proteins **8,** 202–210 (2016).

387. Bogaert, D., De Groot, R. & Hermans, P. W. M. Streptococcus pneumoniae colonisation: The key to pneumococcal disease. *Lancet Infectious Diseases* **4,** 144–154 (2004).
388. Kemp, P., Gupta, M. & Molineux, I. J. Bacteriophage T7 DNA ejection into cells is initiated by an enzyme-like mechanism. *Mol. Microbiol.* **53,** 1251–1265 (2004).
389. Nurmemmedov, E., Castelnovo, M., Medina, E., Catalano, C. E. & Evilevitch, A. Challenging packaging limits and infectivity of phage λ. *J. Mol. Biol.* **415,** 263–273 (2012).
390. RA, L. & GD, P. Fields' Virology, Volume 1. in *Paramyxoviridae: the viruses and their replication* 3177 (2007).
391. García, P., Madera, C., Martínez, B. & Rodríguez, A. Biocontrol of Staphylococcus aureus in curd manufacturing processes using bacteriophages. *Int. Dairy J.* **17,** 1232–1239 (2007).
392. Lee, J. H., Shin, H., Son, B. & Ryu, S. Complete genome sequence of Bacillus cereus bacteriophage BCP78. *J Virol* **86,** 637–638 (2012).
393. Mayer, M. J., Narbad, A. & Gasson, M. J. Molecular characterization of a Clostridium difficile bacteriophage and its cloned biologically active endolysin. *J. Bacteriol.* **190,** 6734–40 (2008).
394. Hu, D.-L., Wang, L., Fang, R., Okamura, M. & Ono, H. K. Staphylococcus aureus Enterotoxins. in *Staphylococcus aureus* 39–55 (2018). doi:10.1016/B978-0-12-809671-0.00003-6
395. Gruet, P., Maincent, P., Berthelot, X. & Kaltsatos, V. Bovine mastitis and intramammary drug delivery: Review and perspectives. *Advanced Drug Delivery Reviews* **50,** 245–259 (2001).
396. Obeso, J. M., Mart??nez, B., Rodr??guez, A. & Garc??a, P. Lytic activity of the recombinant staphylococcal bacteriophage ??H5 endolysin active against Staphylococcus aureus in milk. *Int. J. Food Microbiol.* **128,** 212–218 (2008).
397. Takáč, M., Witte, A. & Bläsi, U. Functional analysis of the lysis genes of Staphylococcus aureus phage P68 in Escherichia coli. *Microbiology* **151,** 2331–2342 (2005).
398. Navarre, W. W., Ton-That, H., Faull, K. F. & Schneewind, O. Multiple

enzymatic activities of the murein hydrolase from staphylococcal phage phi11. Identification of a D-alanyl-glycine endopeptidase activity. *J. Biol. Chem.* **274,** 15847–56 (1999).

399. Loessner, M. J., Gaeng, S., Wendlinger, G., Maier, S. K. & Scherer, S. The two-component lysis system of Staphylococcus aureus bacteriophage Twort: A large TTG-start holin and an associated amidase endolysin. *FEMS Microbiol. Lett.* **162,** 265–274 (1998).
400. Kong, M. & Ryu, S. Bacteriophage PBC1 and Its Endolysin as an Antimicrobial Agent against Bacillus cereus. doi:10.1128/AEM.03485-14
401. Szczesny, A. & Martirosian, G. Epidemiology of Clostridium difficile infection. *Przegl. Epidemiol.* **56,** 49–56 (2002).
402. Lawley, T. D. *et al.* Use of purified clostridium difficile spores to facilitate evaluation of health care disinfection regimens. *Appl. Environ. Microbiol.* **76,** 6895–6900 (2010).
403. Dunne, M. *et al.* The CD27L and CTP1L Endolysins Targeting Clostridia Contain a Built-in Trigger and Release Factor. *PLoS Pathog.* **10,** (2014).
404. Mayer, M. J., Garefalaki, V., Spoerl, R., Narbad, A. & Meijers, R. Structure-based modification of a Clostridium difficile-targeting endolysin affects activity and host range. *J. Bacteriol.* **193,** 5477–5486 (2011).
405. Wang, F. J. & Ripley, L. S. The spectrum of acridine resistant mutants of bacteriophage T4 reveals cryptic effects of the tsL141 DNA polymerase allele on spontaneous mutagenesis. *Genetics* **148,** 1655–1665 (1998).
406. Kellenberger, E. & Séchaud, J. Electron microscopical studies of phage multiplication. II. Production of phage-related structures during multiplication of phages T2 and T4. *Virology* **3,** 256–274 (1957).
407. Silver, S. Acridine sensitivity of bacteriophage T2: A virus gene affecting cell permeability. *J. Mol. Biol.* **29,** 191–202 (1967).
408. Koch, R. E. & Drake, J. W. Ligase-defective bacteriophage T4. I. Effects on mutation rates. *J Virol* **11,** 35–40 (1973).
409. Hobbs, L. J. & Nossal, N. G. Either bacteriophage T4 RNase H or Escherichia coli DNA polymerase I is essential for phage replication. *J. Bacteriol.* **178,** 6772–6777 (1996).
410. Reha-Krantz, L. J. & Lambert, J. K. J. Structure-function studies of the bacteriophage T4 DNA polymerase. Isolation of a novel suppressor

mutant. *J. Mol. Biol.* **186,** 505–514 (1985).

411. Elowitz, M. B. & Leibier, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403,** 335–338 (2000).
412. Gaeng, S., Scherer, S., Neve, H. & Loessner, M. J. Gene Cloning and Expression and Secretion of Listeria monocytogenes Bacteriophage-Lytic Enzymes in Lactococcus lactis. *Appl. Environ. Microbiol.* **66,** 2951–2958 (2000).
413. Loessner, M. J., Schneider, A. & Scherer, S. Modified Listeria bacteriophage lysin genes (ply) allow efficient overexpression and one-step purification of biochemically active fusion proteins. *Appl. Environ. Microbiol.* **62,** 3057–3060 (1996).
414. Son, B. *et al.* Characterization of LysB4, an endolysin from the Bacillus cereus-infecting bacteriophage B4. *BMC Microbiol.* **12,** (2012).
415. Lim, J. A., Shin, H., Kang, D. H. & Ryu, S. Characterization of endolysin from a Salmonella Typhimurium-infecting bacteriophage SPN1S. *Res. Microbiol.* **163,** 233–241 (2012).
416. Oliveira, H. *et al.* A thermostable salmonella phage endolysin, Lys68, with broad bactericidal properties against gram-negative pathogens in presence of weak acids. *PLoS One* **9,** (2014).
417. Calamita, H. G., Ehringer, W. D., Koch, A. L. & Doyle, R. J. Evidence that the cell wall of Bacillus subtilis is protonated during respiration. *Proc. Natl. Acad. Sci.* **98,** 15260–15263 (2001).
418. Goodell, E. W. and L. R. and T. A. Suppression of lytic effect of beta lactams on Escherichia coli and other bacteria. *Proc. Natl. Acad. Sci.* **73,** 3293–3297 (1976).
419. Zhang, W. *et al.* Characterization of Enterococcus faecalis phage IME-EF1 and its endolysin. *PLoS One* **8,** (2013).
420. Yuan, Y., Peng, Q. & Gao, M. Characteristics of a broad lytic spectrum endolysin from phage BtCS33 of Bacillus thuringiensis. *BMC Microbiol.* **12,** (2012).
421. Son, J. S. *et al.* Complete genome sequence of a newly isolated lytic bacteriophage, EFAP-1 of Enterococcus faecalis, and antibacterial activity of its endolysin EFAL-1. *J. Appl. Microbiol.* **108,** 1769–1779 (2010).
422. Lai, M. J. *et al.* Antibacterial activity of Acinetobacter baumannii phage

ϕAB2 (LysAB2) against both Gram-positive and Gram-negative bacteria. *Appl. Microbiol. Biotechnol.* **90,** 529–539 (2011).

423. Orito, Y., Morita, M., Hori, K., Unno, H. & Tanji, Y. Bacillus amyloliquefaciens phage endolysin can enhance permeability of Pseudomonas aeruginosa outer membrane and induce cell lysis. *Appl. Microbiol. Biotechnol.* **65,** 105–109 (2004).
424. Düring, K., Porsch, P., Mahn, A., Brinkmann, O. & Gieffers, W. The non-enzymatic microbicidal activity of lysozymes. *FEBS Lett.* **449,** 93–100 (1999).
425. Silva, M. T. & Silva Pestana, N. T. The in vivo extracellular life of facultative intracellular bacterial parasites: Role in pathogenesis. *Immunobiology* **218,** 325–337 (2013).
426. Armstrong, J. A. & Hart, P. D. Response of cultured macrophages to Mycobacterium tuberculosis, with observations on fusion of lysosomes with phagosomes. *J. Exp. Med.* **134,** 713–40 (1971).
427. Ibarra, J. A. & Steele-Mortimer, O. Salmonella - the ultimate insider. Salmonella virulence factors that modulate intracellular survival. *Cellular Microbiology* **11,** 1579–1586 (2009).
428. Dikshit, N. *et al.* Intracellular Uropathogenic E. coli Exploits Host Rab35 for Iron Acquisition and Survival within Urinary Bladder Cells. *PLoS Pathog.* **11,** (2015).
429. Monack, D. M., Bouley, D. M. & Falkow, S. Salmonella typhimurium Persists within Macrophages in the Mesenteric Lymph Nodes of Chronically Infected Nramp1 \r. *J. Exp. Med.* **199,** 231–241 (2004).
430. Winstanley, C., O'Brien, S. & Brockhurst, M. A. Pseudomonas aeruginosa Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. *Trends in Microbiology* **24,** 327–337 (2016).
431. Kamaruzzaman, N. F., Kendall, S. & Good, L. Targeting the hard to reach: challenges and novel strategies in the treatment of intracellular bacterial infections. *British Journal of Pharmacology* **174,** 2225–2236 (2017).
432. Yosef, I., Edgar, R. & Qimron, U. Phenotypic heterogeneity in a bacteriophage population only appears as stress - induced mutagenesis. *Curr. Genet.* 2–4 (2016). doi:10.1007/s00294-016-0606-2

433. Laakkonen, P. & Vuorinen, K. Homing peptides as targeted delivery vehicles. *Integr. Biol* **2,** 326–337 (2010).
434. Smith, G. P. Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science (80-.).* **228,** 1315–1317 (1985).
435. Arap, W. *et al.* Targeting the prostate for destruction through a vascular address. *Proc. Natl. Acad. Sci.* **99,** 1527–1531 (2002).
436. Zhang, L., Hoffman, J. A. & Ruoslahti, E. Molecular profiling of heart endothelial cells. *Circulation* **112,** 1601–1611 (2005).
437. Pasqualini, R., Koivunen, E. & Ruoslahti, E. αv Integrins as Receptors for Tumor Targeting by Circulating Ligands. *Nat. Biotechnol.* **15,** 542–546 (1997).
438. Porkka, K., Laakkonen, P., Hoffman, J. A., Bernasconi, M. & Ruoslahti, E. A fragment of the HMGN2 protein homes to the nuclei of tumor cells and tumor endothelial cells in vivo. *Proc. Natl. Acad. Sci.* **99,** 7444–7449 (2002).
439. Laakkonen, P. *et al.* Antitumor activity of a homing peptide that targets tumor lymphatics and tumor cells. *Proc. Natl. Acad. Sci.* **101,** 9381–9386 (2004).
440. Zhang, L., Giraudo, E., Hoffman, J. A., Hanahan, D. & Ruoslahti, E. Lymphatic zip codes in premalignant lesions and tumors. *Cancer Res.* **66,** 5696–5706 (2006).
441. Rajotte, D. & Ruoslahti, E. Membrane Dipeptidase Is the Receptor for a Lung-targeting Peptide Identified by in Vivo Phage Display*.
442. Samoylova, T. I. & Smith, B. F. Elucidation of muscle-binding peptides by phage display screening. *Muscle and Nerve* **22,** 460–466 (1999).
443. Kolonin, M. G., Saha, P. K., Chan, L., Pasqualini, R. & Arap, W. Reversal of obesity by targeted ablation of adipose tissue. *Nat. Med.* **10,** 625–632 (2004).
444. Ellerby, H. M. *et al.* Anti-cancer activity of targeted pro-apoptotic peptides. *Nat. Med.* **5,** 1032–1038 (1999).
445. Corti, A., Curnis, F., Rossoni, G., Marcucci, F. & Gregorc, V. Peptide-mediated targeting of cytokines to tumor vasculature: The NGR-hTNF example. *BioDrugs* **27,** 591–603 (2013).

446. Ruoslahti, E. Peptides as targeting elements and tissue penetration devices for nanoparticles. *Advanced Materials* **24,** 3747–3756 (2012).
447. Bowers, R. M. *et al.* Sources of bacteria in outdoor air across cities in the midwestern United States. *Appl. Environ. Microbiol.* **77,** 6350–6356 (2011).
448. Wilson, L. G. Commentary: Medicine, population, and tuberculosis. *Int. J. Epidemiol.* **34,** 521–524 (2005).
449. Kennedy, W. A. *et al.* Incidence of bacterial meningitis in Asia using enhanced CSF testing: Polymerase chain reaction, latex agglutination and culture. *Epidemiol. Infect.* **135,** 1217–1226 (2007).
450. Sethi, S. & Murphy, T. F. Infection in the pathogenesis and course of chronic obstructive pulmonary disease. *N. Engl. J. Med.* **359,** 2355–65 (2008).
451. Sche, P. P., McKenzie, K. M., White, J. D. & Austin, D. J. Display cloning: functional identification of natural product receptors using cDNA-phage display. *Chem. Biol.* **6,** 707–716 (1999).
452. Sheu, T.-J. *et al.* A phage display technique identifies a novel regulator of cell differentiation. *J. Biol. Chem.* **278,** 438–443 (2003).
453. Krajcikova, D. & Hartley, R. W. A new member of the bacterial ribonuclease inhibitor family from Saccharopolyspora erythraea. *FEBS Lett.* **557,** 164–168 (2004).
454. Fukunaga, K., Hatanaka, T., Ito, Y. & Taki, M. Gp10 based-thioetherification (10BASEd-T) on a displaying library peptide of bacteriophage T7. *Mol. Biosyst.* **9,** 2988 (2013).
455. Guo, F. *et al.* Capsid expansion mechanism of bacteriophage T7 revealed by multistate atomic models derived from cryo-EM reconstructions. *Proc. Natl. Acad. Sci.* **111,** E4606–E4614 (2014).
456. Condron, B. G., Atkins, J. F. & Gesteland, R. F. Frameshifting in gene 10 of bacteriophage T7. *J. Bacteriol.* **173,** 6998–7003 (1991).
457. Dunn, J. J., Studier, F. W. & Gottesman, M. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **166,** 477–535 (1983).
458. Fukunaga, K. & Taki, M. Practical tips for construction of custom peptide libraries and affinity selection by using commercially available phage

display cloning systems. *Journal of Nucleic Acids* **2012,** (2012).

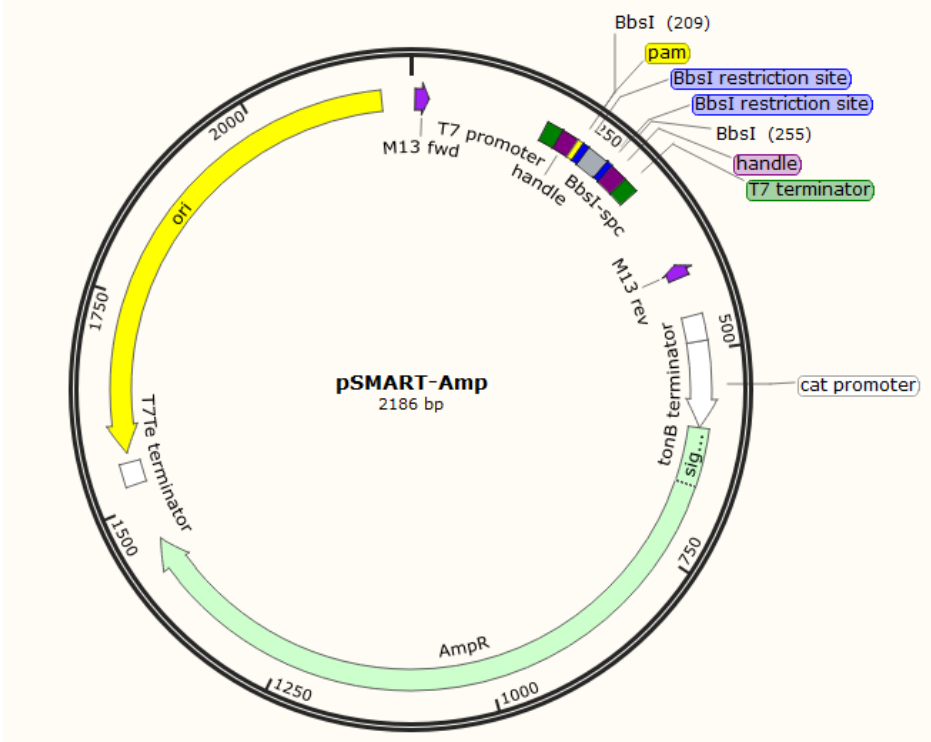
459. Slootweg, E. J. *et al.* Fluorescent T7 display phages obtained by translational frameshift. *Nucleic Acids Res.* **34,** (2006).
460. Akerman, M. E., Chan, W. C. W., Laakkonen, P., Bhatia, S. N. & Ruoslahti, E. Nanocrystal targeting in vivo. *Proc. Natl. Acad. Sci.* **99,** 12617–12621 (2002).
461. Blanco, M. A. *et al.* Identification of staphylococcal nuclease domain-containing 1 (SND1) as a metadherin-interacting protein with metastasis-promoting functions. *J. Biol. Chem.* **286,** 19982–19992 (2011).
462. Rajotte, D. & Ruoslahti, E. Membrane dipeptidase is the receptor for a lung-targeting peptide identified by in vivo phage display. *J. Biol. Chem.* **274,** 11593–11598 (1999).
463. Brown, D. M. & Ruoslahti, E. Metadherin, a cell surface protein in breast tumors that mediates lung metastasis. *Cancer Cell* **5,** 365–374 (2004).
464. Rajotte, D. *et al.* Molecular Heterogeneity of the Vascular Endothelium Revealed by In Vivo Phage Display. *J. Clin. Invest* **102,** 430–437 (1998).
465. Lieber, M., Smith, B., Szakal, a, Nelson-Rees, W. & Todaro, G. A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. *Int. J. Cancer* **17,** 62–70 (1976).
466. Giard, D. J. *et al.* In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J. Natl. Cancer Inst.* **51,** 1417–1423 (1973).
467. Fang, R. & Aust, A. E. Induction of ferritin synthesis in human lung epithelial cells treated with crocidolite asbestos. *Arch. Biochem. Biophys.* **340,** 369–375 (1997).
468. Wallace, W. a & Howie, S. E. Upregulation of tenascin and TGFbeta production in a type II alveolar epithelial cell line by antibody against a pulmonary auto-antigen. *J. Pathol.* **195,** 251–256 (2001).
469. Serwer, P. A metrizamide-impermeable capsid in the DNA packaging pathway of bacteriophage T7. *J. Mol. Biol.* **138,** 65–91 (1980).
470. King, A. *et al.* Tumor-homing peptides as tools for targeted delivery of payloads to the placenta. *Sci. Adv.* **2,** (2016).
471. Dunstan, R. A., Weston, T. A., Schlosser, K. & Bailey, B. Bacteriophage Transcytosis Provides a. *Am. socity Microbiol.* **8,** 1–14 (2017).

472. Karmali, P. P. *et al.* Targeting of albumin-embedded paclitaxel nanoparticles to tumors. *Nanomedicine Nanotechnology, Biol. Med.* **5,** 73–82 (2009).
473. Uchida, M. *et al.* Protein cage nanoparticles bearing the LyP-1 peptide for enhanced imaging of macrophage-rich vascular lesions. *ACS Nano* **5,** 2493–2502 (2011).
474. Soda, Y. *et al.* Transdifferentiation of glioblastoma cells into vascular endothelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 4274–80 (2011).
475. Leon-Sicairos, N. *et al.* Strategies of intracellular pathogens for obtaining iron from the environment. *Biomed Res. Int.* **2015,** (2015).
476. Winer, I. *et al.* F3-targeted cisplatin-hydrogel nanoparticles as an effective therapeutic that targets both murine and human ovarian tumor endothelial cells in vivo. *Cancer Res.* **70,** 8674–83 (2010).
477. Wartosch, L., Bright, N. A. & Luzio, J. P. Lysosomes. *Curr. Biol.* **25,** R315–R316 (2015).
478. Ivanenkov, V. V., Felici, F. & Menon, A. G. Uptake and intracellular fate of phage display vectors in mammalian cells. *Biochim. Biophys. Acta - Mol. Cell Res.* **1448,** 450–462 (1999).
479. Xu, H. *et al.* Sequence determinants of improved CRISPR sgRNA design. *Genome Res.* **25,** 1147–1157 (2015).
480. Cong, L., Ran, F. A., Cox, D., Lin, S. & Barretto, R. Multiplex genome engineering using CRISPR/Cas systems. *Mult. genome Eng. using Cris. Syst.* **339,** 819–823 (2013).
481. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343,** 80–4 (2014).
482. Chen, X., Zaro, J. L. & Shen, W. C. Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews* **65,** 1357–1369 (2013).
483. Pyne, M. E., Bruder, M. R., Moo-Young, M., Chung, D. A. & Chou, C. P. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in Clostridium. *Sci. Rep.* **6,** (2016).

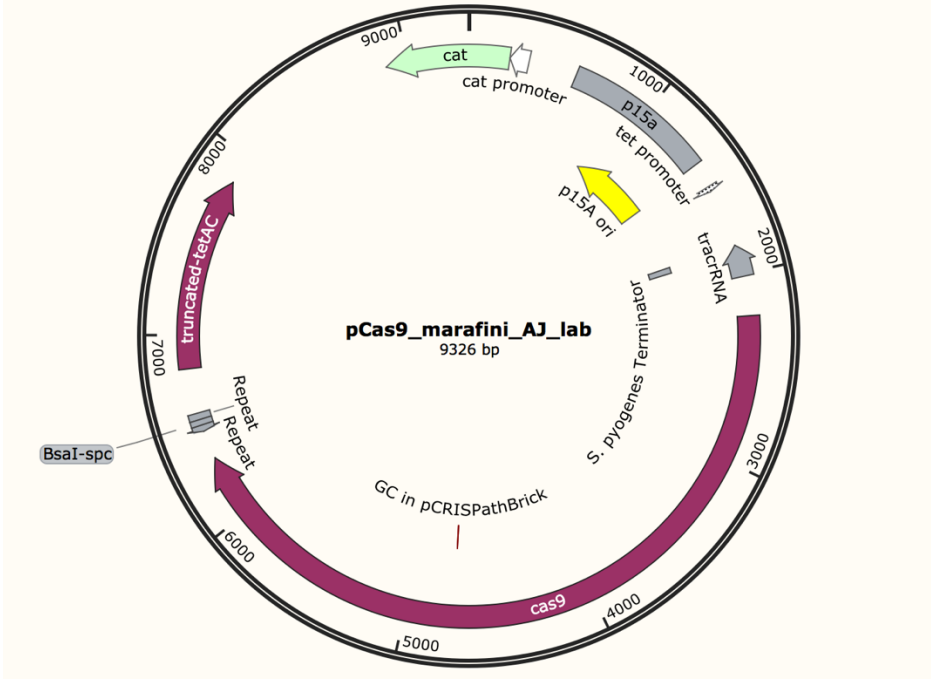
Appendices

Appendix 1

IDT synthesised pSMART vector (pAG1) used to generate pAG2-pAG11



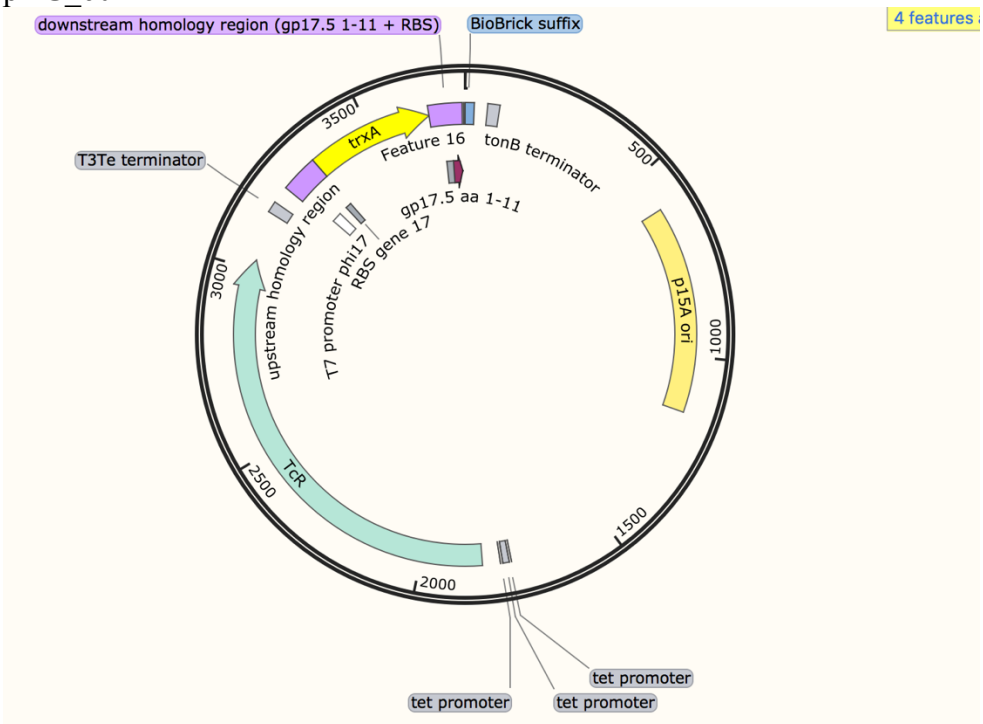
pCas9 vector



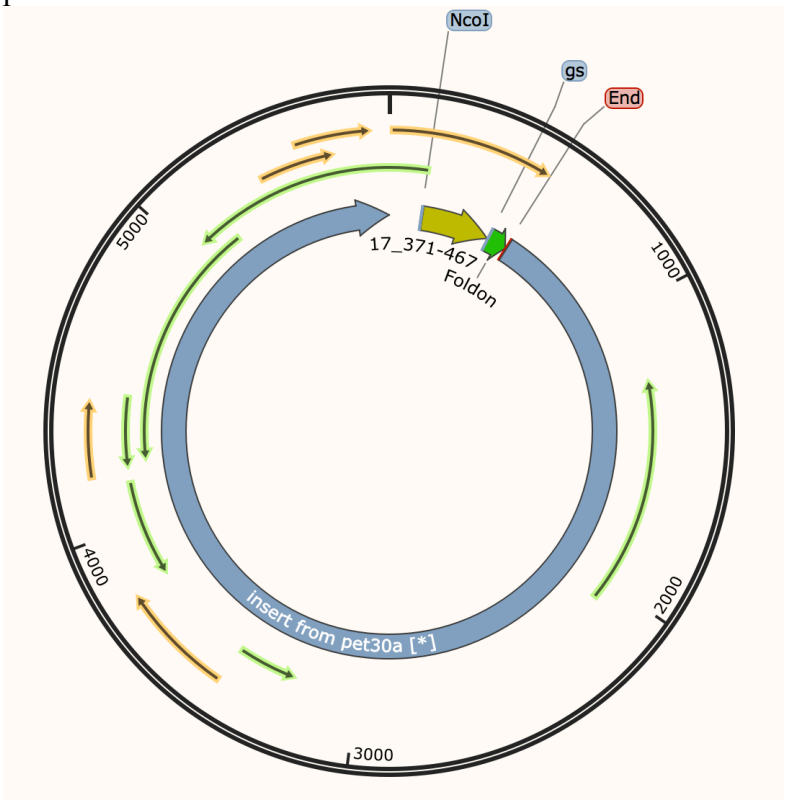
pSEVA551


NheI site
F24
7000
T7 terminator
T0
gp17
tetA
1000
nucleotide correction (now gene...
Fragment A
6000
spacer
spacer
RBS
T1
T7 promoter
nucleotide correction
oriT
2000
placUV5
5000
repB
repA
Fragment B
3000
4000

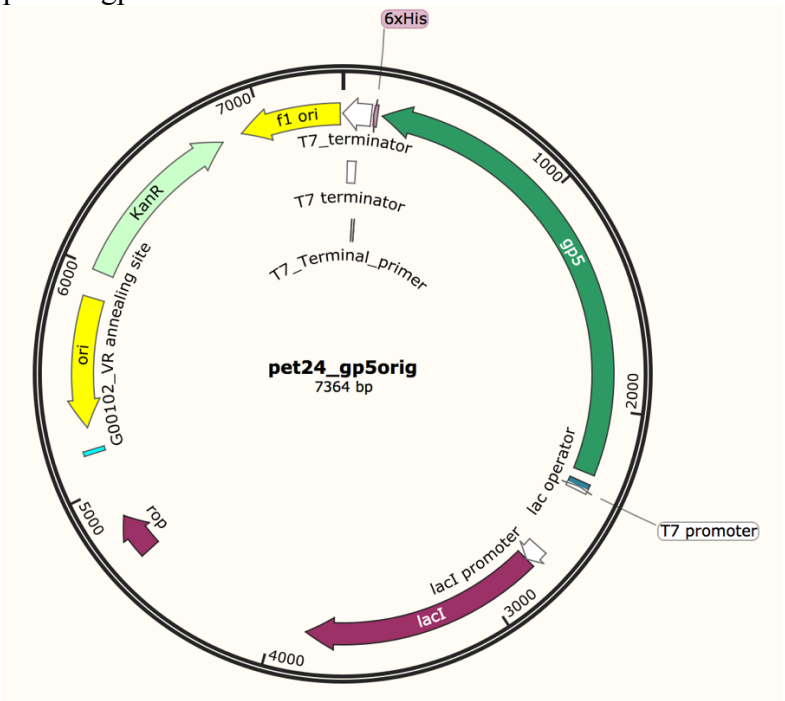
pAG_60


downstream homology region (gp17.5 1-11 + RBS)
BioBrick suffix
4 features
3500
trxA
Feature 16
tonB terminator
500
T3Te terminator
gp17.5 aa 1-11
T7 promoter phi17
RBS gene 17
upstream homology region
3000
p15A ori
1000
TcR
2500
1500
2000
tet promoter
tet promoter
tet promoter

pET30a+



pET24-gp5



Appendix 2

Name	Description	Sequence
HRg17_MtdC_HRg17.5		
AG MTD 001	g17_371-467_foldon_MtdC FWR	gtgacagcttcgttgcgaagAAGTTTCGCCCGGCTGCGCT
AG MTD 002	pSB6A1_yepphi_g_block REV	AGCGCAGCCGGGCGAAACTTcttcgcaacgaagctgtcac
AG MTD 003	g17_371-467_foldon_MtdC REV	ctttccttgtgatttaccaatTAACAAAAAACCCCTAGCCGCCC
AG MTD 004	pSB6A1_yepphi_gblock FWD	GGCTAGGGGTTTTTTGTTAattggtaaatcacaaggaaagacg
AG MTD 005	HR MTDC FWD_1	atcacgaggccctttcgtcttc
AG MTD 006	HR MTD FWD_2	CATGACGCTGGTTGCCGGC
AG MTD 007	HR MTD REV_1	CCCCTAGCCGCCCGATAAG
AG MTD 008	HR MTDC REV_2	gatgcgtccggcgtagagc
HRg17_MtdC_trx_HRg17.5		
AG MTD 009	pSB6A1 + yep-phi g-block FWD	gactttaacaacgaattgatACTAGtagcggccgctgcag
AG MTD 010	pSB6A1 + yep-phi g-block REV	AGCGCAGCCGGGCGAAACTTcttcgcaacgaagctgtcac
AG MTD 011	gblock_His_g17_foldon_MtdC FWD	gtgacagcttcgttgcgaagAAGTTTCGCCCGGCTGCGCT
AG MTD 012	gblock_His_g17_foldon_MtdC REV	tgaatgattttatcgctcatACAAAAAACCCCTAGCCGCCC
AG MTD 013	trxA gp17 homology (flanking gp17) no bioBrick sites g-block FWD	GGCGGCTAGGGGTTTTTTGTatgagcgataaaatcattcacc
AG MTD 014	trxA gp17 homology (flanking gp17) no bioBrick sites g-block REV	ctgcagcggccgctaCTAGTatcaattcgttgttaaagtctaatg
AG MTD 015	HR MTD TRX FWD	ccgacgtgctgaaagctgatgg
AG MTD 016	HR MTD TRX REV	gccaggttggcgtccaggaa
	*these primers are not valid since they produce an insert with no rbs before trxA; the MTD 037-MTD 042 primers should be used instead	
Gibson_InTrans_COg17_MtdC		
AG MTD 017	gblock_His_g17_foldon_MtdC REV	gcgggctaggggttcgagtaACAAAAAACCCCTAGCCGCCC
AG MTD 018	gblock_His_g17_foldon_MtdC FWD	GAGATTCTTTTGTCGCTAAAAAGTTTCGCCCGGCTGCGCT
AG MTD 019	pSEVA551_COgp17 REV	AGCGCAGCCGGGCGAAACTTTTTAGCGACAAAAGAATCTCGCA
AG MTD 020	pSEVA551_COgp17 FWD	GGCGGCTAGGGGTTTTTTGTtactcgaacccctagcccgc
Gibson_InTrans_COg17_MtdC_RBS_Trx		
AG MTD 021	gblock_His_g17_foldon_MtdC FWD	GAGATTCTTTTGTCGCTAAAAAGTTTCGCCCGGCTGCGCT
AG MTD 022	gblock_His_g17_foldon_MtdC REV	tgacctccttaaagtaaatcCGGTTGCCTTGGCGGGGCCT
AG MTD 023	pSEVA551_COg17 REV	AGCGCAGCCGGGCGAAACTTTTTAGCGACAAAAGAATCTCGCA
AG MTD 024	gblock_trxA FWD	AGGCCCCGCCAAGGCAACCGgatttactttaaggaggtcaaatgag
AG MTD 025	pSEVA551_COg17 FWD	acgccaacctggcttaatgatactcgaacccctagcccgc
AG MTD 026	gblock_trxA REV	gcgggctaggggttcgagtatcattaagccaggttggcgt
Gibson_InTrans_WTg17_MtdC		
AG MTD 027	His_foldon_MtdC REV	gtgacagcttcgttgcgaagAAGTTTCGCCCGGCTGCGCT
AG MTD 028	His_foldon_MtdC FWD	agaccgagcgttctgaacaaACAAAAAACCCCTAGCCGCCC
AG MTD 029	pSB6A1_WTg17 REV	GGCGGCTAGGGGTTTTTTGTttgttcagaacgctcggtct
AG MTD 030	pSB6A1_WTg17 FWD	AGCGCAGCCGGGCGAAACTTcttcgcaacgaagctgtcac
Gibson_InTrans_WTg17_MtdC_RBS_Trx		
AG MTD 031	His_foldon_MtdC REV	gtgacagcttcgttgcgaagAAGTTTCGCCCGGCTGCGCT
AG MTD 032	His_foldon_MtdC FWD	tgacctccttaaagtaaatcCGGTTGCCTTGGCGGGGCCT
AG MTD 033	gblock_trxA FWD	agaccgagcgttctgaacaattaagccaggttggcgtcca
AG MTD 034	pSB6A1_WTg17 REV	tggacgccaacctggcttaattgttcagaacgctcggtct
AG MTD 035	gblock_trxA REV	AGGCCCCGCCAAGGCAACCGgatttactttaaggaggtcaaatgagc
AG MTD 036	pSB6A1_WTg17 FWD	AGCGCAGCCGGGCGAAACTTcttcgcaacgaagctgtcac
HRg17_MtdC_rbs_trx_HRg17.5		
AG MTD 037	gblock_His_g17_foldon_MtdC FWD	gtgacagcttcgttgcgaagAAGTTTCGCCCGGCTGCGCT
AG MTD 038	gblock_His_g17_foldon_MtdC REV	tgacctccttaaagtaaatcCGGTTGCCTTGGCGGGGCCT
AG MTD 039	pSB6A1+yrp-phi g-block REV	AGCGCAGCCGGGCGAAACTTcttcgcaacgaagctgtcac
AG MTD 040	trxAgp17 homology g-block FWD	AGGCCCCGCCAAGGCAACCGgatttactttaaggaggtcaaatgag
AG MTD 041	pSB6A1+yep-phi g block FWD	acgccaacctggcttaatgattggtaaatcacaaggaaagacg
AG MTD 042	trxA gp17 homology g-block REV	ctttccttgtgatttaccaatcattaagccaggttggcgt
AG 043	Paj57(trxA-gp17 homology) FWD	TtATGGTaGAGAAtGAGTAAgatttactttaaggaggtcaaatga
AG 044	Paj57(trx-gp17 homology) REV	ATTTCCCCTCTTTCTCTAGTtaagagactacagggagaaga
AG 045	pSEVA551+COg17 FWD	cttctccctgtagtctcttaACTAGAGAAAGAGGGGAAATAC
AG 046	pSEVA551+COg17 REV	tgacctccttaaagtaaatcTTACTCaTTCTCtACCATaATTGC
AG 047	pAJ57 FWD 1	gaaaaggaatattcagcaatttgcccg
AG 048	pAJ57 FWD 2	gaaatacggtatccgtggcatccc
AG 049	pAJ57 FWD 3	cacgacggggagtcaggcaac
AG 050	pAJ57 REV 1	cctccttgagagtccatccgtg
AG 051	pAJ57 REV 2	cgctcagtggaacgaaaactcac
AG 052	pAJ57 REV 3	gaaagcggtcctcgccgaaaatg
AG 061	g17_rev	ccccaccgccagcactac
AG 062	g17_rev	cgggggccagtacggaagaa
AG 063	g17 FRW (before gp17)	tcggctggctttgtggctaacg
AG 064	g17 REV (after gp17)	acctccttgagagtccatccgtgg
AG065	Michail	TCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTAGtcagtggcaaatcgcccaatt
AG066	Michail	tgccggataatacgactcactataggATAAACAGAGGAGATATCACatgatcgtttctgacatcgaag
AG074	Mtd_g_block02.02.16 FWD	gtgacagcttcgttgcgaagACGGCTTTCATCAAGGCCGA
AG075	Mtd_g_block02.02.16 REV	tgacctccttaaagtaaatcacaaaaaaccccctagccgcc
AG076	pSB6A1 + yep-phi g-block FWD	acgccaacctggcttaatgattggtaaatcacaaggaaagacg
AG077	pSB6A1 + yep-phi g-block REV	TCGGCCTTGATGAAAGCCGTcttcgcaacgaagctgtcac
AG078	trxA gp17 homology (flanking gp17)	ggcggctagggggtttttgtgatttactttaaggaggtcaaatgag
AG079	trxA gp17 homology (flanking gp17)	ctttccttgtgatttaccaatcattaagccaggttggcgt
AG080	Mtd_g_block02.02.16 FWD	gtgacagcttcgttgcgaagAACGAATACAGCCTGTGGGA
AG081	Mtd_g_block02.02.16 REV	tgacctccttaaagtaaatcacaaaaaaccccctagccgcc
AG082	pSB6A1 + yep-phi g-block FWD	acgccaacctggcttaatgattggtaaatcacaaggaaagacg
AG083	pSB6A1 + yep-phi g-block REV	TCCCACAGGCTGTATTCGTTcttcgcaacgaagctgtcac
AG084	Mtd_g_block02.02.16 FWD	gtgacagcttcgttgcgaagATCAAGTTtCGCCCGGCTGC
AG085	Mtd_g_block02.02.16 REV	tgacctccttaaagtaaatcacaaaaaaccccctagccgcc
AG086	pSB6A1 + yep-phi g-block FWD	acgccaacctggcttaatgattggtaaatcacaaggaaagacg
AG087	pSB6A1 + yep-phi g-block REV	GCAGCCGGGCGaAACTTGATcttcgcaacgaagctgtcac
AG088	plasmid: pCDF-1b_Qimrons_pWUR400, pRSF-1b_Qimrons_pWUR397, pACYCD	TCGACGCTCTCCCTTATGCGAC
AG089	plasmid: pCDF-1b_Qimrons_pWUR400, pRSF-1b_Qimrons_pWUR397, pACYCD	GAAGCAGTGTGACCGTGTGC

Name	Description	Sequence
AG090	plasmid: pCDF-1b_Qimrons_pWUR400	GTCTAACAATTCGTTCAAGCCG
AG091	plasmid: pRSF-1b_Qimrons_pWUR397	GGAGACCCGGTCGTCAGC
AG092	plasmid: pACYCDuet-1_Qimron's_pWUR478	GCTAGTTATTGCTCAGCGG
AG093	plasmid: pCAS9_Marafini_JaramilloLab	cagacaggcggattctccaagg
AG094	plasmid: pCAS9_Marafini_JaramilloLab	ccgcaaggaatggtgcatgc
AG095	plasmid: forward primer for pET30a+-MVR-gp17-foldon467, pET30a+-MVR-gp17-mtd_171	gtatgaaagaaaccgctgctgc
AG096	plasmid: reverse primer for pET30a+-MVR-gp17-foldon467, pET30a+-MVR-gp17-mtd_171	ctttcgggctttgttagcagc
AG097	plasmid: pSB6A1_HRgp17_Mtd_HRgp17, pSB6A1_HRgp17_Mtd_trx_HRgp17, pSB6A1-WT	caccagcgtttctgggtgagc
AG098	plasmid: pSB6A1_HRgp17_Mtd_HRgp17, pSB6A1_HRgp17_Mtd_trx_HRgp17	ctcaagggcatcggtcgacg
AG099	plasmid: pSEVA551-cogp17 (S189I) _Mtd_trx_intrans	ctagcgcagcgaatagaccag
AG100	plasmid: pSEVA551-cogp17 (S189I) _Mtd_trx_intrans	cacggtgcctgactgcgtta
AG101	plasmid: HR_COg17_trx_HR	gccagtgagttgattgctacgt
AG102	plasmid: HR_COg17_trx_HR	caaaagcctccggtcggagg
AG103	plasmid: HR_COg17_trx_HR	gccagtgagttgattgctacgt
AG104		gcagcggccgctactagtt
AG105	to make QNQ_trxA(AKOS), HR_trxA_2Stopcodons FWD	aattcgcggccgcttCTAGAtgttaacttgagggagcgta
AG106	to make QNQ_trxA(AKOS), HR_trxA_2Stopcodons REV	ctgcagcggccgctaCTAGTatcaattcgttgttaaagtctaatg
AG107	to make QNQ_trxA(AKOS), pSB6A1_yepphi FWD	gactttaacaacgaattgatACTAGtagcggccgctgcag
AG108	to make QNQ_trxA(AKOS), pSB6A1_yepphi REV	tacgctccctcaagttaacaTCTAGaagcggccgcgaatt
AG109	to make QNQ_gp17 & stop_N_stop_gp17 (AKOS), HR_QNQ_COgp17_RBS_trxA_HR FWD	tcaagaattcgcggccgcttgataacaacaatgactgtacct
AG110	to make QNQ_gp17 & stop_N_stop_gp17 (AKOS), HR_QNQ_COgp17_RBS_trxA_HR REV	ctgcagcggccgctaCTAGTatcaattcgttgttaaagtctaatg
AG111	to make QNQ_gp17 & stop_N_stop_gp17 (AKOS), pSB6A1_yepphi FWD	gactttaacaacgaattgatACTAGtagcggccgctgcag
AG112	to make QNQ_gp17 & stop_N_stop_gp17 (AKOS)), pSB6A1_yepphi REV	gtacagtcattgttgttatcaagcggccgcgaattcttga
	Gibson to make HRg1_mKate2_trxA: pAJ57 (trxA-gp17 homology) FWD	ataaacagaggagatatcacatgagcgataaaatcattcacctg
	Gibson to make HRg1_mKate2_trxA:pAJ57 (trxA-gp17 homology) REV	tgagtcgtattgatttggcGtcattaagccaggttggcgt
	Gibson to make HRg1_mKate2_trxA:pIDTSmartAmp_T7HRmini1 FWD	acgccaacctggcttaatgaCgccaaatcaatacgactcact
	Gibson to make HRg1_mKate2_trxA:pIDTSmartAmp_T7HRmini1 REV	GTTTCCTGTGTGACTCTAGCttacgcgaacgcgaagtccg
	Gibson to make HRg1_mKate2_trxA:pLtet-mKate2 FWD	cggacttcgcgttcgcgtaaGCTAGAGTCACACAGGAAAC
	Gibson to make HRg1_mKate2_trxA:pLtet-mKate2 REV	GTGATATCTCCTCTGTTTATTTATTAACGGTGACCCAGCT
	Gibson to make HRg6.3_mKate2_trxA:pAJ57 (trxA-gp17 homology) FWD	ataaacagaggagatatcacatgagcgataaaatcattcacctg
	Gibson to make HRg6.3_mKate2_trxA:pAJ57 (trxA-gp17 homology) REV	atttagggaccttgagtttAtcattaagccaggttggcgt
	Gibson to make HRg6.3_mKate2_trxA:pIDTSmartAmp_T7HRmini2 FWD	acgccaacctggcttaatgaTaaactcaaggtccctaaattaatac
	Gibson to make HRg6.3_mKate2_trxA:pIDTSmartAmp_T7HRmini2 REV	GTTTCCTGTGTGACTCTAGCtcacttagtgtcgtgtaacg
	Gibson to make HRg6.3_mKate2_trxA:pLtet-mKate2 FWD	cgttacacgacactaagtgaGCTAGAGTCACACAGGAAAC
	Gibson to make HRg6.3_mKate2_trxA:pLtet-mKate2 REV	GTGATATCTCCTCTGTTTATTTATTAACGGTGACCCAGCT
	Gibson to make HRg17_mKate2_trxA:pAJ57 (trxA-gp17 homology) FWD	ataaacagaggagatatcacatgagcgataaaatcattcacctg
	Gibson to make HRg17_mKate2_trxA:pAJ57 (trxA-gp17 homology) REV	ctttccttgtgatttaccaAtcattaagccaggttggcgt
	Gibson to make HRg17_mKate2_trxA:pIDTSmartAmp_T7HRmini3 FWD	acgccaacctggcttaatgaTtggtaaatcacaaggaaagac
	Gibson to make HRg17_mKate2_trxA:pIDTSmartAmp_T7HRmini3 REV	GTTTCCTGTGTGACTCTAGCttactcgttctccaccatga
	Gibson to make HRg17_mKate2_trxA:pLtet-mKate2 FWD	tcatggtggagaacgagtaaGCTAGAGTCACACAGGAAAC
	Gibson to make HRg17_mKate2_trxA:pLtet-mKate2 REV	GTGATATCTCCTCTGTTTATTTATTAACGGTGACCCAGCT
	Gibson to make HR_deltagp17:pAJ57 (trxA-gp17 homology) FWD	ataaacagaggagatatcacatgagcgataaaatcattcacctg
	Gibson to make HR_deltagp17:pAJ57 (trxA-gp17 homology) REV	TTGATCAGTTCGCTTACCATttgacctccttaaagtaaatctaag
	Gibson to make HR_deltagp17:pLtet-mKate2 FWD	atttactttaaggaggtcaaATGGTAAGCGAACTGATCAA
	Gibson to make HR_deltagp17:pLtet-mKate2 REV	GTGATATCTCCTCTGTTTATTTATTAACGGTGACCCAGCT

AG113	primers for gRNAs for CRISPR type I/saved in LAB/CRISPR/Additional_Primers	accgtccttacgattaatacagactatcgctttgct
AG114		actcAGCAAAGCGATAGTCTGTATTAATCGTAAGGA
AG115		accgaaatattcacgctaacgggcgcctttacatga
AG116		actcTCATGTAAAGGCGCCCGTTAGCGTGAATATTT
AG117		accgcggtaacatccagttagtagtaaacggacaga
AG118		actcTCTGTCCGTTTACTACTAACTGGATGTTACCG
AG119		accgTTACTCGACGTAACTCGATGGTCGTGTAGCCA
AG120		actcTGGCTACACGACCATCGAGTTACGTCGAGTAA
AG121		accgTACAGTCATTGTTGTTATCTGACCCTCTACCA
AG122		actcTGGTAGAGGGTCAGATAACAACAATGACTGTA
AG123		accgtgtggtctggtagtgctggcggtggggtaagt
AG124		actcACTTACCCCACCGCCAGCACTACCAGACCACA
AG125		accgatctccgcttccgcaatatctggattaagtgt
AG126		actcACACTTAATCCAGATATTGCGGAAGCGGAGAT
AG127		accgcgtggactcaggtgtggtctggtagtgctggc
AG128		actcGCCAGCACTACCAGACCACACCTGAGTCCACG
AG129		accgCTATGAAGTAGATTCCATCGGGGCCAGTACGG
AG130		actcCCGTACTGGCCCCGATGGAATCTACTTCATAG
AG131		accgTACTGAACGACTGTCTGCAATATTCTTGAATC
AG132		actcGATTCAAGAATATTGCAGACAGTCGTTCAGTA
AG133	primers for gRNAs for CRISPR type II/saved in LAB/CRISPR/Additional_Primers	AAACgttccgctgcgcatcaatctg
AG134		AAAACAGATTGATGCGCAGCGGAAC
AG135		AAACgacgctacgaacacaaagcag
AG136		AAAACTGCTTTGTGTTCGTAGCGTC
AG137		AAACgagcagagttagcggatgctg
AG138		AAAACAGCATCCGCTAACTCTGCTC
AG139		AAACgagcctcattacggaactgta
AG140		AAAATACAGTTCCGTAATGAGGCTC
AG141		AAACggtaagtgtgactgtttcac
AG142		AAAAGTGAAACAGTCACACTTACC
AG143		AAACgatctccgcttccgcaatatc
AG144		AAAAGATATTGCGGAAGCGGAGATC
AG145		AAACGTAATCCAGATATTGCGGAAG
AG146		AAAACTTCCGCAATATCTGGATTAC
AG147		AAACGCGAGACCGTTGGAGTGTATT
AG148		AAAAAATACACTCCAACGGTCTCGC
AG149	Gibson primers to make construct to insert H5lysin into T4 (inside ac gene), name: Gibson_HR1_RBS_H5Lysin_HR2_T4, **Primer name: H5Lysin_RBS FWD**	aaagaggagaaatactagATGCAAGCAAAACTA
AG150	Gibson primers to make construct to insert H5lysin into T4 (inside ac gene), name: Gibson_HR1_RBS_H5Lysin_HR2_T4, **Primer name: H5Lysin_RBS REV**	TGCGAAAACACACTGCTAGCTTAACTGATTTCT
AG151	Gibson primers to make construct to insert H5lysin into T4 (inside ac gene), name: Gibson_HR1_RBS_H5Lysin_HR2_T4, **primer name: ac flaniking g_block FWD**	TATGGGGAGAAATCAGTTAAGCTAGCAGTGTGT
AG152	Gibson primers to make construct to insert H5lysin into T4 (inside ac gene), name: Gibson_HR1_RBS_H5Lysin_HR2_T4, **primer name: ac flaniking g_block REV**	ctagtatttctcctctttATCGATACAGATAGT
AG153	Gibson primers to make construct to insert H5lysin into T7 (inside ac gene), name: Gibson_HR1_trxA_RBS_H5Lysin_HR2_T7, **Primer name: H5Lysin_RBS FWD**	aaagaggagaaatactagATGCAAGCAAAACTA
AG154	Gibson primers to make construct to insert H5lysin into T7 (inside ac gene), name: Gibson_HR1_trxA_RBS_H5Lysin_HR2_T7, **Primer name: H5Lysin_RBS REV**	tgacctccttaaagtaaatcTTAACTGATTTCT
AG155	Gibson primers to make construct to insert H5lysin into T7 (inside ac gene), name: Gibson_HR1_trxA_RBS_H5Lysin_HR2_T7, **Primer name: HR1(end of gp17)_trxA_HR2(after gp17)(new g_bock ordered) FWD**	TATGGGGAGAAATCAGTTAAgatttactttaag
AG156	Gibson primers to make construct to insert H5lysin into T7 (inside ac gene), name: Gibson_HR1_trxA_RBS_H5Lysin_HR2_T7, **Primer name: HR1(end of gp17)_trxA_HR2(after gp17)(new g_bock ordered) REV**	ctagtatttctcctcttttactcgttctccac
AG157	primer to sequence g_block: HR1(end of gp17)_trxA_HR2(after gp17)	ggtggatggttacgattcc
AG158	primer to sequence g_block: HR1(end of gp17)_trxA_HR2(after gp17)	ggagcagccttaatcaattcg
AG159	Gibson primer to make HR plasmid to insert PBC1 lysin, amplifies g_block PBC1 FWD	GTTTTAACTATCTGTATCGATaaagaggagaaa
AG160	Gibson primer to make HR plasmid to insert PBC1 lysin, amplifies g_block PBC1 REV	TGCGAAAACACACTGCTAGCTTAGTATTCGATT
AG161	FWD	AGGAAGTAATCGAATACTAAGCTAGCAGTGTGT
AG162	REV	ATctagtatttctcctctttATCGATACAGATA
AG163	Gibson primer to make HR plasmid to insert CD27L lysin, amplifies g_block CD27L FWD	TGTTTTAACTATCTGTATCGATaaagaggagaa
AG164	Gibson primer to make HR plasmid to insert CD27L lysin, amplifies g_block CD27L REV	TGCGAAAACACACTGCTAGCCTAGCGGTTGATA
AG165	FWD	TGGATTTTATCAACCGCTAGGCTAGCAGTGTGT
AG166	REV	ATctagtatttctcctctttATCGATACAGATA
AG167	sequencing primer for PBC1 FWD	CAAGTTTACCGTCCATGCCGGGC
AG168	sequencing primer for CD27L FWD	GCATTACGGTGGGCCACAGC
AG169	sequencing primer for phoh FWD	CACGCGATGCTCGTGTTGCTATTCG
AG170	for sequencing after AC	ccaaggtccgctacaccacc
AG171	for sequencing before AC	CAGCTTGCGCGTACTATTGC
AG172	inside PBC1 lysin forward	CGCTTGGCGGAGGAGCTAGTGG
AG173	inside PBC1 lysin reverse	CACGCGAAATACTCGCGAGCGG
AG174	inside H5 forward	GCGCAACCTGGCGACATGGTTG
AG175	inside H5 reverse	CGTTTCCATGCACTTGCAACTGGC
AG176	inside CD27L frw	CTGAAGAGTGGTGCTTGCACTTCCG
AG177	inside CD27L rev	CAGCGGTAGGCTTGCTTGAGTTAAGG
AG178	AC_2 fwd	cactatgaggaagaatgtaagttgc
AG179	AC_2 rev	cgtttactcgtaatcatcattctgag
AG180	10B_HR_GFE-2 g block in a plsamid FWD	ctgccaccgctgagcaataaataaacagaggagatatcacatga
AG181	10B_HR_GFE-2 g block in a plsamid REV	atatctccttcttaaagttaTTAACAAGTCTCTAGCTCAAATCC
AG182	SEVA631_gp10 FWD AB plasmid for GFE-2	TTGAGCTAGAGACTTGTTAAtaactttaagaaggagatatacatatggc
AG183	SEVA631_gp10 REV AB plasmid for GFE-2	gtgatatctcctctgtttatttattgctcagcggtggcag
AG184	10B_metadherin in a plsmid FWD	ctgccaccgctgagcaataaataaacagaggagatatcacatga
AG185	10B_metadherin in a plsmid REV	atatctccttcttaaagttaTTAtttggatttcccagttgga
AG186	**SEVA631_gp10 FWD for the AB plasmid to be inserted in metadherin plasmid**	caactgggaaatccaaaTAAtaactttaagaaggagatatacatatggc
AG187	SEVA631_gp10 REV for the AB plasmid to be inserted in metadherin plasmid	gtgatatctcctctgtttatttattgctcagcggtggcag
AG188	10B_HR_GFE-1 in plasmid FWD	ctgccaccgctgagcaataaataaacagaggagatatcacatga
AG189	10B_HR_GFE-1 in plasmid REV	atatctccttcttaaagttaTTAGCATCTTTCGGGGCATT
AG190	SEVA631_gp10 FWD for the AB plasmid to be inserted in GFE-1	AATGCCCCGAAAGATGCTAAtaactttaagaaggagatatacatatggc
AG191	SEVA631_gp10 REV for the AB plasmid to be inserted in GFE-1	gtgatatctcctctgtttatttattgctcagcggtggcag

Appendix 3

Sequence inserted in pSMART amp to generate pAG_1 vector schematic representation given in Chapter 3.2.1.

TAATACGACTCACTATAGGGAGTTCCCCGCGCCAGCGGGGATAAACCGTGGTCTTCGACCAGT
CTCGGAAGCTCAAAGGTCTGAAGACCAGAGTTCCCCGCGCCAGCGGGGGCTAGTTATTGCTCA
GCGG

Insert for pAG2
TCCTTACGATTAATACAGACTATCGCTTTGCT

Insert for pAG3
AAATATTCACGCTAACGGGCGCCTTTACATGA

Insert for pAG4
CGGTAACATCCAGTTAGTAGTAAACGGACAGA

Insert for pAG5
TTACTCGACGTAACTCGATGGTCGTGTAGCCA

Insert for pAG6
TACAGTCATTGTTGTTATCTGACCCTCTACCA

Insert for pAG7
CGTGGACTCAGGTGTGGTCTGGTAGTGCTGGC

Insert for pAG8
TGTGGTCTGGTAGTGCTGGCGGTGGGGTAAGT

Insert for pAG9
ATCTCCGCTTCCGCAATATCTGGATTAAGTGT

Insert for pAG10
CTATGAAGTAGATTCCATCGGGGCCAGTACGG

Insert for pAG11
TACTGAACGACTGTCTGCAATATTCTTGAATC

Insert for pAG12
AAGTGTGACTGTTTCACAGG

Insert for pAG13
AGGCGTGGACTCAGGTGTGG

Insert for pAG14
AGTGTGCCAACAACTCTTGG

Insert for pAG15
TTCCGCTGCGCATCAATCTG

Insert for pAG16
ACGCTACGAACACAAAGCAG

Insert for pAG17
CAGCATCCGCTAACTCTGCTC

Insert for pAG18
TACAGTTCCGTAATGAGGCT

Insert for pAG19
GGTAAGTGTGACTGTTTCAC

Insert for pAG20
GATCTCCGCTTCCGCAATAT

Insert pAG21
CTTCCGCAATATCTGGATTA

Insert pAG22
AATACACTCCAACGGTCTCG

pAG_25 insert

TGGACTCACAAAGAAAAAACGCCCGGTGTGCAAGACCGAGCGTTCTGAACAATTACTCGTTCT
CCACCATGATTGCATTAGGTACTGAACGACTGTCTGCAATATTCTTGAATCCGAGACCGTTGGA
GTGTATTTGGAATCGTAACCATCCACCATCAGAGGCTATGAAGTAGATTCCATCGGGGCCAGT
ACGGAAGAAGTTCCAAGAGTTGTTGGCACACTTAATCCAGATATTGCGGAAGCGGAGATCCTG
TGAAACAGTCACACTTACCCCACCGCCAGCACTACCAGACCACACCTGAGTCCACTCCTTGGA
CTTCGCAACGAAGCTGTCACGTAGGTAAGCATCCAGCCATTTACCTCCCCACTTAGTACCTTGA
ATATTACCATCAGTGGCCACAACGGCCTGACCTACGTGGAAGTGTTTGTTAACTACTGCATAGT
CCTGCTTGAGTGTTAAGGTCGTACCATGTACATAGGAGTGGAAGGTACAGTCATTGTTGTTATC
TGACCCTCTACCAATGTACCAGTTATTCCTGTTACCATCTTTAGATAGAATATAGTGCGCCTTGT
CGGATGCGGACTCTAATTGAAGAACATGCCCATTCTGCAATTTTAGCTGACCGGTCATGGCTCC
ACCTTGGGTGATGATCTGTCCGTTTACTACTAACTGGATGTTACCGCCTATCGCTGTTGTCCACT
CTCTACGTTGAACATACATCAGCCATCCGTTCTCATCTCCCCACTCCATGACAGAGTAACGATT
AGTGACACCACCAAAGAACTGTTGATACTGGCCACAGTCAAAACCGTTTGTGGTCATGTAAAG
GCGCCCGTTAGCGTGAATATTTCCTTTCCAGTACACATTGGTTCCATCTACCTTATCAATTGCAC
CAGCCAATCCATTGTAATTTTCCAGCTTGTCTGCCTCACGTTCCGCACGGTCTGCTTGCTGTTCT
GCCAAATGAGCAGAGTTAGCGGATGCTGTGGCAGAGTTCTCAGCGTTTACCTCAGATTGATGC
GCAGCGGAAGCAGAGTTCCCAGCAGATGTAGCGTATTGACCAGCCGTATTCTTGAACCGCTTG
GCTTCGTCTCGGAAACCCTTGGTCTCATCGCGCCACTGCTTTGTGTTCGTAGCGTTGGTACTGG
ACTCGTTCTTAAAGCCCTCCGCTTGGTTTCTGAAAGTCTCAGCCTCATTACGGAACTGTAAGGC
TTCATTACGTGCTTGCCATGAGTTCTGGTTCATGGTCTTTAGTTGACCAAACGGAACAGCATCG
CGGTCATCCACGGCGTTCGCTAGGTTCACAATTCGACGACCACGAGCATCCAAGTGACCATCG
TTATTGACACCGATAGTATCCGTAGTGAGGTCACGGGCCTCTTCCGCTACGTGCATCGTTTGAA
TCTGAGCGACGTTAAGGTCATACGCGCGGAGGATTGAACCATCCGTAAAGTCAACCAATCGGT
CGGTAGTGGAGGTTACTCGACGTAACTCGATGGTCGTGTAGCCATCGGCTGGACCCCAAGCCT
TTGTCAGAGAGATAGTAGTACGTGTAGCAAAGCGATAGTCTGTATTAATCGTAAGGACCTTTC
GGTCTACACCAATAAGAGTTACCACTACGAACTTACGGGCTAGATACTCAAACGGGATATTAA
AATCACGATTGGAGCCATCTAACTGGTAAGTCAAAACGGTTTTAATTACGTTAGCCATCTAGTA
TTTCTCCTCTTTCTCTAGTTCTCCCTATAGTGAGTCGTATTATCTAGTAGCGGCCGCTGCAG

pAG_26 insert

TGGACTCACAAAGAAAAAACGCCCGGTGTGCAAGACCGAGCGTTCTGAACAATTAAGCCAGG
TTGGCGTCCAGGAACTCTTTCAGTTGGCCTTTAGACAGAGCGCCTACTTTGGTTGCTGCCACTT
CACCGTTTTTGAACAGCAGCAGAGTTGGGATGCCACGGATACCGTATTTCGGTGCGGTGCCAG
GGTTCTGGTCAATGTTCAGTTTGGCCACAGTCAGCTTACCCTGATATTCGTCCGCAATTTCGTCC
AGGATTGGAGCGATCATTTTGCAAGGGCCGCACCACTCTGCCCAGAAATCAACCAGAATTGCA
CCATCAGCTTTCAGCACGTCGGTATCAAAAGAGTCATCGGTCAGGTGAATGATTTTATCGCTCA
TTTGACCTCCTTAAAGTAAATCCGGTTGCCTTGGCGGGGCCTTTCGGCCCCGCTACTCAAGAAT
CAGGTGGTCACAGACGCCGCGCGCCCCGAAGAACGCGAACGAGAACGACGGCCCGCTGTACC
AGAGCGCAGCGCGAGAACCCGAGAGCGACGTGCCGTTCCAGGCGCCGCCGAATAGCGCAGCA
GCGGGCTGGGCGTACACCGATCCTCTGCCGCCAGTGTTGGCCGTGTATTCGGATGCGCCATTCA
CGCCGCCGAACTCGTTACCCCACGTCCACAAGCAACCGGACGCCTGCACAACGCCCCACTTGG
ACGTGAAGATGTTCCACGCGCTCGTTGCGCCCGTGCCGTTCACGCCGGTGGTGGGCACGTCGG
TGCCGCCGCTGGACGTAGCCTCGGTCGTGCCGAAAGCCAGCGCCTGGAATTCGTTGTAGTTAG
GCAGGCGCTTACCGTGATGAGTCATGACCTCAGCGAAGTTGTACCAAGCTCCGTCACTGTAGG
CCGCGCTGCCGTCTCCACCGAACTTGGTAGATTTCTTAGGTGATGCACTACCATCTGCAATTGT
CACGTTGTATTTGCTGGTGCCATCGGTCAGGTGGTTGACGCCTAGCAGATAGATGTCTGCCCAA
AACGCGCCGGCAACCAGCGTCATGCCGCGCGGGTCGAGCGCAGCCGGGCGAAACTTCTTCGCA
ACGAAGCTGTCACGTAGGTAAGCATCCAGCCATTTACCTCCCCACTTAGTACCTTGAATATTAC
CATCAGTGGCCACAACGGCCTGACCTACGTGGAAGTGTTTGTTAACTACTGCATAGTCCTGCTT
GAGTGTTAAGGTCGTACCATGTACATAGGAGTGGAAGGTACAGTCATTGTTGTTATCTGACCCT
CTACCAATGTACCAGTTATTCCTGTTACCATCTTTAGATAGAATATAGTGCGCCTTGTCGGATG
CGGACTCTAATTGAAGAACATGCCCATTCTGCAATTTTAGCTGACCGGTCATGGCTCCACCTTG
GGTGATGATCTGTCCGTTTACTACTAACTGGATGTTACCGCCTATCGCTGTTGTCCACTCTCTAC
GTTGAACATACATCAGCCATCCGTTCTCATCTCCCCACTCCATGACAGAGTAACGATTAGTGAC
ACCACCAAAGAACTGTTGATACTGGCCACAGTCAAAACCGTTTGTGGTCATGTAAAGGCGCCC
GTTAGCGTGAATATTTCCTTTCCAGTACACATTGGTTCCATCTACCTTATCAATTGCACCAGCCA
ATCCATTGTAATTTTCCAGCTTGTCTGCCTCACGTTCCGCACGGTCTGCTTGCTGTTCTGCCAAA
TGAGCAGAGTTAGCGGATGCTGTGGCAGAGTTCTCAGCGTTTACCTCAGATTGATGCGCAGCG
GAAGCAGAGTTCCCAGCAGATGTAGCGTATTGACCAGCCGTATTCTTGAACCGCTTGGCTTCGT
CTCGGAAACCCTTGGTCTCATCGCGCCACTGCTTTGTGTTCGTAGCGTTGGTACTGGACTCGTT
CTTAAAGCCCTCCGCTTGGTTTCTGAAAGTCTCAGCCTCATTACGGAACTGTAAGGCTTCATTA

CGTGCTTGCCATGAGTTCTGGTTCATGGTCTTTAGTTGACCAAACGGAACAGCATCGCGGTCAT
CCACGGCGTTCGCTAGGTTCACAATTCGACGACCACGAGCATCCAAGTGACCATCGTTATTGA
CACCGATAGTATCCGTAGTGAGGTCACGGGCCTCTTCCGCTACGTGCATCGTTTGAATCTGAGC
GACGTTAAGGTCATACGCGCGGAGGATTGAACCATCCGTAAAGTCAACCAATCGGTCGGTAGT
GGAGGTTACTCGACGTAACTCGATGGTCGTGTAGCCATCGGCTGGACCCCAAGCCTTTGTCAG
AGAGATAGTAGTACGTGTAGCAAAGCGATAGTCTGTATTAATCGTAAGGACCTTTCGGTCTAC
ACCAATAAGAGTTACCACTACGAACTTACGGGCTAGATACTCAAACGGGATATTAAAATCACG
ATTGGAGCCATCTAACTGGTAAGTCAAAACGGTTTTAATTACGTTAGCCAT

pAG_27 insert

ATGGCAAACGTGATTAAGACCGTATTGACCTACCAATTAGATGGTAGTAACAGGGACTTCAAC
ATTCCATTTGAATACCTAGCGCGTAAATTCGTTGTGGTTACTCTGATTGGCGTAGATCGTAAAG
TTTTAACAATCAACACGGATTATCGCTTCGCGACGCGCACGACTATCAGTTTAACGAAAGCGT
GGGGGCCAGCGGATGGCTATACGACAATCGAATTACGCCGAGTGACCTCCACGACCGACCGAC
TGGTAGATTTCACTGACGGCAGTATTTTGCGCGCCTACGATTTAAATGTGGCACAAATCCAGAC
TATGCACGTTGCCGAGGAAGCGCGTGATTTAACAACCGACACAATCGGCGTCAACAATGACGG
CCATCTCGACGCACGTGGGCGTCGCATCGTTAATCTCGCGAACGCGGTCGATGACCGTGATGC
CGTACCCTTCGGTCAGCTCAAAACAATGAACCAAAACTCCTGGCAGGCTCGTAATGAGGCCCT
CCAATTTCGTAACGAAGCGGAAACATTTAGAAACCAGGCGGAAGGCTTTAAAAACGAGATCTC
CACGAATGCGACTAATACCAAACAGTGGAGGGACGAAACCAAGGGCTTCCGAGATGAGGCAA
AAAGATTTAAAAACACAGCAGGCCAGTATGCAACGAGCGCAGGAAATAGCGCAAGTGCAGCC
CACCAGAGTGAAGTTAATGCAGAAAATAGCGCGACTGCTAGTGCAAATAGTGCCCACCTAGCT
GAGCAACAGGCTGACCGCGCGGAACGCGAGGCCGACAAACTGGAGAACTATAACGGTCTTGC
AGGAGCTATAGATAAGGTCGACGGAACGAACGTTTATTGGAAGGGTAACATACATGCTAACG
GACGCCTCTACATGACAACAAACGGCTTCGATTGCGGACAATACCAGCAATTTTTTGGTGGGG
TTACTAATCGCTACAGCGTAATGGAATGGGGTGACGAAAATGGTTGGTTAATGTACGTCCAGA
GACGTGAATGGACTACTGCCATTGGGGGAAATATTCAACTCGTTGTTAATGGTCAAATTATTAC
TCAAGGAGGCGCAATGACTGGCCAATTGAAATTGCAAAATGGTCATGTATTGCAGCTTGAAAG
TGCTAGTGATAAAGCACACTATATCCTATCGAAAGACGGTAACCGAAACAATTGGTATATAGG
ACGAGGAAGCGATAATAACAACGACTGCACCTTCCATTCCTACGTTCACGGCACCACACTCAC
CTTAAAACAGGACTACGCAGTAGTAAATAAGCATTTTCATGTTGGCCAAGCTGTCGTCGCCACT
GACGGCAACATCCAGGGAACCAAATGGGGCGGCAAGTGGCTTGACGCCTATTTGCGAGATTCT
TTTGTCGCTAAAAAGTTTCGCCCGGCTGCGCTCGACCCGCGCGGCATGACGCTGGTTGCCGGCG
CGTTTTGGGCAGACATCTATCTGCTAGGCGTCAACCACCTGACCGATGGCACCAGCAAATACA
ACGTGACAATTGCAGATGGTAGTGCATCACCTAAGAAATCTACCAAGTTCGGTGGAGACGGCA
GCGCGGCCTACAGTGACGGAGCTTGGTACAACTTCGCTGAGGTCATGACTCATCACGGTAAGC
GCCTGCCTAACTACAACGAATTCCAGGCGCTGGCTTTCGGCACGACCGAGGCTACGTCCAGCG
GCGGCACCGACGTGCCCACCACCGGCGTGAACGGCACGGGCGCAACGAGCGCGTGGAACATC
TTCACGTCCAAGTGGGGCGTTGTGCAGGCGTCCGGTTGCTTGTGGACGTGGGGTAACGAGTTC
GGCGGCGTGAATGGCGCATCCGAATACACGGCCAACACTGGCGGCAGAGGATCGGTGTACGC
CCAGCCCGCTGCTGCGCTATTCGGCGGCGCCTGGAACGGCACGTCGCTCTCGGGTTCTCGCGCT
GCGCTCTGGTACAGCGGGCCGTCGTTCTCGTTCGCGTTCTTCGGGGCGCGCGGCGTCTGTGACC
ACCTGATTCTTGAGTAGCGGGGCCGAAAGGCCCCGCCAAGGCAACCG

pAG_28 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCAAGTTTCGCCCGGCTGCGCTCGACCC
GCGCGGCATGACGCTGGTTGCCGGCGCGTTTTGGGCAGACATCTATCTGCTAGGCGTCAACCA
CCTGACCGATGGCACCAGCAAATACAACGTGACAATTGCAGATGGTAGTGCATCACCTAAGAA
ATCTACCAAGTTCGGTGGAGACGGCAGCGCGGCCTACAGTGACGGAGCTTGGTACAACTTCGC
TGAGGTCATGACTCATCACGGTAAGCGCCTGCCTAACTACAACGAATTCCAGGCGCTGGCTTTC
GGCACGACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCACCACCGGCGTGAACGGCAC
GGGCGCAACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGTGCAGGCGTCCGGTTG
CTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGAATACACGGCCAACAC
TGGCGGCAGAGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGCGGCGCCTGGAACGG
CACGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCGTTCTCGTTCGCGTTC
TTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAGTAG

pAG_29 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCGGTTCTGGTTATATTCCTGAAGCTCC
AAGAGATGGGCAAGCATACGTTCGTAAAGATGGCGAATGGGTATTCCTTTCTACCTTTTTAAA
GTTTCGCCCGGCTGCGCTCGACCCGCGCGGCATGACGCTGGTTGCCGGCGCGTTTTGGGCAGA
CATCTATCTGCTAGGCGTCAACCACCTGACCGATGGCACCAGCAAATACAACGTGACAATTGC
AGATGGTAGTGCATCACCTAAGAAATCTACCAAGTTCGGTGGAGACGGCAGCGCGGCCTACAG
TGACGGAGCTTGGTACAACTTCGCTGAGGTCATGACTCATCACGGTAAGCGCCTGCCTAACTA
CAACGAATTCCAGGCGCTGGCTTTCGGCACGACCGAGGCTACGTCCAGCGGCGGCACCGACGT
GCCCACCACCGGCGTGAACGGCACGGGCGCAACGAGCGCGTGGAACATCTTCACGTCCAAGTG
GGGCGTTGTGCAGGCGTCCGGTTGCTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGG
CGCATCCGAATACACGGCCAACACTGGCGGCAGAGGATCGGTGTACGCCCAGCCCGCTGCTGC
GCTATTCGGCGGCGCCTGGAACGGCACGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGC
GGGCCGTCGTTCTCGTTCGCGTTCTTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAGT
AGCGGGGCCGAAAGGCCCCGCCAAGGCAACCGTACTCGAACCCCTAGCCCGCTCTTATCGGGC
GGCTAGGGGTTTTTTGT

pAG_30 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCAAGTTTCGCCCGGCTGCGCTCGACCC
GCGCGGCATGACGCTGGTTGCCGGCGCGTTTTGGGCAGACATCTATCTGCTAGGCGTCAACCA
CCTGACCGATGGCACCAGCAAATACAACGTGACAATTGCAGATGGTAGTGCATCACCTAAGAA
ATCTACCAAGTTCGGTGGAGACGGCAGCGCGGCCTACAGTGACGGAGCTTGGTACAACTTCGC
TGAGGTCATGACTCATCACGGTAAGCGCCTGCCTAACTACAACGAATTCCAGGCGCTGGCTTTC
GGCACGACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCACCACCGGCGTGAACGGCAC
GGGCGCAACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGTGCAGGCGTCCGGTTG
CTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGAATACACGGCCAACAC
TGGCGGCAGAGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGCGGCGCCTGGAACGG
CACGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCGTTCTCGTTCGCGTTC
TTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAGTAG

pAG_31 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCGGTTCTGGTTATATTCCTGAAGCTCC
AAGAGATGGGCAAGCATACGTTCGTAAAGATGGCGAATGGGTATTCCTTTCTACCTTTTTAACG
GCTTTCATCAAGGCCGACAAGTCGGCCGTCGCCTTCACGCGCACCGGCAACGCAACGGCCAGC
ATCAAGGCTGGCACCATCGTGGAGGTCAACGGCAAGCTGGTGCAGTTCACCGCCGACACGGCC
ATCACCATGCCGGCGCTGACGGCCGGCACCGACTACGCCATCTACGTCTGCGACGATGGCACG
GTGCGCGCCGATTCCAACTTTTCGGCGCCCACTGGCTACACCTCGACCACGGCGCGCAAGGTG
GGCGGCTTCCACTATGCGCCGGGAAGCAACGCTGCAGCGCAGGCTGGTGGAAACACCACGGC
GCAGATCAACGAATACAGCCTGTGGGACATCAAGTTTCGCCCGGCTGCGCTCGACCCGCGCGG
CATGACGCTGGTTGCCGGCGCGTTTTGGGCAGACATCTATCTGCTAGGCGTCAACCACCTGACC
GATGGCACCAGCAAATACAACGTGACAATTGCAGATGGTAGTGCATCACCTAAGAAATCTACC
AAGTTCGGTGGAGACGGCAGCGCGGCCTACAGTGACGGAGCTTGGTACAACTTCGCTGAGGTC
ATGACTCATCACGGTAAGCGCCTGCCTAACTACAACGAATTCCAGGCGCTGGCTTTCGGCACG
ACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCACCACCGGCGTGAACGGCACGGGCGC
AACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGTGCAGGCGTCCGGTTGCTTGTG
GACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGAATACACGGCCAACACTGGCG
GCAGAGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGCGGCGCCTGGAACGGCACGT
CGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCGTTCTCGTTCGCGTTCTTCGG
GGCGCGCGGCGTCTGTGACCACCTGATTCTTGAGTAG

pAG_32 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCAACGAATACAGCCTGTGGGACATCA
AGTTtCGCCCGGCTGCGCTCGACCCGCGCGGCATGACGCTGGTtGCCGGCGCGTTTTGGGCaGAC
ATCTATCTGCTAGGCGTCAACCACCTGACCGATGGCACCAGCAAATACAACGTGACaATtGCaG
AtGGtAGtGCaTCaCCtAAGAAATCTACCAAGTTCGGtGGaGACGGCAGCGCGGCCTACAGtGACGG
aGCtTGGTACAACTTCGCtGAGGTCATGACTCAtCACGGtAAGCGCCTGCCtAACTACAACGAATT
CCAGGCGCTGGCTTTCGGCACGACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCACCAC
CGGCGTGAACGGCACGGGCGCaACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGT
GCAGGCGTCCGGTTGCTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGA
ATACACGGCCAACACTGGCGGCagaGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGC
GGCGCCTGGAACGGCACGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCG
TTCTCGTTCGCGTTCTTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAG

pAG_33 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCGGTTCTGGTTATATTCCTGAAGCTCC
AAGAGATGGGCAAGCATACGTTCGTAAAGATGGCGAATGGGTATTCCTTTCTACCTTTTTAAAC
GAATACAGCCTGTGGGACATCAAGTTtCGCCCGGCTGCGCTCGACCCGCGCGGCATGACGCTGG
TtGCCGGCGCGTTTTGGGCaGACATCTATCTGCTAGGCGTCAACCACCTGACCGATGGCACCAG
CAAATACAACGTGACaATtGCaGAtGGtAGtGCaTCaCCtAAGAAATCTACCAAGTTCGGtGGaGACG
GCAGCGCGGCCTACAGtGACGGaGCtTGGTACAACTTCGCtGAGGTCATGACTCAtCACGGtAAGC
GCCTGCCtAACTACAACGAATTCCAGGCGCTGGCTTTCGGCACGACCGAGGCTACGTCCAGCGG
CGGCACCGACGTGCCCACCACCGGCGTGAACGGCACGGGCGCaACGAGCGCGTGGAACATCTT
CACGTCCAAGTGGGGCGTTGTGCAGGCGTCCGGTTGCTTGTGGACGTGGGGTAACGAGTTCGG
CGGCGTGAATGGCGCATCCGAATACACGGCCAACACTGGCGGCagaGGATCGGTGTACGCCCAG
CCCGCTGCTGCGCTATTCGGCGGCGCCTGGAACGGCACGTCGCTCTCGGGTTCTCGCGCTGCGC
TCTGGTACAGCGGGCCGTCGTTCTCGTTCGCGTTCTTCGGGGCGCGCGGCGTCTGTGACCACCT
GATTCTTGAG

pAG_34 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCATCAAGTTTCGCCCGGCTGCGCTCGA
CCCGCGCGGCATGACGCTGGTTGCCGGCGCGTTTTGGGCAGACATCTATCTGCTAGGCGTCAA
CCACCTGACCGATGGCACCAGCAAATACAACGTGACAATTGCAGATGGTAGTGCATCACCTAA
GAAATCTACCAAGTTCGGTGGAGACGGCAGCGCGGCCTACAGTGACGGAGCTTGGTACAACTT
CGCTGAGGTCATGACTCATCACGGTAAGCGCCTGCCTAACTACAACGAATTCCAGGCGCTGGC
TTTCGGCACGACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCACCACCGGCGTGAACGG
CACGGGCGCAACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGTGCAGGCGTCCGG
TTGCTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGAATACACGGCCAA
CACTGGCGGCAGAGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGCGGCGCCTGGAA
CGGCACGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCGTTCTCGTTCGCG
TTCTTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAG

pAG_35 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCGGTTCTGGTTATATTCCTGAAGCTCC
AAGAGATGGGCAAGCATACGTTCGTAAAGATGGCGAATGGGTATTCCTTTCTACCTTTTTAATC
AAGTTTCGCCCGGCTGCGCTCGACCCGCGCGGCATGACGCTGGTTGCCGGCGCGTTTTGGGCA

GACATCTATCTGCTAGGCGTCAACCACCTGACCGATGGCACCAGCAAATACAACGTGACAATT
GCAGATGGTAGTGCATCACCTAAGAAATCTACCAAGTTCGGTGGAGACGGCAGCGCGGCCTAC
AGTGACGGAGCTTGGTACAACTTCGCTGAGGTCATGACTCATCACGGTAAGCGCCTGCCTAAC
TACAACGAATTCCAGGCGCTGGCTTTCGGCACGACCGAGGCTACGTCCAGCGGCGGCACCGAC
GTGCCCACCACCGGCGTGAACGGCACGGGCGCAACGAGCGCGTGGAACATCTTCACGTCCAAG
TGGGGCGTTGTGCAGGCGTCCGGTTGCTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAAT
GGCGCATCCGAATACACGGCCAACACTGGCGGCAGAGGATCGGTGTACGCCCAGCCCGCTGCT
GCGCTATTCGGCGGCGCCTGGAACGGCACGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACA
GCGGGCCGTCGTTCTCGTTCGCGTTCTTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGA
G

pAG_36 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCCTGGCGCGGCACGATCTGGTGAAAA
CGGCTTTCATCAAGGCCGACAAGTCGGCCGTCGCCTTCACGCGCACCGGCAACGCAACGGCCA
GCATCAAGGCTGGCACCATCGTGGAGGTCAACGGCAAGCTGGTGCAGTTCACCGCCGACACGG
CCATCACCATGCCGGCGCTGACGGCCGGCACCGACTACGCCATCTACGTCTGCGACGATGGCA
CGGTGCGCGCCGATTCCAACTTTTCGGCGCCCACTGGCTACACCTCGACCACGGCGCGCAAGG
TGGGCGGCTTCCACTATGCGCCGGGAAGCAACGCTGCAGCGCAGGCTGGCGGAAACACCACG
GCGCAGATCAACGAATACAGCCTGTGGGACATCAAGTTCCGCCCGGCTGCGCTCGACCCGCGC
GGCATGACGCTGGTGGCCGGCGCGTTTTGGGCGGACATCTATCTGCTAGGCGTCAACCACCTG
ACCGATGGCACCAGCAAATACAACGTGACCATCGCGGACGGCAGCGCGTCCCCGAAGAAATC
TACCAAGTTCGGCGGCGACGGCAGCGCGGCCTACAGCGACGGCGCCTGGTACAACTTCGCCGA
GGTCATGACTCACCACGGCAAGCGCCTGCCCAACTACAACGAATTCCAGGCGCTGGCTTTCGG
CACGACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCACCACCGGCGTGAACGGCACGG
GCGCCACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGTGCAGGCGTCCGGTTGCT
TGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGAATACACGGCCAACACTG
GCGGCCGCGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGCGGCGCCTGGAACGGCA
CGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCGTTCTCGTTCGCGTTCTTC
GGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAGTAG

pAG_37 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCGGTTCTGGTTATATTCCTGAAGCTCC
AAGAGATGGGCAAGCATACGTTCGTAAAGATGGCGAATGGGTATTCCTTTCTACCTTTTTACTG
GCGCGGCACGATCTGGTGAAAACGGCTTTCATCAAGGCCGACAAGTCGGCCGTCGCCTTCACG
CGCACCGGCAACGCAACGGCCAGCATCAAGGCTGGCACCATCGTGGAGGTCAACGGCAAGCT
GGTGCAGTTCACCGCCGACACGGCCATCACCATGCCGGCGCTGACGGCCGGCACCGACTACGC
CATCTACGTCTGCGACGATGGCACGGTGCGCGCCGATTCCAACTTTTCGGCGCCCACTGGCTAC
ACCTCGACCACGGCGCGCAAGGTGGGCGGCTTCCACTATGCGCCGGGAAGCAACGCTGCAGCG
CAGGCTGGCGGAAACACCACGGCGCAGATCAACGAATACAGCCTGTGGGACATCAAGTTCCG
CCCGGCTGCGCTCGACCCGCGCGGCATGACGCTGGTGGCCGGCGCGTTTTGGGCGGACATCTA
TCTGCTAGGCGTCAACCACCTGACCGATGGCACCAGCAAATACAACGTGACCATCGCGGACGG
CAGCGCGTCCCCGAAGAAATCTACCAAGTTCGGCGGCGACGGCAGCGCGGCCTACAGCGACG
GCGCCTGGTACAACTTCGCCGAGGTCATGACTCACCACGGCAAGCGCCTGCCCAACTACAACG
AATTCCAGGCGCTGGCTTTCGGCACGACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCA
CCACCGGCGTGAACGGCACGGGCGCCACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGC
GTTGTGCAGGCGTCCGGTTGCTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCA
TCCGAATACACGGCCAACACTGGCGGCCGCGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTA
TTCGGCGGCGCCTGGAACGGCACGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGCGGG
CCGTCGTTCTCGTTCGCGTTCTTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAGTAG

pAG_38 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT

ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCCACGATCTGGTGAAAACGGCTTTCAT
CAAGGCCGACAAGTCGGCCGTCGCCTTCACGCGCACCGGCAACGCAACGGCCAGCATCAAGG
CTGGCACCATCGTGGAGGTCAACGGCAAGCTGGTGCAGTTCACCGCCGACACGGCCATCACCA
TGCCGGCGCTGACGGCCGGCACCGACTACGCCATCTACGTCTGCGACGATGGCACGGTGCGCG
CCGATTCCAACTTTTCGGCGCCCACTGGCTACACCTCGACCACGGCGCGCAAGGTGGGCGGCT
TCCACTATGCGCCGGGAAGCAACGCTGCAGCGCAGGCTGGCGGAAACACCACGGCGCAGATC
AACGAATACAGCCTGTGGGACATCAAGTTCCGCCCGGCTGCGCTCGACCCGCGCGGCATGACG
CTGGTGGCCGGCGCGTTTTGGGCGGACATCTATCTGCTAGGCGTCAACCACCTGACCGATGGC
ACCAGCAAATACAACGTGACCATCGCGGACGGCAGCGCGTCCCCGAAGAAATCTACCAAGTTC
GGCGGCGACGGCAGCGCGGCCTACAGCGACGGCGCCTGGTACAACTTCGCCGAGGTCATGACT
CACCACGGCAAGCGCCTGCCCAACTACAACGAATTCCAGGCGCTGGCTTTCGGCACGACCGAG
GCTACGTCCAGCGGCGGCACCGACGTGCCCACCACCGGCGTGAACGGCACGGGCGCCACGAG
CGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGTGCAGGCGTCCGGTTGCTTGTGGACGTG
GGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGAATACACGGCCAACACTGGCGGCCGCG
GATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGCGGCGCCTGGAACGGCACGTCGCTCT
CGGGTTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCGTTCTCGTTCGCGTTCTTCGGGGCGCG
CGGCGTCTGTGACCACCTGATTCTTGAGTAG

pAG_39 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCGGTTCTGGTTATATTCCTGAAGCTCC
AAGAGATGGGCAAGCATACGTTCGTAAAGATGGCGAATGGGTATTCCTTTCTACCTTTTTACAC
GATCTGGTGAAAACGGCTTTCATCAAGGCCGACAAGTCGGCCGTCGCCTTCACGCGCACCGGC
AACGCAACGGCCAGCATCAAGGCTGGCACCATCGTGGAGGTCAACGGCAAGCTGGTGCAGTTC
ACCGCCGACACGGCCATCACCATGCCGGCGCTGACGGCCGGCACCGACTACGCCATCTACGTC
TGCGACGATGGCACGGTGCGCGCCGATTCCAACTTTTCGGCGCCCACTGGCTACACCTCGACC
ACGGCGCGCAAGGTGGGCGGCTTCCACTATGCGCCGGGAAGCAACGCTGCAGCGCAGGCTGG
CGGAAACACCACGGCGCAGATCAACGAATACAGCCTGTGGGACATCAAGTTCCGCCCGGCTGC
GCTCGACCCGCGCGGCATGACGCTGGTGGCCGGCGCGTTTTGGGCGGACATCTATCTGCTAGG
CGTCAACCACCTGACCGATGGCACCAGCAAATACAACGTGACCATCGCGGACGGCAGCGCGTC
CCCGAAGAAATCTACCAAGTTCGGCGGCGACGGCAGCGCGGCCTACAGCGACGGCGCCTGGT
ACAACTTCGCCGAGGTCATGACTCACCACGGCAAGCGCCTGCCCAACTACAACGAATTCCAGG
CGCTGGCTTTCGGCACGACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCACCACCGGCG
TGAACGGCACGGGCGCCACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGTGCAGG
CGTCCGGTTGCTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGAATACA
CGGCCAACACTGGCGGCCGCGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGCGGCG
CCTGGAACGGCACGTCGCTCTCGGGTTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCGTTCTC
GTTCGCGTTCTTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAGTAG

pAG_40 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCGGTTCTGGTTATATTCCTGAAGCTCC
AAGAGATGGGCAAGCATACGTTCGTAAAGATGGCGAATGGGTATTCCTTTCTACCTTTTTAAA
GGCGTGGACTCAGGTGTGGTCTGGTAGTGCTGGCGGTGGGGTAAGTGTGACTGTTTCACAGGA
TCTCCGCTTCCGCAATATCTGGATTAAGTGTGCCAACAACTCTTGGAACTTCTTCCGTACTGGC
CCCGATGGAATCTACTTCATAGCCTCTGATGGTGGATGGTTACGATTCCAAATACACTCCAACG
GTCTCGGATTCAAGAATATTGCAGACAGTCGTTCAGTACCTAATGCAATCATGGTGGAGAACG
AGTAA

pAG_41 insert

GGGCATGTTCTTCAATTAGAGTCCGCATCCGACAAGGCGCACTATATTCTATCTAAAGATGGTA
ACAGGAATAACTGGTACATTGGTAGAGGGTCAGATAACAACAATGACTGTACCTTCCACTCCT
ATGTACATGGTACGACCTTAACACTCAAGCAGGACTATGCAGTAGTTAACAAACACTTCCACG
TAGGTCAGGCCGTTGTGGCCACTGATGGTAATATTCAAGGTACTAAGTGGGGAGGTAAATGGC
TGGATGCTTACCTACGTGACAGCTTCGTTGCGAAGTCCGGTTCTGGTTATATTCCTGAAGCTCC
AAGAGATGGGCAAGCATACGTTCGTAAAGATGGCGAATGGGTATTCCTTTCTACCTTTTTAAA
AGCATGGACACAAGTTTGGTCGGGTAGTGCAGGCGGTGGGGTTTCTGTAACAGTCAGTCAAGA
TCTTCGCTTTCGCAACATCTGGATCAAATGCGCAAATAATAGTTGGAACTTTTTCCGCACTGGC
CCAGATGGAATATATTTCATAGCATCTGACGGTGGCTGGTTGCGTTTTCAGATCCATTCTAATG
GCCTTGGCTTCAAGAACATTGCAGATAGCCGTTCCGTACCCAATGCAATTATGGTAGAGAATG
AGTAA

PBC1 lysin insert (codon optimised)

AAAGAGGAGAAATACTAGATGGACAAGTTTACCGTCCATGCCGGGCATACGCTTGGCGGAGG
AGCTAGTGGAAGTGGTTTCGAAGAAAGCGCCGTGGCCCGCCAATTCTGTCCTGTCCTTATCGAC
GCATTACGCAAAGTCGGGGCAACCGTGACGGATTGTACAGATAACGTGAGCACCACGCAATCA
GCAAATCTTACCCGTCTTATCAACCTTTGCAACGCAGTCAACCGCGATTTAGACATCAGTCTGC
ACTTCAATAGTTCGGATGACCCAGCCGCCACGGGGGTCGAAGTCTTATATTACGATCAACGTG
CGCTGGCGGCTAAAGTGTCCGACGCAATCTCAAAAGCCAGTGGGCTGCGTGACCGCGGCCCCA
AGGAGCGTAAGGATCTGGCAGTATTAAATGGGACAAAGGCACCCGCAATCCTTATTGAGCTTG
CATTCATCAGTAACGCATCTGATATGAAAAAGTTCTTCGACAATATGCAGGCCATCGCGAATG
CCATCGTCCAAGCTGTAACCGGTAAGGCAGTTAACGTGGACCCACCGGCGCCCAAGCGTACTG
CTACTATTATTACAACCGGCGGATTAGGACAGGACGCGGCCAAACAGGCGGTTGACATGATGT
TCGCAAAAGGATGGTACGGGACCCTGACTTTTCAGACCGACGGGATTGCCGTGTTGAAGACAG
GTGGGCTGTCAGGCGACAAATTGCAGGCCGCTCGCGAGTATTTCGCGTGGCGCGGTTGGTGGT
ATCAGGAGGAAGTAATCGAATACTAA

H5 lysin insert (codon optimised)

ATGCAAGCAAAACTAACTAAAAAAGAGTTTATAGAGTGGTTGAAAACATCTGAGGGAAAACA
ATATAATGCGGACGGATGGTATGGATTTCAATGCTTTGATTATGCTAATGCTGGTTGGAAAGCT
TTGTTCGGATTACTTTTAAAAGGTGTAGGTGCAAAAGATATTCCATTCGCCAACAATTTTGACG
GACTAGCTACTGTATACCAAAATACACCAGACTTCTTAGCGCAACCTGGCGACATGGTTGTATT
CGGTAGTAATTATGGTGCAGGATACGGTCATGTTGCATGGGTAATTGAAGCAACTTTAGATTAT
ATCATTGTATATGAGCAGAATTGGCTCGGCGGTGGCTGGACAGACGGTGTACAACAACCTGGC
TCTGGTTGGGAAAAAGTTACAAGACGCCAACACGCTTACGACTTCCCTATGTGGTTTATCCGTC
CTAACTTCAAAAGCGAAACAGCTCCACGATCAGTACAATCTCCTACGCAAGCATCTAAAAAGG
AAACGGCTAAGCCACAACCTAAAGCGGTAGAACTTAAAATCATCAAAGATGTGGTTAAAGGTT
ATGATCTACCTAAGCGTGGTAGTAACCCTAACTTTATAGTTATTCACAACGACGCAGGAAGCA
AAGGAGCAACAGCAGAAGCATATCGTAATGGATTAGTTAACGCGCCATTATCGAGACTAGAG
GCAGGTATTGCGCATAGTTACGTATCAGGTAACACGGTTTGGCAAGCCTTAGACGAATCGCAA
GTAGGTTGGCATACAGCGAACCAAATAGGCAATAAATATGGTTACGGTATTGAAGTGTGCCAA
TCAATGGGAGCAGATAATGCGACGTTTTTAAAAAATGAACAGGCGACTTTCCAAGAATGTGCT
AGATTGTTGAAAAAATGGGGATTACCAGCAAACCGTAACACAATCCGATTACACAACGAATTC
ACTTCAACATCATGCCCACACAGAAGCTCAGTATTGCACACTGGTTTTGACCCAGTAACTCGTG
GTCTATTGCCAGAAGATAAACGACTACAACTTAAAGACTACTTTATCAAGCAAATTAGAGCAT
ACATGGATGGTAAAATACCGGTTGCTACTGTCTCAAATGATTCAAGTGCTTCAAGTAATACAGT
TAAGCCAGTTGCAAGTGCATGGAAACGTAATAAATATGGTACTTACTACATGGAAGAAAGTGC
TAGATTCACAAACGGCAATCAACCAATCACAGTAAGAAAAGTGGGGCCATTTTTATCTTGTCC
AGTGGGTTATCAGTTCCAACCTGGTGGATATTGTGATTATACAGAAGTGATGTTACAAGATGG
CCACGTTTGGGTAGGATATACATGGGAGGGGCAACGTTATTACTTGCCTATTAGAACATGGAA
TGGTTCTGCCCCGCCTAATCAGATATTAGGTGACTTATGGGGAGAAATCAGTTAA

CD27L insert (codon optimised)

AAAGAGGAGAAATACTAGATGAAGATTTGCATTACGGTGGGCCACAGCATTCTGAAGAGTGGT
GCTTGCACTTCCGCAGACGGTGTGGTGAACGAATATCAGTACAATAAATCCTTGGCCCCCGTTT
TGGCAGATACGTTCCGTAAGGAGGGCCACAAGGTCGATGTAATTATCTGTCCCGAAAAACAGT
TTAAGACAAAAAATGAAGAAAAGTCATACAAAATCCCGCGCGTTAATAGTGGGGGTTATGATC
TGTTGATCGAGTTGCACTTGAATGCCTCTAACGGACAAGGAAAGGGCTCGGAGGTGCTTTATT
ACTCGAATAAGGGTTTGGAGTATGCGACCCGTATTTGCGATAAATTAGGAACCGTTTTCAAGA
ACCGTGGGGCGAAATTGGACAAACGCCTGTATATCCTTAACTCAAGCAAGCCTACCGCTGTTC
TGATCGAAAGTTTCTTTTGCGACAATAAGGAGGATTATGACAAAGCAAAGAAACTGGGGCACG
AGGGGATCGCCAAACTTATCGTGGAGGGGGTGTTGAATAAGAACATCAATAATGAAGGAGTC
AAGCAGATGTATAAACACACAATTGTATACGATGGCGAGGTTGACAAAATTAGCGCAACTGTC
GTCGGGTGGGGATACAATGATGGGAAAATCCTGATTTGCGACATCAAGGATTATGTGCCCGGA
CAAACACAGAACCTTTACGTTGTTGGAGGAGGAGCTTGCGAAAAAATCAGTAGTATTACGAAG
GAGAAATTTATCATGATTAAAGGGAATGATCGTTTCGACACTCTTTATAAAGCGCTGGATTTTA
TCAACCGCTAG

2 in 1 (H5 and PBC1) insert

CGATATGCAAGCAAAACTAACTAAAAAAGAGTTTATAGAGTGGTTGAAAACATCTGAGGGAA
AACAATATAATGCGGACGGATGGTATGGATTTCAATGCTTTGATTATGCTAATGCTGGTTGGAA
AGCTTTGTTCGGATTACTTTTAAAAGGTGTAGGTGCAAAAGATATTCCATTCGCCAACAATTTT
GACGGACTAGCTACTGTATACCAAAATACACCAGACTTCTTAGCGCAACCTGGCGACATGGTT
GTATTCGGTAGTAATTATGGTGCAGGATACGGTCATGTTGCATGGGTAATTGAAGCAACTTTAG
ATTATATCATTGTATATGAGCAGAATTGGCTCGGCGGTGGCTGGACAGACGGTGTACAACAAC
CTGGCTCTGGTTGGGAAAAAGTTACAAGACGCCAACACGCTTACGACTTCCCTATGTGGTTTAT
CCGTCCTAACTTCAAAAGCGAAACAGCTCCACGATCAGTACAATCTCCTACGCAAGCATCTAA
AAAGGAAACGGCTAAGCCACAACCTAAAGCGGTAGAACTTAAAATCATCAAAGATGTGGTTA
AAGGTTATGATCTACCTAAGCGTGGTAGTAACCCTAACTTTATAGTTATTCACAACGACGCAGG
AAGCAAAGGAGCAACAGCAGAAGCATATCGTAATGGATTAGTTAACGCGCCATTATCGAGACT
AGAGGCAGGTATTGCGCATAGTTACGTATCAGGTAACACGGTTTGGCAAGCCTTAGACGAATC
GCAAGTAGGTTGGCATACAGCGAACCAAATAGGCAATAAATATGGTTACGGTATTGAAGTGTG
CCAATCAATGGGAGCAGATAATGCGACGTTTTTAAAAAATGAACAGGCGACTTTCCAAGAATG
TGCTAGATTGTTGAAAAAATGGGGATTACCAGCAAACCGTAACACAATCCGATTACACAACGA
ATTCACTTCAACATCATGCCCACACAGAAGCTCAGTATTGCACACTGGTTTTGACCCAGTAACT
CGTGGTCTATTGCCAGAAGATAAACGACTACAACTTAAAGACTACTTTATCAAGCAAATTAGA
GCATACATGGATGGTAAAATACCGGTTGCTACTGTCTCAAATGATTCAAGTGCTTCAAGTAATA
CAGTTAAGCCAGTTGCAAGTGCATGGAAACGTAATAAATATGGTACTTACTACATGGAAGAAA
GTGCTAGATTCACAAACGGCAATCAACCAATCACAGTAAGAAAAGTGGGGCCATTTTTATCTT
GTCCAGTGGGTTATCAGTTCCAACCTGGTGGATATTGTGATTATACAGAAGTGATGTTACAAGA
TGGCCACGTTTGGGTAGGATATACATGGGAGGGGCAACGTTATTACTTGCCTATTAGAACATG
GAATGGTTCTGCCCCGCCTAATCAGATATTAGGTGACTTATGGGGAGAAATCAGTTAAAAAGA
GGAGAAATACTAGATGGACAAGTTTACCGTCCATGCCGGGCATACGCTTGGCGGAGGAGCTAG
TGGAAGTGGTTTCGAAGAAAGCGCCGTGGCCCGCCAATTCTGTCCTGTCCTTATCGACGCATTA
CGCAAAGTCGGGGCAACCGTGACGGATTGTACAGATAACGTGAGCACCACGCAATCAGCAAA
TCTTACCCGTCTTATCAACCTTTGCAACGCAGTCAACCGCGATTTAGACATCAGTCTGCACTTC
AATAGTTCGGATGACCCAGCCGCCACGGGGGTCGAAGTCTTATATTACGATCAACGTGCGCTG
GCGGCTAAAGTGTCCGACGCAATCTCAAAAGCCAGTGGGCTGCGTGACCGCGGCCCCAAGGA
GCGTAAGGATCTGGCAGTATTAAATGGGACAAAGGCACCCGCAATCCTTATTGAGCTTGCATT
CATCAGTAACGCATCTGATATGAAAAAGTTCTTCGACAATATGCAGGCCATCGCGAATGCCAT
CGTCCAAGCTGTAACCGGTAAGGCAGTTAACGTGGACCCACCGGCGCCCAAGCGTACTGCTAC
TATTATTACAACCGGCGGATTAGGACAGGACGCGGCCAAACAGGCGGTTGACATGATGTTCGC
AAAAGGATGGTACGGGACCCTGACTTTTCAGACCGACGGGATTGCCGTGTTGAAGACAGGTGG
GCTGTCAGGCGACAAATTGCAGGCCGCTCGCGAGTATTTCGCGTGGCGCGGTTGGTGGTATCA
GGAGGAAGTAATCGAA

10A_HR_GFE-1

CGTGCTAACTTCCAAGCGGACCAGATTATCGCTAAGTACGCAATGGGCCACGGTGGTCTTCGC
CCAGAAGCTGCTGGTGCAGTGGTTTTCAAAGTGGAGTGCGGGTTCGAGTGTGTACGCCAATGC
CCCGAAAGATGCTAAATAAACAGAGGAGATATCACATGAGCGATAAAATCATTCACCTGACCG
ATGACTCTTTTGATACCGACGTGCTGAAAGCTGATGGTGCAATTCTGGTTGATTTCTGGGCAGA
GTGGTGCGGCCCTTGCAAAATGATCGCTCCAATCCTGGACGAAATTGCGGACGAATATCAGGG
TAAGCTGACTGTGGCCAAACTGAACATTGACCAGAACCCTGGCACCGCACCGAAATACGGTAT
CCGTGGCATCCCAACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCAGCAACCAAAGTAGGCGC
TCTGTCTAAAGGCCAACTGAAAGAGTTCCTGGACGCCAACCTGGCTTAATGACTAGCATAACC
CCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATGCGCTCAT
ACGATATGAACGTTGAGACTGCCG

10A_HR_GFE-2

CGTGCTAACTTCCAAGCGGACCAGATTATCGCTAAGTACGCAATGGGCCACGGTGGTCTTCGC
CCAGAAGCTGCTGGTGCAGTGGTTTTCAAAGTGGAGTGCGGATTTGAGCTAGAGACTTGTTAA
ATAAACAGAGGAGATATCACATGAGCGATAAAATCATTCACCTGACCGATGACTCTTTTGATA
CCGACGTGCTGAAAGCTGATGGTGCAATTCTGGTTGATTTCTGGGCAGAGTGGTGCGGCCCTTG
CAAAATGATCGCTCCAATCCTGGACGAAATTGCGGACGAATATCAGGGTAAGCTGACTGTGGC
CAAACTGAACATTGACCAGAACCCTGGCACCGCACCGAAATACGGTATCCGTGGCATCCCAAC
TCTGCTGCTGTTCAAAAACGGTGAAGTGGCAGCAACCAAAGTAGGCGCTCTGTCTAAAGGCCA
ACTGAAAGAGTTCCTGGACGCCAACCTGGCTTAATGACTAGCATAACCCCTTGGGGCCTCTAA
ACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATGCGCTCATACGATATGAACGTT
GAGACTGCCG

10A_HR_metadherin

CGTGCTAACTTCCAAGCGGACCAGATTATCGCTAAGTACGCAATGGGCCACGGTGGTCTTCGC
CCAGAAGCTGCTGGTGCAGTGGTTTTCAAAGTGGAGGGGTTAAATGGTCTGTCTTCTGCTGATC
CCAACTCTGATTGGAATGCACCAGCAGAAGAGTGGGGCAATTGGGTAGACGAAGAAAGAGCT
TCACTTCTAAAGTCCCAGGAACCAATTCCTGATGATCAAAAGGTCTCAGATGATGATAAAGAA
AAGGGAGAGGGAGCTCTTCCAACTGGGAAATCCAAATAAATAAACAGAGGAGATATCACATG
AGCGATAAAATCATTCACCTGACCGATGACTCTTTTGATACCGACGTGCTGAAAGCTGATGGT
GCAATTCTGGTTGATTTCTGGGCAGAGTGGTGCGGCCCTTGCAAAATGATCGCTCCAATCCTGG
ACGAAATTGCGGACGAATATCAGGGTAAGCTGACTGTGGCCAAACTGAACATTGACCAGAACC
CTGGCACCGCACCGAAATACGGTATCCGTGGCATCCCAACTCTGCTGCTGTTCAAAAACGGTG
AAGTGGCAGCAACCAAAGTAGGCGCTCTGTCTAAAGGCCAACTGAAAGAGTTCCTGGACGCCA
ACCTGGCTTAATGACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCT
GAAAGGAGGAACTATATGCGCTCATACGATATGAACGTTGAGACTGCCG

10B_HR_GFE-1

ACAGAAGAAACCTTAACGCCAGCACAGGAGGCCGCACGCACCCGCGCTGCTAACAAAGCCCG
AAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATGCGGGTTCGAGTGTGTACGCCAATG
CCCCGAAAGATGCTAAATAAACAGAGGAGATATCACATGAGCGATAAAATCATTCACCTGACC
GATGACTCTTTTGATACCGACGTGCTGAAAGCTGATGGTGCAATTCTGGTTGATTTCTGGGCAG
AGTGGTGCGGCCCTTGCAAAATGATCGCTCCAATCCTGGACGAAATTGCGGACGAATATCAGG
GTAAGCTGACTGTGGCCAAACTGAACATTGACCAGAACCCTGGCACCGCACCGAAATACGGTA
TCCGTGGCATCCCAACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCAGCAACCAAAGTAGGCG
CTCTGTCTAAAGGCCAACTGAAAGAGTTCCTGGACGCCAACCTGGCTTAATGACTAGCATAAC
CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATGCGCTCA
TACGATATGAACGTTGAGACTGCCG

10B_HR_GFE-2

ACAGAAGAAACCTTAACGCCAGCACAGGAGGCCGCACGCACCCGCGCTGCTAACAAAGCCCG
AAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATGCGGATTTGAGCTAGAGACTTGTTA
AATAAACAGAGGAGATATCACATGAGCGATAAAATCATTCACCTGACCGATGACTCTTTTGAT
ACCGACGTGCTGAAAGCTGATGGTGCAATTCTGGTTGATTTCTGGGCAGAGTGGTGCGGCCCTT
GCAAAATGATCGCTCCAATCCTGGACGAAATTGCGGACGAATATCAGGGTAAGCTGACTGTGG
CCAAACTGAACATTGACCAGAACCCTGGCACCGCACCGAAATACGGTATCCGTGGCATCCCAA

CTCTGCTGCTGTTCAAAAACGGTGAAGTGGCAGCAACCAAAGTAGGCGCTCTGTCTAAAGGCC
AACTGAAAGAGTTCCTGGACGCCAACCTGGCTTAATGACTAGCATAACCCCTTGGGGCCTCTA
AACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATGCGCTCATACGATATGAACGT
TGAGACTGCCG

10B_HR_metadherin

ACAGAAGAAACCTTAACGCCAGCACAGGAGGCCGCACGCACCCGCGCTGCTAACAAAGCCCG
AAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAAGGGTTAAATGGTCTGTCTTCTGCTGA
TCCCAACTCTGATTGGAATGCACCAGCAGAAGAGTGGGGCAATTGGGTAGACGAAGAAAGAG
CTTCACTTCTAAAGTCCCAGGAACCAATTCCTGATGATCAAAAGGTCTCAGATGATGATAAAG
AAAAGGGAGAGGGAGCTCTTCCAACTGGGAAATCCAAATAAATAAACAGAGGAGATATCACA
TGAGCGATAAAATCATTCACCTGACCGATGACTCTTTTGATACCGACGTGCTGAAAGCTGATG
GTGCAATTCTGGTTGATTTCTGGGCAGAGTGGTGCGGCCCTTGCAAAATGATCGCTCCAATCCT
GGACGAAATTGCGGACGAATATCAGGGTAAGCTGACTGTGGCCAAACTGAACATTGACCAGA
ACCCTGGCACCGCACCGAAATACGGTATCCGTGGCATCCCAACTCTGCTGCTGTTCAAAAACG
GTGAAGTGGCAGCAACCAAAGTAGGCGCTCTGTCTAAAGGCCAACTGAAAGAGTTCCTGGACG
CCAACCTGGCTTAATGACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTT
GCTGAAAGGAGGAACTATATGCGCTCATACGATATGAACGTTGAGACTGCCG

Appendix 4

pAG_2

TCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCGCGG
GGAACTCAGCAAAGCGATAGTCTGTATTAATCGTAAGGACGGTTTATCCCCGCTGGCGCGGGG
AACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCCAGC
GTAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAAGCT
GACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAGGAT
CTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACC
AGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAGCA
GAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTC
TGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGC
TGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAC
CTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGT
ATCTTTATAGTCCTGTCG

pAG_3

TCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCGCGG
GGAACTCTCATGTAAAGGCGCCCGTTAGCGTGAATATTTCGGTTTATCCCCGCTGGCGCGGGG
AACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCCAGC
GTAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAAGCT
GACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAGGAT
CTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACC
AGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAGCA
GAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTC
TGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGC
TGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAC
CTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGT
ATCTTTATAGTCCTGTCG

pAG_4

GATCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCGC
GGGGAACTCTCTGTCCGTTTACTACTAACTGGATGTTACCGCGGTTTATCCCCGCTGGCGCGGG
GAACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCCAG
CGTAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAAGCT
GACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAGGAT
CTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACC
AGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAGCA
GAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTC
TGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGC
TGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAC
CTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGT
ATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCATCGATTTTTGTGATGCTCGTCA
GGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCAGAAAGGCCCACCCGAAGGTGAGCCAG
GTGATTACATTTAGGTCCTCAAGAAAGGCCCACCCGAAGGTGANCCAGGTGATTACATTTAGG
TCCTCAAGAAAGGCCCACC

pAG_5

GATCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCGC
GGGGAACTCTGGCTACACGACCATCGAGTTACGTCGAGTAACGGTTTATCCCCGCTGGCGCGG
GGAACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCCA
GCGTAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAAG

CTGACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAGG
ATCTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTA
CCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAG
CAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAA
CTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGC
GATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCG
GGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAG
ATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGT
ATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCC
TGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCATCGATTTTTGTGATGCTC
GTCAGGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCAGAAAGGCCCACCCGAAGGTGA
GCCAGGTGATTACATTTAGGTCCTCAAGAAAGGCCCACCCGAAGGT

pAG_6

TACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCGCGGGGA
ACTCTGGTAGAGGGTCAGATAACAACAATGACTGTACGGTTTATCCCCGCTGGCGCGGGGAAC
TCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCCAGCGTA
ACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAAGCTGACT
AGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAGGATCTTCT
TGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGG
TGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAGCAGAGCG
CAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAG
CACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTC
GTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAAC
GGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACA
GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAA
GCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTT
TATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCATCGATTTTTGTGATGCTCGTCAGGGG
GGGCGGANCCTATGGAAAAAACGCCAGCAACGCAAAANGGCCCACCCGAAAGGTG

pAG_7

CCGATCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGC
GCGGGGAACTCACTTACCCCACCGCCAGCACTACCAGACCACACGGTTTATCCCCGCTGGCGC
GGGGAACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTC
CAGCGTAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCA
AGCTGACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAA
GGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGC
TACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTC
AGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAG
AACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTG
GCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGT
CGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTG
AGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAG
GTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACG
CCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCATCGATTTTTGTGATGC
TCGTCAGGGGGGGCGGAGCCTATGGAAAAAACGCCAGCAAC

pAG_8

CGATCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCG
CGGGGAACTCACACTTAATCCAGATATTGCGGAAGCGGAGATCGGTTTATCCCCGCTGGCGCG
GGGAACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCC
AGCGTAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAA
GCTGACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAG
GATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCT
ACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCA
GCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGA
ACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGG
CGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGA
GATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG
TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGANGGANCTTCCAGGGGGAAACGC

CTGGTATCTTTATAGTCCTGTCNGGTTTCGCCACCTCTGACTTGAGCATCGATTTTTGTGATGCT
CGTCNGGGGGGCGGANCCNATGGAAAAACGCCA

pAG_9

TCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCGCGG
GGAACTCACACTTAATCCAGATATTGCGGAAGCGGAGATCGGTTTATCCCCGCTGGCGCGGGG
AACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCCAGC
GTAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAAGCT
GACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAGGAT
CTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACC
AGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAGCA
GAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTC
TGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGC
TGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAC
CTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGT
ATCTTTATAGTCCTGTCG

pAG_10

TCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCGCGG
GGAACTCCTATGAAGTAGATTCCATCGGGGCCAGTACGGCGGTTTATCCCCGCTGGCGCGGGG
AACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCCAGC
GTAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAAGCT
GACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAGGAT
CTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACC
AGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAGCA
GAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTC
TGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGC
TGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAC
CTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGT
ATCTTTATAGTCCTGTCG

pAG_11

TCATACGTATAATGCCGTAAGATCACGCGCCCGCTGAGCAATAACTAGCCCCCGCTGGCGCGG
GGAACTCTACTGAACGACTGTCTGCAATATTCTTGAATCCGGTTTATCCCCGCTGGCGCGGGGA
ACTCCCTATAGTGAGTCGTATTAGAGGCCGCATATCATTCAGGACGAGCCTCAGACTCCAGCG
TAACTGGACTGCAATCAACTCACTGGCTCACCTTCCGGTCCACGATCAGCTAGAATCAAGCTG
ACTAGATAAACTGGCCGTCGTTTTACACGGGTGGGCCTTTCTTCGGTAGAAAATCAAAGGATCT
TCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAG
CGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAGCAGA
GCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTG
TAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAA
GTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTG
AACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGG
TAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTAT
CTTTATAGTCCTGTCG

pAG_12

ATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGT
CCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCA
AGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACAGTGCTCCGAGAACGGG
TGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCCATAGTGACTGGCGATGCT
GTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACATGAAAACGCTTTAACAGAA
AGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCAAGTTATCATCGGCAATGTTG
AATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACTCAACAAGTCTCAGTGTGCTG
AAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTGTGAAACAGTCACACTTGTTTTGG
GACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTTTTGTCTAAAAAATTTCGTAAT

CGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCTGTATTACTGCATTTATTAAGA
GTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCATC

pAG_13

CACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTC
GGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGC
GTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAG
CGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACAGTGCTCCGAGAACGGGTGCGC
ATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCCATAGTGACTGGCGATGCTGTCGG
AATGGACGATCACACTACTCTTCTTTTGCCTATTATAACATGAAAACGCTTTAACAGAAAGAAT
AGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCAAGTTATCATCGGCAATGTTGAATGG
AGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACTCAACAAGTCTCAGTGTGCTGAAGTT
TTGGGACCATTCAAAACAGCATAGCTCTAAAACAGGCGTGGACTCAGGTGTGGGTTTTGGGAC
CATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTTTTGTCTAAAAAATTTCGTAATCGC
ACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCTGTATTACTGCATTTATTAAGAGTA
TTATACCATATTTTTAGTTATTAAGAAATAATCTTCNTCTAAAATATACTTCAGTCACCTCCTA

pAG_14

CCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGT
GATGCCGGCCACGATGCGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGC
GTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGG
ACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGC
CATAGTGACTGGCGATGCTGTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACA
TGAAAACGCTTTAACAGAAAGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCA
AGTTATCATCGGCAATGTTGAATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACT
CAACAAGTCTCAGTGTGCTGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCAAGA
GTTGTTGGCACACTGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTT
TTGTCTAAAAAATTTCGTAATCGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCT
GTATTACTGCATTTATTAAGAGTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCATC

pAG_15

CCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGT
GATGCCGGCCACGATGCGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGC
GTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGG
ACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGC
CATAGTGACTGGCGATGCTGTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACA
TGAAAACGCTTTAACAGAAAGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCA
AGTTATCATCGGCAATGTTGAATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACT
CAACAAGTCTCAGTGTGCTGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACAGATT
GATGCGCAGCGGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATT
TTTGTCTAAAAAATTTCGTAATCGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCC
TGTATTACTGCATTTATTAAGAGTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCATC

pAG_16

CCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGT
GATGCCGGCCACGATGCGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGC
GTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGG
ACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGC
CATAGTGACTGGCGATGCTGTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACA
TGAAAACGCTTTAACAGAAAGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCA
AGTTATCATCGGCAATGTTGAATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACT
CAACAAGTCTCAGTGTGCTGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACTGCTTT
GTGTTCGTAGCGTGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTTT
TGTCTAAAAAATTTCGTAATCGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCTG
TATTACTGCATTTATTAAGAGTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCATC

pAG_17

CCNCCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCAT
CGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGA
TGCGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGG
CTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACAGTGCTCCGAGA
ACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCCATAGTGACTGGCG
ATGCTGTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACATGAAAACGCTTTAA
CAGAAAGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCAAGTTATCATCGGCA
ATGTTGAATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACTCAACAAGTCTCAGT
GTGCTGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCAGCATCCGCTAACTCTGCT
CGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTTTTGTCTAAAAAAT
TTCGTAATCGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCTGTATTACTGCATTT
ATTAAGAGTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCNTCTAAAATATACTTCAG
TCACCTCACT

pAG_18

NCCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCG
GTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATG
CGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCT
CCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACAGTGCTCCGAGAAC
GGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCCATAGTGACTGGCGAT
GCTGTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACATGAAAACGCTTTAACA
GAAAGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCAAGTTATCATCGGCAATG
TTGAATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACTCAACAAGTCTCAGTGTG
CTGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACTACAGTTCCGTAATGAGGCTGTT
TTGGGACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTTTTGTCTAAAAAATTTCG
TAATCGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCTGTATTACTGCATTTATTA
AGAGTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCNTCTAAAATAT

pAG_19

CCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGT
GATGCCGGCCACGATGCGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGC
GTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGG
ACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGC
CATAGTGACTGGCGATGCTGTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACA
TGAAAACGCTTTAACAGAAAGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCA
AGTTATCATCGGCAATGTTGAATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACT
CAACAAGTCTCAGTGTGCTGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAAGTGAAA
CAGTCACACTTACCGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTT
TTGTCTAAAAAATTTCGTAATCGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCT
GTATTACTGCATTTATTAAGAGTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCATC

pAG_20

CCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGT
GATGCCGGCCACGATGCGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGC
GTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGG
ACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGC
CATAGTGACTGGCGATGCTGTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACA
TGAAAACGCTTTAACAGAAAGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCA
AGTTATCATCGGCAATGTTGAATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACT
CAACAAGTCTCAGTGTGCTGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAAGATATT
GCGGAAGCGGAGATCGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTAT
TTTTGTCTAAAAAATTTCGTAATCGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCC
CTGTATTACTGCATTTATTAAGAGTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCATC

pAG_21

GCCTGCCNCCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTC
CCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGC
CACGATGCGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGAT
AGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACAGTGCTC

CGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCCATAGTGAC
TGGCGATGCTGTCGGAATGGACGATCACACTACTCTTCTTTTGCCTATTATAACATGAAAACGC
TTTAACAGAAAGAATAGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCAAGTTATCATC
GGCAATGTTGAATGGAGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACTCAACAAGTCT
CAGTGTGCTGAAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTTCCGCAATATCTGG
ATTAGTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTTTTGTCTAAAA
AATTTCGTAATCGCACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCTGTATTACTGC
ATTTATTAAGAGTATTATACCATATTTTTAGTTATTAAGAAATAATCTTCNTCTAAAATATACTT
CAGT

pAG_22

CACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTC
GGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGC
GTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAG
CGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACAGTGCTCCGAGAACGGGTGCGC
ATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCCATAGTGACTGGCGATGCTGTCGG
AATGGACGATCACACTACTCTTCTTTTGCCTATTATAACATGAAAACGCTTTAACAGAAAGAAT
AGGAAGGTATCCGACTGCTGGTATTAACCCTCTTTCTCAAGTTATCATCGGCAATGTTGAATGG
AGTCCATTCAAAACAGCATAGCTCTAAAACATGGAACTCAACAAGTCTCAGTGTGCTGAAGTT
TTGGGACCATTCAAAACAGCATAGCTCTAAAACAATACACTCCAACGGTCTCGGTTTTGGGAC
CATTCAAAACAGCATAGCTCTAAAACCTCGTAGACTATTTTTGTCTAAAAAATTTCGTAATCGC
ACTATTTGTCTCAGCTAGACTTCAGTCTTGAAAAGCCCCTGTATTACTGCATTTATTAAGAGTA
TTATACCATATTTTTAGTTATTAAGAAATAATCTTCNTCTAAAATATACTTCAGTCACCTCCT

pAG_25

CTTAGTACCTTGAATATTACCATCAGTGGCCACAACGGCCTGACCTACGTGGAAGTGTTTGTTA
ACTACTGCATAGTCCTGCTTGAGTGTTAAGGTCGTACCATGTACATAGGAGTGGAAGGTACAG
TCATTGTTGTTATCTGACCCTCTACCAATGTACCAGTTATTCCTGTTACCATCTTTAGATAGAAT
ATAGTGCGCCTTGTCGGATGCGGACTCTAATTGAAGAACATGCCCATTCTGCAATTTTAGCTGA
CCGGTCATGGCTCCACCTTGGGTGATGATCTGTCCGTTTACTACTAACTGGATGTTACCGCCTA
TCGCTGTTGTCCACTCTCTACGTTGAACATACATCAGCCATCCGTTCTCATCTCCCCACTCCATG
ACAGAGTAACGATTAGTGACACCACCAAAGAACTGTTGATACTGGCCACAGTCAAAACCGTTT
GTGGTCATGTAAAGGCGCCCGTTAGCGTGAATATTTCCTTTCCAGTACACATTGGTTCCATCTA
CCTTATCAATTGCACCAGCCAATCCATTGTAATTTTCCAGCTTGTCTGCCTCACGTTCCGCACGG
TCTGCTTGCTGTTCTGCCAAATGAGCAGAGTTAGCGGATGCTGTGGCAGAGTTCTCAGCGTTTA
CCTCAGATTGATGCGCAGCGGAAGCAGAGTTCCCAGCAGATGTAGCGTATTGACCAGCCGTAT
TCTTGAACCGCTTGGCTTCGTCTCGGAAACCCTTGGTCTCATCGCGCCACTGCTTTGTGTTCGTA
GCGTTGGTACTGGACTCGTTCTTAAAGCCCTCCGCTTGGTTTCTGAAAGTCTCAGCCTCATTAC
GGAACTGTAAGGCTTCATTACGTGCTTGCCATGAGTTCTGGTTCATGGTCTTTAGTTGACCAAA
CGGAACAGCATCGCGGTCATCCACGGCGTTCGCTAGGTTCACAATTCGACGACCACGAGCATC
CAAGTGA

pAG_26

CGACCGCGCGGCATGACGCTGGTTGCCGGCGCGTTTTGGGCAGACATCTATCTGCTAGGCGTC
AACCACCTGACCGATGGCACCAGCAAATACAACGTGACAATTGCAGATGGTAGTGCATCACCT
AAGAAATCTACCAAGTTCGGTGGAGACGGCAGCGCGGCCTACAGTGACGGAGCTTGGTACAA
CTTCGCTGAGGTCATGACTCATCACGGTAAGCGCCTGCCTAACTACAACGAATTCCAGGCGCT
GGCTTTCGGCACGACCGAGGCTACGTCCAGCGGCGGCACCGACGTGCCCACCACCGGCGTGAA
CGGCACGGGCGCAACGAGCGCGTGGAACATCTTCACGTCCAAGTGGGGCGTTGTGCAGGCGTC
CGGTTGCTTGTGGACGTGGGGTAACGAGTTCGGCGGCGTGAATGGCGCATCCGAATACACGGC
CAACACTGGCGGCAGAGGATCGGTGTACGCCCAGCCCGCTGCTGCGCTATTCGGCGGCGCCTG
GAACGGCACGTCGCTCTCGGGTCTCGCGCTGCGCTCTGGTACAGCGGGCCGTCGTTCTCGTTCG
CGTTCTTCGGGGCGCGCGGCGTCTGTGACCACCTGATTCTTGAGTAGCGGGGCCGAAAGGCCC
CGCCAAGGCAACCGGATTTACTTTAAGGAGGTCAAATGAGCGATAAAATCATTCACCTGACCG
ATGACTCTTTTGATACCGACGTGCTGAAAGCTGATGGTGCAATTCTGGTTGATTTCTGGGCAGA
GTGGTGCGGCCCTTGCAAAATGATCGCTCCAATCCTGGACGAAATTGCGGACGAATATCAGGG
TAAGCTGACTGTGGCCAAACTGAACATTGACCAGAACCCTGGCACCGCACCGAAATACGGTAT
CCGTGGCATCCCAACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCAGCAACCAAAGTAGGCGC
TCTGTCTAAAGGCCAACTGAAAGAGTTCCTGGACGCCAACCTGGCTTAATGATACTCGAACCC
CTAGCCCGCTCTTATCGGGCGGCTAAGGGGTTTTTTGTCNAGCTAGCTAGGGATCCTCTNNAGT
C

pAG_27

CTAGGGGTTCGAGTACGGTTGCCTTGGCGGGGCCTTTCGGCCCCGCTACNNNNNAATCAGGTG
GTCACAGACGCCGCGCGCCCCGAAGAACGCGAACGAGAACGACGGCCCGCTGTACCAGAGCG
CAGCGCGAGAACCCGAGAGCGACGTGCCGTTCCAGGCGCCGCCGAATAGCGCAGCAGCGGGC
TGGGCGTACACCGATCCTCTGCCGCCAGTGTTGGCCGTGTATTCGGATGCGCCATTCACGCCGC
CGAACTCGTTACCCCACGTCCACAAGCAACCGGACGCCTGCACAACGCCCCACTTGGACGTGA
AGATGTTCCACGCGCTCGTTGCGCCCGTGCCGTTCACGCCGGTGGTGGGCACGTCGGTGCCGCC
GCTGGACGTAGCCTCGGTCGTGCCGAAAGCCAGCGCCTGGAATTCGTTGTAGTTAGGCAGGCG
CTTACCGTGATGAGTCATGACCTCAGCGAAGTTGTACCAAGCTCCGTCACTGTAGGCCGCGCTG
CCGTCTCCACCGAACTTGGTAGATTTCTTAGGTGATGCACTACCATCTGCAATTGTCACGTTGT
ATTTGCTGGTGCCATCGGTCAGGTGGTTGACGCCTAGCANATAGATGTCTGCCCAAAACGCGC
CGGCAACCAGCGTCNTGCCGCGCGGGTCGAGCGCAGCCGGGCGAAACTTCTTCNCAACGAAGC
TGTCACGTAGGTAAGCATCCAGCCATTTACCTCCCCACTTAGNACCTTGAATATTACCATCAGT

pAG_50

AGCTAGTGGAAGTGGTTTCGAAGAAAGCGCCGTGGCCCGCCAATTCTGTCCTGTCCTTATCGAC
GCATTACGCAAAGTCGGGGCAACCGTGACGGATTGTACAGATAACGTGAGCACCACGCAATCA
GCAAATCTTACCCGTCTTATCAACCTTTGCAACGCAGTCAACCGCGATTTAGACATCAGTCTGC
ACTTCAATAGTTCGGATGACCCAGCCGCCACGGGGGTCGAAGTCTTATATTACGATCAACGTG
CGCTGGCGGCTAAAGTGTCCGACGCAATCTCAAAAGCCAGTGGGCTGCGTGACCGCGGCCCCA
AGGAGCGTAAGGATCTGGCAGTATTAAATGGGACAAAGGCACCCGCAATCCTTATTGAGCTTG
CATTCATCAGTAACGCATCTGATATGAAAAAGTTCTTCGACAATATGCAGGCCATCGCGAATG
CCATCGTCCAAGCTGTAACCGGTAAGGCAGTTAACGTGGACCCACCGGCGCCCAAGCGTACTG
CTACTATTATTACAACCGGCGGATTAGGACAGGACGCGGCCAAACAGGCGGTTGACATGATGT
TCGCAAAAGGATGGTACGGGACCCTGACTTTTCAGACCGACGGGATTGCCGTGTTGAAGACAG
GTGGGCTGTCAGGCGACAAATTGCAGGCCGCTCGCGAGTATTTCGCGTGGCGCGGTTGGTGGT
ATCAGGAGGAAGTAATCGAATACTAAGCTAGCAGTGTGTTTTCGCAAGCTTTAGCTCAGGGTA
TGTGTGCATATTTAACATTTGTGTTGTTATCTAATGATTAAGAAAATCTTGGGCTATTCATTAGC
CCTTGCTACTTTATTAGTAGCACTATATTATGGAGTAATGTTCGGATTAATTCAAGTCGTGCTTT
TCATTTCTGATGTTATTATGGCACTACATTCACGGTACCTGGAGCACAAGACTGGCCTCATGGG
CCTTCCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAACATGGTCA
TAGCTGTTTCCTTGCGTATTGGGCGCTCTCCGCTTCCTCGCTC

pAG_51

ATTCCATGTTCTAATACCACAACTAACCGAGGAAGTAAACAACTTTTTATCGTTTTGGCTAGCA
GTGTGTTTTCGCAAGCTTTAGCTCAGGGTATGTGTGCATATTTAACATTTGTGTTGTTATCTAAT
GATTAAGAAAATCTTGGGCTATTCATTAGCCCTTGCTACTTTATTAGTAGCACTATATTATGGA
GTAATGTTCGGATTAATTCAAGTCGTGCTTTTCATTTCTGATGTTATTATGGCACTACATTCACG
GTACCTGGAGCACAAGACTGGCCTCATGGGCCTTCCGCTCACTGCCCGCTTTCCAGTCGGGAA
ACCTGTCGTGCCAGCTGCATTAACATGGTCATAGCTGTTTCCTTGCGTATTGGGCGCTCTCCGC
TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGGTAAAGCCTGGGGTGCCTAATGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCC
GCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTtCTCATAGCTCACGC
TGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG
TTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGA
CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAagAaCAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGA

pAG_52

TTCTGAaGAGTGGTGCTTGCACTTCCGCagACGGTGTGGTGAACGAATATCAGTACAATAAATC
CTTGGCCCCCGTTTTGGCAGATACGTTCCGTAAGGAGGGCCACAAGGTCGATGTAATTATCTGT
CCCGAAAAACAGTTTAAGACAAAAAATGAAGAAAAGTCATACAAAATCCCGCGCGTTAATAG
TGGGGGTTATGATCTGTTGATCGAGTTGCACTTGAATGCCTCTAACGGACAAGGAAAGGGCTC
GGAGGTGCTTTATTACTCGAATAAGGGTTTGGAGTATGCGACCCGTATTTGCGATAAATTAGG
AACCGTTTTCAAGAACCGTGGGGCGAAATTGGACAAACGCCTGTATATCCTTAACTCAAGCAA
GCCTACCGCTGTTCTGATCGAAAGTTTCTTTTGCGACAATAAGGAGGATTATGACAAAGCAAA
GAAACTGGGGCACGAGGGGATCGCCAAACTTATCGTGGAGGGGGTGTTGAATAAGAACATCA
ATAATGAAGGAGTCAAGCAGATGTATAAACACACAATTGTATACGATGGCGAGGTTGACAAAT
TAGCAACTGTCGTCGGGTGGGGATACAATGATGGGAAAATCCTGATTTGCGACATCAAGGATT
ATGTGCCCGGACAAACACAGAACCTTTACGTTGTTGGAGGAGGAGCTTGCGAAAAAaTCAGTA
GTATTACGAAGGAGAAATTTATCATGATTAAAGGGAATGATCGTTTCGACACTCTTTATAAAG
CGCTGGATTTTACAACCGCTAGGCTAGCAGTGTGTTTTCGCAAGCTTTAGCTCAGGGTATGTGT
GCATATTTAACATTTGTGTTGTTATCTAATGATTAAGAAAATCTTGGGCTATTCATTAGCCCTTG
CTACTTTATTAGTAGCACTATATTATGGAGTAATGTTCGGATTAATTCAAGTCGTGCTTTTCATT
TCTGATGTtATTATGGCACTACATTCACGGTACCTGgAGCACAAGACTGGCCTCATGGGCCTTCC
GCTCACTgCCcGCTTtCCAGTCGGgAAaCCTGTCGTGCCAGCTGCATTAACATGGTCATAGCTGTT
tCCTTGCGTaTTGGGCGCTCtccGCTTCCTCGC

pAG_53

GTGGTTTCGAAGAAAGCGCCGTGGCCCGCCAATTCTGTCCTGTCCTTATCGACGCATTACGCAA
AGTCGGGGCAACCGTGACGGATTGTACAGATAACGTGAGCACCACGCAATCAGCAAATCTTAC
CCGTCTTATCAACCTTTGCAACGCAGTCAACCGCGATTTAGACATCAGTCTGCACTTCAATAGT
TCGGATGACCCAGCCGCCACGGGGGTCGAAGTCTTATATTACGATCAACGTGCGCTGGCGGCT
AAAGTGTCCGACGCAATCTCAAAAGCCAGTGGGCTGCGTGACCGCGGCCCCAAGGAGCGTAA
GGATCTGGCAGTATTAAATGGGACAAAGGCACCCGCAATCCTTATTGAGCTTGCATTCATCAG
TAACGCATCTGATATGAAAAAGTTCTTCGACAATATGCAGGCCATCGCGAATGCCATCGTCCA
AGCTGTAACCGGTAAGGCAGTTAACGTGGACCCACCGGCGCCCAAGCGTACTGCTACTATTAT
TACAACCGGCGGATTAGGACAGGACGCGGCCAAACAGGCGGTTGACATGATGTTCGCAAAAG
GATGGTACGGGACCCTGACTTTTCAGACCGACGGGATTGCCGTGTTGAAGAC

Appendix 5

Representation of post-hoc analysis of the wash step

Control Phage	Comparison Phage	Mean Difference	p-value	95% Confidence Interval	
				Lower Bound	Upper Bound
T7	HP-1	-1398500.00*	.011	-2451880.94	-345119.07
	HP-2	-1715166.67*	.002	-2768547.60	-661785.73
	HP-3	-3548500.00*	.000	-4601880.94	-2495119.07
	HP-4	-4065166.67*	.000	-5118547.60	-3011785.73
	HP-5	-2415166.67*	.000	-3468547.60	-1361785.73
	HP-6	-865166.67	.104	-1918547.60	188214.27
HP-1	HP-2	-316666.67	.543	-1370047.60	736714.27
	HP-3	-2150000.00*	.000	-3203380.94	-1096619.07
	HP-4	-2666666.67*	.000	-3720047.60	-1613285.73
	HP-5	-1016666.67	.058	-2070047.60	36714.27
	HP-6	533333.33	.309	-520047.60	1586714.27
HP-2	HP-3	-1833333.33*	.001	-2886714.27	-779952.40
	HP-4	-2350000.00*	.000	-3403380.94	-1296619.07
	HP-5	-700000.00	.184	-1753380.94	353380.93
	HP-6	850000.00	.110	-203380.93	1903380.94
HP-3	HP-4	-516666.67	.324	-1570047.60	536714.27
	HP-5	1133333.33*	.036	79952.40	2186714.27
	HP-6	2683333.33*	.000	1629952.40	3736714.27
HP-4	HP-5	1650000.00*	.003	596619.07	2703380.94
	HP-6	3200000.00*	.000	2146619.07	4253380.94
HP-5	HP-6	1550000.00*	.005	496619.07	2603380.94

Representation of post-hoc analysis of the lysis step

Control Phage	Comparison Phage	Mean Difference	p-value	95% Confidence Interval	
				Lower Bound	Upper Bound
T7	HP1	-65416.67*	.000	-74026.73	-56806.60
	HP2	-42750.00*	.000	-51360.06	-34139.94
	HP3	-32250.00*	.000	-40860.06	-23639.94
	HP4	-10800.00*	.016	-19410.06	-2189.94
	HP5	-13250.00*	.004	-21860.06	-4639.94
	HP6	-10916.67*	.015	-19526.73	-2306.60
HP1	HP2	22666.67*	.000	14056.60	31276.73
	HP3	33166.67*	.000	24556.60	41776.73
	HP4	54616.67*	.000	46006.60	63226.73
	HP5	52166.67*	.000	43556.60	60776.73
	HP6	54500.00*	.000	45889.94	63110.06
HP2	HP3	10500.00*	.019	1889.94	19110.06
	HP4	31950.00*	.000	23339.94	40560.06
	HP5	29500.00*	.000	20889.94	38110.06
	HP6	31833.33*	.000	23223.27	40443.40
HP3	HP4	21450.00*	.000	12839.94	30060.06
	HP5	19000.00*	.000	10389.94	27610.06
	HP6	21333.33*	.000	12723.27	29943.40
HP4	HP5	-2450.00	.565	-11060.06	6160.06
	HP6	-116.67	.978	-8726.73	8493.40
HP5	HP6	2333.33	.583	-6276.73	10943.40

Summary of wash step t-tests, relationship between incubation time and type of phage.

Equal variances assumed

		Levene's Test for Equality of Variances				
Phage_No		F	Sig.	t	df	p-value
1	W	4.030	0.115	-2.476	4	0.068
				-2.476	2.026	0.130
2	W	4.000	0.116	1.732	4	0.158
				1.732	2.000	0.225
3	W	11.253	0.028	-12.583	4	0.000
				-12.583	2.000	0.006
4	W	0.396	0.563	-2.568	4	0.062
				-2.568	3.604	0.069
5	W	1.665	0.266	2.755	4	0.051
				2.755	2.669	0.080
6	W	3.826	0.122	-4.427	4	0.011
				-4.427	2.427	0.033
7	W	0.216	0.667	2.822	4	0.048
				2.822	3.931	0.049

Summary of t-tests for the lysis step for the relationship between time and type of phage.

	Levene's Test for Equality of Variances				
Phage_No	F	Sig.	t	df	Sig. (2-tailed)
T7	13.892	0.020	0.387	4	0.719
			0.387	2.015	0.736
HP1	14.530	0.019	-7.834	4	0.001
			-7.834	2.000	0.016
3	0.746	0.437	-3.268	4	0.031
			-3.268	3.571	0.036
4	5.918	0.072	-1.273	4	0.272
			-1.273	2.025	0.330
5	15.747	0.017	-3.836	4	0.019
			-3.836	2.000	0.062
6	6.897	0.058	-7.667	4	0.002
			-7.667	2.278	0.011
7	6.606	0.062	-4.152	4	0.014
			-4.152	2.358	0.040

Wash for time 1 and time 2

Time			Mean Difference (I-J)	Std. Error	Sig.
1	HP_3	HP_4	1900000.00*	873602.267	0.047
		HP_5	-866666.67	873602.267	0.338
		HP_6	-4333333.33*	873602.267	0.000
		HP_7	733333.33	873602.267	0.415
		HP_8	533333.33	873602.267	0.551
		T7	1970333.33*	873602.267	0.041
	HP_4	HP_3	-1900000.00*	873602.267	0.047
		HP_5	-2766666.67*	873602.267	0.007
		HP_6	-6233333.33*	873602.267	0.000
		HP_7	-1166666.67	873602.267	0.203
		HP_8	-1366666.67	873602.267	0.140
		T7	70333.33	873602.267	0.937
	HP_5	HP_3	866666.67	873602.267	0.338
		HP_4	2766666.67*	873602.267	0.007
		HP_6	-3466666.67*	873602.267	0.001
		HP_7	1600000.00	873602.267	0.088
		HP_8	1400000.00	873602.267	0.131
		T7	2837000.00*	873602.267	0.006
	HP_6	HP_3	4333333.33*	873602.267	0.000
		HP_4	6233333.33*	873602.267	0.000
		HP_5	3466666.67*	873602.267	0.001
		HP_7	5066666.67*	873602.267	0.000
		HP_8	4866666.67*	873602.267	0.000
		T7	6303666.67*	873602.267	0.000
	HP_7	HP_3	-733333.33	873602.267	0.415
		HP_4	1166666.67	873602.267	0.203
		HP_5	-1600000.00	873602.267	0.088

		HP_6	-5066666.67*	873602.267	0.000
		HP_8	-200000.00	873602.267	0.822
		T7	1237000.00	873602.267	0.179
	HP_8	HP_3	-533333.33	873602.267	0.551
		HP_4	1366666.67	873602.267	0.140
		HP_5	-1400000.00	873602.267	0.131
		HP_6	-4866666.67*	873602.267	0.000
		HP_7	200000.00	873602.267	0.822
		T7	1437000.00	873602.267	0.122
	T7	HP_3	-1970333.33*	873602.267	0.041
		HP_4	-70333.33	873602.267	0.937
		HP_5	-2837000.00*	873602.267	0.006
		HP_6	-6303666.67*	873602.267	0.000
		HP_7	-1237000.00	873602.267	0.179
		HP_8	-1437000.00	873602.267	0.122
2	HP_3	HP_4	-2533333.33*	542776.519	0.000
		HP_5	-3433333.33*	542776.519	0.000
		HP_6	-1000000.00	542776.519	0.087
		HP_7	-2766666.67*	542776.519	0.000
		HP_8	533333.33	542776.519	0.342
		T7	826666.67	542776.519	0.150
	HP_4	HP_3	2533333.33*	542776.519	0.000
		HP_5	-900000.00	542776.519	0.120
		HP_6	1533333.33*	542776.519	0.013
		HP_7	-233333.33	542776.519	0.674
		HP_8	3066666.67*	542776.519	0.000
		T7	3360000.00*	542776.519	0.000
	HP_5	HP_3	3433333.33*	542776.519	0.000
		HP_4	900000.00	542776.519	0.120
		HP_6	2433333.33*	542776.519	0.001

	HP_7	666666.67	542776.519	0.240
	HP_8	3966666.67*	542776.519	0.000
	T7	4260000.00*	542776.519	0.000
HP_6	HP_3	1000000.00	542776.519	0.087
	HP_4	-1533333.33*	542776.519	0.013
	HP_5	-2433333.33*	542776.519	0.001
	HP_7	-1766666.67*	542776.519	0.006
	HP_8	1533333.33*	542776.519	0.013
	T7	1826666.67*	542776.519	0.005
HP_7	HP_3	2766666.67*	542776.519	0.000
	HP_4	233333.33	542776.519	0.674
	HP_5	-666666.67	542776.519	0.240
	HP_6	1766666.67*	542776.519	0.006
	HP_8	3300000.00*	542776.519	0.000
	T7	3593333.33*	542776.519	0.000
HP_8	HP_3	-533333.33	542776.519	0.342
	HP_4	-3066666.67*	542776.519	0.000
	HP_5	-3966666.67*	542776.519	0.000
	HP_6	-1533333.33*	542776.519	0.013
	HP_7	-3300000.00*	542776.519	0.000
	T7	293333.33	542776.519	0.597
T7	HP_3	-826666.67	542776.519	0.150
	HP_4	-3360000.00*	542776.519	0.000
	HP_5	-4260000.00*	542776.519	0.000
	HP_6	-1826666.67*	542776.519	0.005
	HP_7	-3593333.33*	542776.519	0.000
	HP_8	-293333.33	542776.519	0.597

Lysis both lysis time 1 and 2

Time			Mean Difference (I-J)	Std. Error	Sig.
1	HP_3	HP_4	2666.67	5592.058	0.641
		HP_5	5666.67	5592.058	0.328
		HP_6	35566.67*	5592.058	0.000
		HP_7	28666.67*	5592.058	0.000
		HP_8	30666.67*	5592.058	0.000
		T7	32933.33*	5592.058	0.000
	HP_4	HP_3	-2666.67	5592.058	0.641
		HP_5	3000.00	5592.058	0.600
		HP_6	32900.00*	5592.058	0.000
		HP_7	26000.00*	5592.058	0.000
		HP_8	28000.00*	5592.058	0.000
		T7	30266.67*	5592.058	0.000
	HP_5	HP_3	-5666.67	5592.058	0.328
		HP_4	-3000.00	5592.058	0.600
		HP_6	29900.00*	5592.058	0.000
		HP_7	23000.00*	5592.058	0.001
		HP_8	25000.00*	5592.058	0.001
		T7	27266.67*	5592.058	0.000
	HP_6	HP_3	-35566.67*	5592.058	0.000
		HP_4	-32900.00*	5592.058	0.000
		HP_5	-29900.00*	5592.058	0.000
		HP_7	-6900.00	5592.058	0.238
		HP_8	-4900.00	5592.058	0.396
		T7	-2633.33	5592.058	0.645
	HP_7	HP_3	-28666.67*	5592.058	0.000

		HP_4	-26000.00*	5592.058	0.000
		HP_5	-23000.00*	5592.058	0.001
		HP_6	6900.00	5592.058	0.238
		HP_8	2000.00	5592.058	0.726
		T7	4266.67	5592.058	0.458
	HP_8	HP_3	-30666.67*	5592.058	0.000
		HP_4	-28000.00*	5592.058	0.000
		HP_5	-25000.00*	5592.058	0.001
		HP_6	4900.00	5592.058	0.396
		HP_7	-2000.00	5592.058	0.726
		T7	2266.67	5592.058	0.691
	T7	HP_3	-32933.33*	5592.058	0.000
		HP_4	-30266.67*	5592.058	0.000
		HP_5	-27266.67*	5592.058	0.000
		HP_6	2633.33	5592.058	0.645
		HP_7	-4266.67	5592.058	0.458
		HP_8	-2266.67	5592.058	0.691
2	HP_3	HP_4	42666.67*	6276.917	0.000
		HP_5	60666.67*	6276.917	0.000
		HP_6	73666.67*	6276.917	0.000
		HP_7	75666.67*	6276.917	0.000
		HP_8	78333.33*	6276.917	0.000
		T7	97900.00*	6276.917	0.000
	HP_4	HP_3	-42666.67*	6276.917	0.000
		HP_5	18000.00*	6276.917	0.012
		HP_6	31000.00*	6276.917	0.000
		HP_7	33000.00*	6276.917	0.000
		HP_8	35666.67*	6276.917	0.000
		T7	55233.33*	6276.917	0.000

HP_5	HP_3	-60666.67*	6276.917	0.000
	HP_4	-18000.00*	6276.917	0.012
	HP_6	13000.00	6276.917	0.057
	HP_7	15000.00*	6276.917	0.031
	HP_8	17666.67*	6276.917	0.014
	T7	37233.33*	6276.917	0.000
HP_6	HP_3	-73666.67*	6276.917	0.000
	HP_4	-31000.00*	6276.917	0.000
	HP_5	-13000.00	6276.917	0.057
	HP_7	2000.00	6276.917	0.755
	HP_8	4666.67	6276.917	0.469
	T7	24233.33*	6276.917	0.002
HP_7	HP_3	-75666.67*	6276.917	0.000
	HP_4	-33000.00*	6276.917	0.000
	HP_5	-15000.00*	6276.917	0.031
	HP_6	-2000.00	6276.917	0.755
	HP_8	2666.67	6276.917	0.677
	T7	22233.33*	6276.917	0.003
HP_8	HP_3	-78333.33*	6276.917	0.000
	HP_4	-35666.67*	6276.917	0.000
	HP_5	-17666.67*	6276.917	0.014
	HP_6	-4666.67	6276.917	0.469
	HP_7	-2666.67	6276.917	0.677
	T7	19566.67*	6276.917	0.008
T7	HP_3	-97900.00*	6276.917	0.000
	HP_4	-55233.33*	6276.917	0.000
	HP_5	-37233.33*	6276.917	0.000
	HP_6	-24233.33*	6276.917	0.002
	HP_7	-22233.33*	6276.917	0.003

HP_8	-19566.67*	6276.917	0.008